

The Implementation and Use of Benchmarking in Local Government: A Case Study of the Translation of a Management Accounting Innovation

SVEN SIVERBO*

Abstract: Benchmarking is a management accounting innovation (MAI) that can be used for performance measurement and management in both the private and the public sectors. Although public sector accounting researchers have reported some success with the use of benchmarking, frequently charged problems exist in implementing and using this management technique. To look beyond the technical and institutional explanations, this paper takes a translation approach and presents a case study of a local government benchmarking network. We conclude that there is a link between benchmarking *implementation* problems and initiators' failure to build a strong network of benchmarking allies. Implementation is facilitated if actors, other than the initiators, recognize the possibility of making benchmarking more relevant and less cost focused. However, even when a network of actors has a favourable attitude towards benchmarking, benchmarking may still appear as an unruly 'actant'. Furthermore, the perception of implementation failure and success is heterogeneous and connected to various actors' adoption of benchmarking. We also conclude that there is a connection between the *use* of benchmarking and 1) actors' possibilities to use benchmarking in the struggle for resources and 2) the perception of benchmarking information as 'factual' or 'factual enough'. However, the perception of benchmarking information as 'factual' or 'factual enough' seems not only a matter of correct or incorrect ratios but also of whether such information serves actors' interests. A final conclusion is that the use of benchmarking increases when actors other than the initiators complement the original idea and 'counter interest' the initiators.

Keywords: benchmarking, management accounting innovation, public sector, local government, translation, use, implementation

*The author is from Karlstad Business School at Karlstad University.

Address for correspondence: Sven Siverbo, Karlstad Business School at Karlstad University, 651 88 Karlstad, Sweden.
e-mail: sven.siverbo@kau.se

INTRODUCTION

Benchmarking is a management accounting innovation (MAI) that can be used for performance measurement and improvement in both the private and public sectors. It can be used to compare performance (e.g., costs, productivity or results) and processes. In the private sector, benchmarking is one of the most adopted and popular management techniques (Rigby and Bilodeau, 2007). In the public sector, interest in benchmarking in the last decade has greatly increased in several countries. One aspect of this interest is the external use of benchmarking, that is, organizations' use of various ratios to compare their units' performance with similar units in other organizations (Northcott and Llewellyn, 2003). The United Kingdom is at the forefront of this movement. The British government has used benchmarking since the end of the 1990s to control and stimulate performance improvement at the local government level and in health care organizations (Broadbent et al., 1999; Bowerman and Ball, 2000; Bowerman et al., 2002; Jackson and Lapsley, 2003; and Northcott and Llewellyn, 2005). Public authorities in many other countries are also involved in benchmarking activities. Among these countries are the United States (Dorsch and Yasin, 1998; and Cavalluzzo and Ittner, 2004), New Zealand (Broadbent et al., 1999), the Netherlands (de Bruijn and van Helden, 2006; and ter Bogt, 2008), Norway (Johnsen, 2007; and Askim et al., 2008), Denmark (Triantafillou, 2007) and Sweden (Siverbo and Johansson, 2006; and Johansson and Siverbo, 2009).

Although some researchers have reported success by users of benchmarking and similar techniques (see Bowerman and Ball, 2000; Lapsley and Wright, 2004; Llewellyn and Northcott, 2005; Siverbo and Johansson, 2006; Askim et al., 2008; and Johnsen, 2007), other researchers have frequently observed problems in their implementation and use (see Bowerman and Ball, 2000; Ammons et al., 2001; Northcott and Llewellyn, 2003; Askim, 2004; Northcott and Llewellyn, 2005; Llewellyn and Northcott, 2005; Johnsen, 2005; Pollit, 2006; Östergren, 2006; Triantafillou, 2007; Askim et al., 2008; and Johansson and Siverbo, 2009). This situation is similar to that facing other MAIs in the public sector (see, for example, Preston et al., 1992; Chua, 1995; Geiger and Ittner, 1996; de Lancer Julnes and Holtzer, 2001; Cavalluzzo and Ittner, 2004; ter Bogt, 2004; and Arnaboldi and Azzone, 2010) and MAIs in general (see, for example, Gosselin, 1997; Langfield-Smith, 2008; and Otley, 2008).

In previous research on benchmarking in the public sector, problems with the use of this technique are largely explained as technical problems, as poor implementation strategies and even as commitment problems (see a literature review in Siverbo and Johansson, 2006). Arnaboldi and Azzone (2010) present similar reasoning about performance measurement systems in the public sector. In general, this research, which takes an upper management perspective, reveals a lack of interest in social and political factors and in how the MAI unfolds over time (cf. Modell, 2007; and Arnaboldi and Azzone, 2010). To some extent, New Institutional Sociological (NIS) theory has been used to explain

the implementation and use of performance measurement and management. Although several studies have supported the NIS explanation (e.g., Geiger and Ittner, 1996; Gosselin, 1997; de Lancer Julnes and Holzer, 2001; Cavalluzzo and Ittner, 2004; van Dooren, 2005; and Johansson and Siverbo, 2009), the theory has been criticized for its silence on micro dynamics, that is, the interplay between different institutional logics, rationality, agency and power (Modell, 2009).

Therefore, in this paper we use a translation approach to increase the understanding of the implementation and use of benchmarking in public sector organizations (PSOs). On a general level, a translation approach means that a critical and sociological perspective is the point of departure in the research (Baxter and Chua, 2003). An axiom in this research is that the popularity of an innovation is less the result of its technical characteristics or contextual conditions and more the result of the translation process that involves a complex and unpredictable interaction between the innovation and the actors who have changeable interests and identities (Briers and Chua, 2001; Baxter and Chua, 2003; and Alcouffe et al., 2008).

In this paper, we view the introduction of a MAI not as a neutral activity but rather as an attempt by an actor or actor group to maintain or improve organizational position. A distinguishing feature of MAIs is that, by their creation of inscriptions, they have the potential to change the way organizational reality is constructed (Robson, 1992). Thus MAIs are likely to provoke controversy and cause power struggles that affect their implementation and use (Baxter and Chua, 2003; and Alcouffe et al., 2008). The translation approach provides a useful theoretical 'lens' for explaining interrelated organizational and technological phenomena that occur when MAIs are adopted in PSOs (Papadopoulos et al., 2011). With the translation approach as our starting point, we focus less on the technical efficacy of benchmarking or on the conditions in the implementation process *per se* (the foci in previous research) and focus more on if and how benchmarking is 'made to be good' (Briers and Chua, 2001, p. 266).

In addition to increasing our understanding of benchmarking implementation and use, our approach is a response to the challenge to improve the general understanding of the use of MAIs in PSOs (Lapsley and Pallot, 2000; de Lancer Julnes and Holzer, 2001; and Jackson and Lapsley, 2003) as well as complement earlier MAI studies on budgeting (Preston et al., 1992; and Ezzamel, 1994), DRG-based accounting (Chua, 1995; and Lowe, 2001a and 2001c), ABC (Briers and Chua, 2001; and Alcouffe et al., 2008), Juran's cost of quality technique (Emsley, 2008), standard costing (Lowe and Koh, 2007) and performance measurement (Arnaboldi and Azzone, 2010) that also use the translation approach.

The paper is structured as follows. In the next section we present the theoretical framework that defines relevant core concepts and describes the translation approach. In the third section, we describe our method, and in the fourth section we present and analyse our results. In the last section, we discuss and present our conclusions.

THEORETICAL FRAMEWORK

Benchmarking: Definition and Design Elements

In an extensive literature review, Francis and Holloway (2007, p.184) identify some basic features of benchmarking:

At its most basic, benchmarking can be conceived as a means of sharing good practice, which has been validated in certain organizational contexts and which logic suggests may bring benefits elsewhere.

Moriarty and Smallman (2009, p. 498) also provide some clarification by defining benchmarking as:

an exemplar driven teleological process operating within an organization with the objective of intentionally changing an existing state of affairs into a superior state of affairs.

In order to analyse how benchmarking evolves in a local context, it is necessary to have a rough understanding of the 'original' idea, that is, which design elements (Bjørnenak and Olson, 1999) to expect in benchmarking. Due to the variety of benchmarking concepts, it is difficult to find such a standard or original idea for benchmarking. However, for analytical purposes we need to develop a provisional public sector 'original', that is, a description of public sector benchmarking design elements derived from the literature.

Public Sector Benchmarking

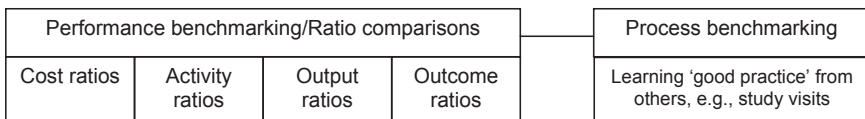
Under normal conditions, PSOs do not compete against each other (however, see Northcott and Llewellyn, 2005). The lack of market pressure and competition provides PSOs with both the opportunity and the motivation to search for best practices among themselves (van Helden and Tillema, 2005; Northcott and Llewellyn, 2005; and Andersen et al., 2008). Since PSOs have no trade secrets, easy access to their information and permission to copy their exemplars are expected (Bowerman et al., 2002). Thus the benchmarking features of *internal benchmarking* (no secrets) and *functional benchmarking* (learning from exemplars) (Moriarty and Smallman, 2009) are associated with PSOs. According to the general understanding among benchmarking researchers, a combination of internal and functional benchmarking results in a two-step benchmarking model that begins with performance benchmarking and is followed by process benchmarking (Elnathan et al., 1996; Wolfram Cox et al., 1997; Bhutta and Huq, 1999; Magd and Curry, 2003; Lee et al., 2006; Francis and Holloway, 2007; Alstete 2008; Andersen et al., 2008; and Moriarty and Smallman, 2009).

Performance benchmarking, which assesses performance, is the activity that potentially identifies a performance gap (the difference between individual achievement and exemplars' achievement). A critical decision is the selection of

measures that must be tailored to the needs of the individual organization or department (Gunasekaran, 2005). For comparisons, these measures are often presented as ratios (Triantafillou, 2007). Obviously, many ratios may be used for PSOs. Furthermore, the opinion of what a good ratio is seems to change over time. In the era of Progressive Public Administration (PPA), the emphasis was on probity. This emphasis was reflected in holding managers accountable for costs and procedures (Hood, 1995; and Broadbent et al., 1999). In that era, the main interest was *cost ratios* (resources committed to an activity) and *activity ratios* (how the activity was performed, for example, by using staff-to-client ratios or formal competences). In the early 1990s, New Public Management (NPM) replaced PPA in many countries. NPM is a more corporate-like management style that, among other things, advocates increased interest in and accountability for outputs and outcomes (Hood, 1995). In this environment, PSOs increasingly complement the traditional cost and activity measurements with output and outcome measurements (Broadbent et al., 1999; Pallot, 2001; Wang and Berman, 2001; Lapsley and Wright, 2004; ter Bogt, 2004 and 2008; Northcott and Llewellyn, 2005; and Hoque, 2008). Based on this reasoning, we can identify four kinds of ratios, or (sub)design elements (see Figure 1).

Figure 1

Expected Design Elements in Public Sector Benchmarking



Process benchmarking is the activity in which organizations learn to narrow their performance gaps through investigation of exemplars' processes, for instance, by study visits (Gunasekaran, 2005; and Alstete, 2008). Benchmarking researchers generally agree that if the benchmarking process ends with performance benchmarking, organizations learn little about improving their operations (see, e.g., Northcott and Llewellyn, 2005; and Alstete, 2008). Some researchers even argue that without process benchmarking aimed at improving activities through learning from others, it is improper to speak of 'true' benchmarking (Alstete, 2008). Several UK researchers who implicitly or explicitly advocate the two-step model have noted the risks and problems with only collecting data and publishing league tables (Longbottom, 2000; Llewellyn and Northcott, 2005; Northcott and Llewellyn, 2005; and Francis and Holloway, 2007). Consequently, we may expect public sector benchmarking to contain the design elements presented in Figure 1.

MAI Adoption, Implementation and Use

Various frameworks and conceptualisations have been developed to describe the stages a MAI must pass through before it can be expected to be effective and before it can be considered as taken-for-granted in an organization (Leseure et al., 2004; and Modell, 2007). To support the analysis in this paper, we distinguish between adoption, implementation and use (cf. Johnsen and Vakkuri, 2006; and Johnsen, 2007). *Adoption* refers to the formal or informal decisions an actor group makes to commence implementation of a MAI, presumably after an encounter with the global idea; *implementation* refers to how the MAI thereafter is conceptualised, designed, developed and tested until it is ready for full-scale use; and *use* refers to how the MAI is used in practice.

As stated in the Introduction, accounting researchers have frequently observed problems in benchmarking implementation and use. Researchers have made the same observation about other MAIs in both the public and the private sectors. In a literature review, Siverbo and Johansson (2006) note that technical problems and poor implementation strategies are major causes of such problems although commitment problems in some degree also inhibit the realization of benchmarking goals. The research on public sector performance measurement systems also presents such technological, functionalistic and implementation strategy-related explanations (Arnaboldi and Azzone, 2010). This research, which is often biased towards an upper management perspective, is silent on how social and political factors affect the evolution of MAIs (cf. Modell, 2007; and Arnaboldi and Azzone, 2010). Calls have been issued for more theoretically informed research that preferably takes non-mainstream and/or critical perspectives (Francis and Holloway, 2007; Moriarty and Smallman, 2009; and Arnaboldi and Azzone, 2010).

Some accounting researchers have answered this call; however, their response has not explicitly focused on benchmarking. An institutional theory perspective has been used to study implementation and use issues related to performance measurement and management. A standard observation in accounting studies based in NIS theory is that organizations sometimes do not follow through on their MAI adoption; instead, they are satisfied merely with a superficial adoption that makes them appear modern and rational (Meyer and Rowan, 1977/1977). Although several studies support the contribution NIS theory makes (e.g., Geiger and Ittner, 1996; Gosselin, 1997; de Lancer Julnes and Holzer, 2001; Cavalluzzo and Ittner, 2004; van Dooren, 2005; and Johansson and Siverbo, 2009), the theory's perspective has been criticized as too limiting. A common criticism is that a perspective based in NIS theory, with its focus on institutional pressures and legitimacy, does not explain the reality in organizations that seemingly decouple adoption from implementation and use. The NIS perspective has problems with micro dynamics, that is, the interplay between different institutional logics, rationality, agency and power (Modell, 2009). Although efforts have been made to fill the gaps in the institutional explanation,

researchers have suggested that NIS should be complemented (Modell, 2009) or substituted (Lounsbury, 2008) by a translation approach. In the following section we describe the translation approach and how it contributes to the understanding of benchmarking implementation and use.

A Translation Approach

In order to understand what happens in organizations that adopt benchmarking, it is important to realize that organizations develop their own versions of MAIs. Latour (1986/1986) states that artefacts – for example, MAIs such as benchmarking – have no diffusion power of their own but live on the energy supplied by users who adapt them to their own needs. MAIs do not spread in unchanged forms among passive users; they are *translated* locally by adaptation to local conditions, interests and preferences. The approach means that a MAI is created in an on-going, complex interaction among structure, actors, organization form and the prevailing situation (cf. Preston et al., 1992; Chua, 1995; Lowe, 2001b; and Mennicken, 2008). In research, this approach has different labels, including the ‘Latourian approach’ (Baxter and Chua, 2003), the ‘Sociology of translation’ (Chua, 1995; and Modell, 2009) and ‘Actor Network Theory (ANT)’ (Briers and Chua, 2001; Lowe, 2001b and 2001c; and Emsley, 2008). Some researchers argue that ANT is a more controversial version of the approach since it recognizes non-human actors, the so-called actants (see e.g., Chua, 1995; Briers and Chua, 2001; and Lowe, 2001b and 2001c). To date, accounting researchers have been quite reluctant to treat actors and actants symmetrically, that is, to treat natural and social objects as equals (Lowe, 2001c).

One important principle of the translation approach is that management accounting has the potential to create organizations’ reality (Preston et al., 1992; Robson, 1992; Chua, 1995; and Lowe, 2001b and 2001c). Accounting visualizes facts or inscriptions with inherent persuasive powers, which, among other things, allows it to control actors from a distance (Preston et al., 1992; Robson, 1992; Chua, 1995; and Lowe, 2001c). In this way, management accounting can influence behaviour, which, according to some researchers, makes it a non-human actor or actant (Briers and Chua, 2001; Lowe, 2001b; and Alcouffe et al., 2008) that is therefore somewhat out of the control of actors (Lowe, 2001c). The translation approach directs the researcher to investigate how management accounting contributes to creating organizations’ reality and how MAIs can change widely held views about reality. The translation approach involves deciding to what extent benchmarking information is perceived as fact, and, consequently, what effect benchmarking has (cf. Preston et al., 1992; Chua, 1995; Briers and Chua, 2001; and Lowe, 2001b). However, this fact building is intimately connected to various actors’ interests:

[I]t is when new accounting technologies can hold diverse ‘facts’ and interests together, stabilise them (temporarily), that these ‘facts’ will start to become ‘true’ (Briers and Chua, 2001, p. 267).

Nevertheless, although powerful interests may influence what is perceived as true, the possibilities for actors to control accounting systems may be limited. According to Lowe (2001c), accounting soon escapes the grasp of actors and starts to produce unintended consequences:

This line of theorizing would look to explain less as a result of human agency and more as a result of the impact of relatively independent and unfettered accounting systems (Lowe, 2001c, p. 85).

In researching accounting this means that accounting systems and techniques need to be considered as equals with the human actors in our studies of organizational interaction. The implication is that rather than looking always for a human or super human interest to explain what we find in organizations, we need to look for more mundane explanations based upon the intimacy of the bonds between people and technology (Lowe, 2001c, p. 94).

Another important principle in the translation approach is that a MAI is neither adopted nor implemented in a rational process. Instead, a MAI is the result of a complex, political process of power struggles among actors with different interests (Preston et al., 1992; Chua, 1995; Briers and Chua, 2001; Lowe, 2001b; and Alcouffe et al., 2008). MAIs are not stable, unchangeable standards but rather mouldable, half-filled packages that actors in organizations fill with content (Røvik, 1998/1998). The success of a MAI depends less on its inherent qualities and more on the outcome of the related negotiations and power struggles among actors with conflicting interests (Preston et al., 1992; Chua, 1995; and Mennicken, 2008). Actors who support the use of a MAI often try to strengthen and consolidate their position in the organization; therefore they have to deal with resistance from other actors (Ezzamel, 1994). Battles and trials of strength arise in which advocates for an innovation must fight various counter actors on various fronts (Preston et al., 1992; and Alcouffe et al., 2008).

Some actors believe that because MAIs are often adjusted or changed in the translation process, they are then distorted. However, this distortion may be necessary since the alternative is that other actors in the organization will disassociate themselves from the MAI (Chua, 1995; and Røvik, 1998/1998).

Using different conceptualizations, several researchers have tried to describe what occurs in the translation process. In this paper, we use the conceptualization originally developed by Callon (1986) in which the translation process is divided into four (overlapping) 'moments': problematization, interessement, enrolment and mobilization. Some accounting researchers (Ezzamel, 1994; Skærbæk and Melander, 2004 and 2007; and Alcouffe et al., 2008) have used this conceptualization although their interpretations differ slightly from Callon's.¹ This paper generally builds on and extends Callon's conceptualization using these interpretations from accounting research that are adapted to the research on MAIs.

Problematization is the moment when an initiating actor's (the initiator) tries to define other actors' problems, objectives and interests, and to make a

MAI (suitable to that actor's interests) indispensable. In the problematization moment, the initiator identifies the relevant actors and demonstrates that it is in their interests to adopt the MAI. The primary aim of initiators, who try to establish associations between themselves and other actors, is to make the MAI an obligatory passage point, that is, a unique and well defined practice that they control and that other actors must follow (Ezzamel, 1994). In benchmarking, this practice means participating in the benchmarking work and using the information it produces.

Interessement is the moment when the initiator tries to attach other actors to the problematization. For the initiator, an important task is to obstruct competing problematizations posed by other initiators. This trial of strength may require force, seduction and/or solicitation. If successful, it confirms the validity of the problematization and of the tentative alliances.

Enrolment is the moment when the alliances are firmly established although a successful interessement may not necessarily create alliances. Enrolment means that actors accept their roles. In benchmarking, this means that the actors accept benchmarking as an obligatory passage point and act in accordance with their roles in the benchmarking process. Callon (1986, p. 211) writes:

Interessement achieves enrolment if it is successful. To describe enrolment is thus to describe the group of multilateral negotiations, trials of strength and tricks that accompany the interessements and enable them to succeed.

This process results in networks of actors with the same interests.

Mobilization is the moment when initiators take on the role of spokesmen for all actors. As spokesmen, they speak and act for these other actors. Mobilization means enrolment is transformed into active support for the initiator and the MAI.

Successful translation means MAIs become obligatory passage points and, eventually, if the actor network is stable, part of the taken-for-granted practices of the organization – the MAI is 'black boxed' (cf. Lowe, 2001b). From the perspective of the translation approach, the most important success criterion is an extensive network of MAI advocates and users (Alcouffe et al., 2008).

In summary, organizations that adopt a MAI such as benchmarking initiate a translation process. The implementation and use of benchmarking is typically a complex and unpredictable process where the benchmarking idea is moulded in the interaction between the idea and the actors who have changeable interests and preferences. Power struggles among actors result because benchmarking is not a neutral technique but rather a technique that primarily serves the interests of the initiators. Benchmarking, with its potential for creating reality, may pose a threat to various actors because comparisons can indicate which actors perform well which do not.

METHOD

We chose the case study method for this study because it is well suited to obtaining an in-depth understanding of how organizations 'experiment' with new accounting systems (Chua, 1995) and to focusing attention on large numbers of variables and their interaction (Lowe, 2001b). Furthermore, case studies enable mapping of organizational processes, which is necessary in research that takes the translation approach (Lowe, 2001b) and which is a required step in benchmarking research (Francis and Holloway, 2007).

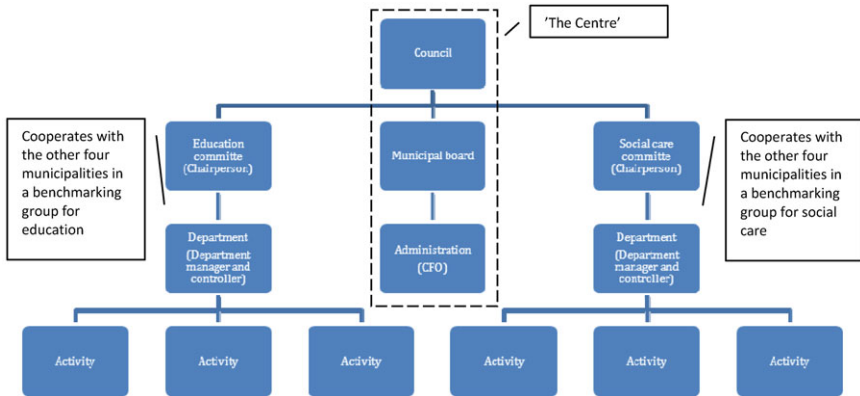
Our case consists of a benchmarking network of five Swedish municipalities in Northern Bohuslän (the Network). All 290 municipalities in Sweden are responsible for important welfare services such as education, social care, community planning, maintenance of local roads and parks, and waste disposal. The largest municipality in Sweden is Stockholm (810,000 inhabitants) and the smallest municipality is Bjurholm with 2,500 inhabitants. The median municipality has 15,400 inhabitants. The five municipalities in the Network have between 9,500 and 12,200 inhabitants, and their annual expenditures are between 55 and 70 million Euro. The directly elected politicians who govern these municipalities make decisions about taxes, fees, service levels, organizations forms and control systems. Accordingly, each municipality decides if and how they want to use benchmarking. Initiatives at the national level have led to the development of national databases that make it easier for municipalities to compare their performance with similar municipalities. However, in accordance with the long-standing tradition of municipal autonomy, the central government encourages, but does not require, the use of benchmarking.

This paper focuses on the two largest benchmarking groups in the Network: the group for social care (eldercare, social welfare, social care and services for the disabled) and the group for education (nine-year compulsory schools and pre-schools). The municipal council for each municipality oversees the committees and departments responsible for social care and education. According to the Municipality Act, these committees and departments are directly accountable to the councils. However, to some degree the committees and departments are also subordinate to the municipality boards that supervise and coordinate all municipal activities. Therefore, we refer to the municipality council and the municipality board – and their administrations – as 'the Centre' of each municipality (see Figure 2).

The paper's empirical data are semi-structured interviews. Such interviews are appropriate in case studies that take a translation approach (Chua, 1995; Lowe, 2001a, and 2001b; also, cf. Yin, 1984). In the selection of respondents, our intention was to capture actor groups that presumably perceive the benchmarking technique differently and have different interests in using it (cf. Malmi, 1997; and van Dooren, 2005). Based on earlier studies about various actor groups' interests and success criteria (Kouzes and Mico, 1979; Wang and Berman, 2001; Lapsley and Wright, 2004; ter Bogt, 2004 and

Figure 2

Organization Scheme of a Municipality (for reasons of clarity, some committees are excluded); Respondents are in brackets



2008; and van Dooren, 2005), we interviewed CFOs, committee chairpersons (the politicians), department managers and department controllers. Although practical circumstances prevented us from interviewing activity managers, we learned their views indirectly from the other respondents.

In 2006, we initially conducted interviews with the committee chairpersons, the department managers and the department controllers for each of the five education departments and the five social care departments. In their positions as committee chairpersons, the politicians are appointed by the municipality councils and work approximately half of the time with their assignments. A complementary group interview was conducted with the CFOs of the five municipalities one year after the first round of interviews. We conducted this group interview because it was clear we needed more information about the benchmarking project's history and the initiators' motives. A weakness in our method was our inability to make an ethnographic study as recommended by Preston et al. (1992) and Chua (1995). Instead, we mapped the process by asking the respondents to describe it. This weakness in method is somewhat remedied by the fact that the five municipalities are well known to us from other research projects, contract teaching and informal contacts (cf. Ahrens and Chapman, 2007).

In research based on a translation approach, it is important to present the ideas held by participating individuals and the definitions and interpretations that emerge from the interactions. 'The construction of technologies is a collective process, involving not only the designers but also the potential users and people affected by the technology' (Preston et al., 1992, p. 577). Therefore,

we interviewed the four actor groups who could provide accounts of public sector benchmarking in action (cf. Lowe, 2001b). Our intention in the interviews was to explore the nature and extent of the benchmarking work in the municipalities. For instance, we wished to learn which design elements had been adopted, implemented and used, how different actors perceived the benchmarking technique, and which factors had affected its implementation and use.

Our original ambition was to conduct 30 interviews, but we were only able to interview 27 respondents. Two committee chairpersons and one department manager declined to participate. We were unable to contact one chairperson by e-post or telephone. The other chairperson said she had nothing to contribute since she was new in her position. She also said it was not meaningful to interview other politicians on the committee since none of them had participated in the development of the benchmarking work. The department manager implicitly declined to participate by not responding to our email or telephone call.

BENCHMARKING IN THE NETWORK

The Establishment of a Benchmarking Network

The economic situation in the Swedish municipal sector in the 1990s was worrisome, and several municipalities struggled with deficits. This was the situation for the five municipalities in the Network because their financial condition had weakened year after year. The CFOs in these municipalities were especially concerned about the high and increasing costs in their two largest departments: the education department and the social care department. They were also puzzled by the fact that – according to the national statistics – costs for the same activities in these departments were conspicuously different among the five municipalities. In order to improve the reliability of the cost ratios, the CFOs realized the controllers needed to meet on a regular basis in order to determine which costs to include in each ratio. The controllers were easily enrolled since they liked to meet and considered the benchmarking work interesting and rather fun. The starting point of the benchmarking project, which may also be identified as the adoption point, occurred during 1996. Previously, the controllers met only on an *ad hoc* basis to discuss resource allocation models and to compare calculations for the obligatory cost ratio reporting to Statistics Sweden.

However, according to the CFOs, the benchmarking implementation work developed slowly in the years following its adoption. The controllers, who were unable to do this work alone, found that others in the departments had limited interest in helping them. From the CFOs' perspective, the slow progress in the benchmarking process was problematic. In the problematization moment, they saw benchmarking as a needed obligatory passage point because they wanted to use benchmarking to examine and even question departmental costs. In order to accelerate the benchmarking implementation, the CFOs began to negotiate with the chairpersons at the Centres. They were quite successful because, by the

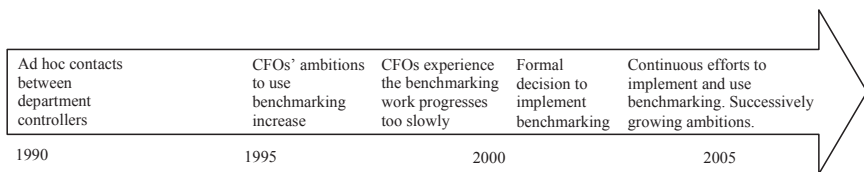
year 2001, all five municipal boards or municipal councils had formally agreed to launch a common benchmarking project. Even though the chairpersons in the Centres were not openly opposed to the idea, they still had to be convinced that the benchmarking project was necessary. One CFO admitted he had used persuasion to convince the chairpersons of his point of view:

It is a matter of letting the politicians on the board and the council understand what decisions they shall make (CFO 2).

Once the Network was formalized, the work with the cost comparisons began to advance. Reports were systematized for several departments, and meetings were held where the comparisons were presented and discussed. In some activities the benchmarking project grew from cost comparisons, which were mainly of interest to 'the Centres', into a wider performance management project involving department managers and activity managers. Figure 3 summarizes the development of the benchmarking work.

Figure 3

Key Events in the Benchmarking Project in Northern Bohuslän



However, there were several difficulties following the formal decision to adopt benchmarking. Moreover, several actors were disappointed about what had been achieved. In the following sections we describe the process from the perspective of the actors in the social care and education departments. The main focus is on their perceptions of – and explanations for – benchmarking implementation and use.

Translating Benchmarking in the Social Care Departments²

As described above, the social care controllers were enrolled early as advocates of benchmarking. They liked the idea of benchmarking and saw that its use might increase efficiency and promote collegial exchanges of knowledge and experience. They enjoyed working with controllers from other municipalities, collecting data, providing quality assurance and compiling benchmarking information.

In the beginning, the controllers collected and compared costs and staff-to-client ratios. After an initial period of problems with comparability, they thought the ratios were fairly reliable, or 'factual enough'. However, after some

time they noticed an intersement failure, that is, a problem with enlarging the benchmarking Network. The benchmarking reports did not interest the chairpersons and department managers whom the controllers thought of as receivers and users. Little attention was paid to the reports. They were not analysed, and ultimately they did not lead to process benchmarking. From the enrolled actors' perspective, the Network of benchmarking users had to be enlarged.

To enrol more actors, the department controllers assumed they had to provide more complete and more relevant descriptions of the reality. Therefore, they attempted to include some outcome (quality) measures in their reports. To that end, they began to measure client satisfaction in eldercare. Although this step was a departure from the original intention of achieving cost savings, they thought it was necessary if they were to include more actors in the network and increase the use of the information.

This translation of the idea appeared to change the social care department managers' attitude towards the benchmarking project. Originally they were sceptical of the project, but gradually they became more positive towards the comparisons. This change of attitude was in part the result of their conviction that the cost information was reasonably reliable ('factual enough') due to several years of effort (by the controller group). The main explanation, however, was their ambition to include quality comparisons in the project. Although the work with quality ratios had not progressed very far, the department managers could see benefit in receiving information that promoted discussions about value for money spent. Such information would counterbalance the Centres' excessive (in their opinion) focus on costs and give them valuable arguments in negotiations with the Centres:

... his [the financial director] interest is purely financial. He says it to my face: You are 10 million [Swedish crowns] too expensive (Social care department manager 2).

The problem is that it has been a financial project where you emphasize the financial ratios. Clearly, we need to know what our costs and prices are, but that is only a part. (...) There are other things we should measure, like the client experience. You should never ignore that. If we spend more resources than another municipality, our clients should think our care is better. It is the best measure (Social care department manager 3).

The department managers' ambition to find and implement outcome ratios seemed to be a successful countermove that moved the focus away from the initiators' original interest in controlling costs. However, neither the controllers nor the CFOs considered the new ambition a 'displacement', indicating that they may have been subject to 'counter intersement' from the department managers.

According to some controllers and department managers, the combination of survey results in eldercare and the cost information provided valuable information. They learnt that the most expensive municipality had the least

satisfied clients, and vice versa. This information was observable mostly in the department that had the lowest costs and the highest client satisfaction. The chairpersons in this municipality were very interested in the information that appeared in the annual report, in budget discussions with the Centre and in follow-up procedures:

In our municipality politicians ask for it [the benchmarking information]. But in other municipalities the interest is very low. We go through the ratios every year with the politicians. [The department] has always had good results and that is probably one explanation for the high interest. (...) We see we have low costs and high client satisfaction (Social care department controller 1).

The department manager said this comparison revealed that the department was efficiently managed, that tax revenues were wisely invested in resources for the department, and that the increase in staff was necessary since the high productivity could not continue without risking employee stress and burnout:

We can convince our politicians [in the Centre] that money is well invested in social care since they know we use tax money efficiently. We can also see we have to take steps in order not to wear the staff out. This body of information has been useful in making this argument (Social care department manager 1).

Interestingly, the chairperson and the department manager from the municipality with the highest costs, lowest client satisfaction and largest deficits throughout the studied period all declined to participate in the study.

However, on the whole there were few outcome measures. The controllers and some department managers seemed to associate this implementation problem with a failure to interest and enrol the activity managers:

We are becoming more interested in the operations, but there have been too few people from the operations involved in the comparison work (Social care department manager 3).

This was a problematic situation since it was necessary for the activity managers to analyse the information if the ratios were to be used and improved (e.g., in the development of relevant outcome indicators). The controllers also wanted the activity managers to do more process benchmarking in order to learn from each other.

There appeared to be a deadlock: outcome ratios were required for enrolling activity managers. Yet, at the same time, enrolled activity managers were prerequisites for outcome ratios.

Obviously, it was difficult to enrol activity managers if the comparisons gave them no useful information. Both controllers and department managers described and problematized the lack of involvement by the activity managers. The activity managers thought they were being asked to analyse nonfactual and incomplete representations of reality. The controllers thought the activity managers felt threatened by the described reality:

We must get the analysis group to analyse. We have said that all four years that I have been here. It is sad that [the activity managers] do not realize the benefit with it. It's as if they don't want it. When we have presented the comparisons, they have perceived it as a threat, and they tell us we have made errors. The focus has been misplaced. No, then we don't want to work with this (Social care department controller 3).

Regrettably, we do not have the activity managers on board. The way I see it, this is the financial director's project, unfortunately. (...) So far the activity managers defend themselves by maintaining there are errors in the figures (Social care department controller 5).

The work has progressed so far that everyone knows which numbers are compared. It has taken many years to create this understanding. (...) When we receive the comparative numbers, we need to analyse the differences presented. The next step is to get the activity's people in to explain the differences and show us where we are a little behind. We stand there now, and maybe we have stood there too long. We have stood there for a number of years (Social care department manager 1).

The controllers thought less priority was given to the implementation and use of benchmarking by people in poorly performing departments. Since such departments were disinclined to reveal the weaknesses that benchmarking might expose, they did not participate in the benchmarking development work. This lack of participation delayed the work. Yet the controllers believed that in the most recent year, the analysis groups had progressed compared to previous years. These groups had begun to meet in order to seek explanations for the variations in the comparison reports. Again, the controllers' explanation of the progress was that the quality comparisons (client satisfaction surveys) had made the benchmarking project more interesting for the activity managers. To a lesser degree, the activity managers thought the project was a cost reduction move initiated by the CFOs.

As described above, benchmarking information attracted the chairpersons' interest when outcome ratios were included and when the comparisons indicated a department was efficiently managed. In general, however, these chairpersons, like many activity managers, seemed to regard the benchmarking project as a cost comparison activity of minor interest. The chairpersons were not very concerned about possible failings in benchmarking implementation and use.

Generally, to what extent actors enrolled as users of benchmarking information seemed to depend on how useful the information was in the power struggle for scarce resources. Two chairpersons and two department managers stated they used the benchmarking information in appropriation discussions with the Centres in order to defend the current levels of funding and to argue that even more resources were needed. One chairperson said he realized that the Centre could use the comparisons to argue for reductions in funding although his department had not yet done so. Rather, the comparisons indicated the department's costs were reasonable and its resources were inadequate (e.g., a personnel shortage):

When we discuss the budget with [the Centre], it is good to have it [the cost comparison information] available. We can show that our requests are not unreasonable. On the other hand, they could also say that in this area, you are very expensive (Social care chairperson 1).

Another controller observed that additional information was not of use in situations when cost comparisons showed the activity was at the midpoint of the cost comparison ranking. Such a ranking implied the activity had achieved a balance between operations and cost, and that the described reality did not require action:

If you are somewhere in the middle, it is rather agreeable, but if you find yourself above or under, the cost situation in the activity is called into question (Social care department controller 4).

In two departments where the initial comparisons revealed high costs, the chairpersons asked for additional ratio comparisons in order to explain cost differences in particular activities. One chairperson, who argued that the cost differences seemed unreasonable, admitted the purpose of the additional analysis was to find a reasonable explanation for the department's high costs that would justify opposing the cutback request. In another department, the department manager and the chairperson said they used the additional information to explain the fact that their activity was more expensive and to defend its costs. In their opinion, on some occasions the comparisons had prevented cutbacks in the department:

We must be able to defend our higher costs for sheltered accommodations (Social care chairperson 2).

The controller in this department was frustrated that the information was not used more widely:

When we present the information, the explanation for the differences is that we have so many sheltered accommodations, and then it stops there. Then it [the report] goes in a desk drawer (Social care department controller 2).

In some situations, the Centres used the cost ratios to pressure the chairpersons, but such pressure seldom forced the chairpersons to make cutbacks. The chairpersons responsible for the high cost activities said reasonable explanations often existed for the high costs so it was rarely necessary to make changes. They offered various justifications for these high costs:

It is not pleasant to be the most expensive, but social care can't be measured with numbers. We decide ourselves how we want things to be in the municipality. There are reasons. We have chosen to be expensive and to place great emphasis on our sheltered accommodation. We do not fight about that. Politically we have agreed that it provides quality (Social care chairperson 2).

The chairpersons also thought that the cost ratios were useful in arguments with their department managers and activity managers. The ratios helped

explain why cutbacks were necessary or why additional allowances could not be granted. A department manager confirmed that the chairpersons had found the ratios useful in prior years when the department was experiencing fiscal stress, but they lost interest when the department's financial situation improved.

Furthermore, in one department (with comparatively high costs) the chairperson said there was criticism of the cost ratios' reliability. In another department (with the best results) the chairperson stated the figures were trustworthy and 'beyond questioning'. However, both chairpersons used the ratios: in one department, as a basis for cutbacks and resource reallocations; in the other, in budget negotiations with the Centre.

Translating Benchmarking in the Education Departments

As in the social care departments, benchmarking in the education departments in its early stages was a matter of cost and activity comparisons. The education department controllers, who were early adopters of benchmarking, used the cost comparisons to inform the chairpersons, department managers and activity managers (headmasters). Unlike the social care controllers, the education controllers had an informal leader. This controller took responsibility for summoning people to meetings and for compiling and assuring the quality of the figures. According to another controller, this controller used benchmarking to show the chairpersons and the Centre that his department had small appropriations.

The controllers thought the main receivers of the benchmarking information were the chairpersons and the activity managers. However, in their experience, these actors were not very interested in using such information. There was support for this observation in the chairpersons' interviews. One chairperson stated that the Centre in his municipality promoted the project, presumably because it wanted greater control over costs and the use of resources in the departments. He claimed this was a well thought-out strategy on the part of the Centre:

From the beginning the financial directors wanted more control over costs and how resources were used in the departments (Education chairperson 2).

In general, the controllers as well as other actors thought that cost and activity comparisons did not interest potential users. One controller said the chairpersons listened to the presentations when the benchmarking figures were presented but did not always grasp their meaning. Furthermore, the controllers said the activity managers sometimes felt threatened by the cost information. Some activity managers had expressed doubts about the reliability of the cost and activity ratios and were therefore unwilling to work with them:

Earlier I presented it at staff meetings, but that was not very successful. The staff looked at it [the benchmarking figures] as if it were something that came from

outside. What the hell does he want? We presented the information and they said: It is not like that here (Education department controller 1).

Like the social care controllers, the education controllers thought that potential users would be more interested if the benchmarking material also included outcome measures, for instance, average marks and client satisfaction. They thought a solution would be to extend the compilation and quality assurance of cost and activity ratios to outcome ratios.

However, some setbacks in the plan to complement the cost and activity information with outcome ratios had occurred. Two education controllers in particular had had an unpleasant experience. Some years ago, when they implemented average marks as a measure of quality, the activity managers reacted negatively. Their complaint was that the controllers lacked the competence to make judgements about quality in the schools and therefore were not qualified to decide on quality measures.

The first quality ratios produced mixed reactions. It was the costs in relation to the quality. We used the marks from Statistics Sweden as a measure of quality. That caused some reaction. Some people were furious that we, the controllers, thought we could measure quality (Education department controller 3).

This controversy showed the education controllers as well as the department managers that the activity managers should be more involved in the benchmarking work. These controllers still thought outcome ratios should be included. On the contrary, the department managers thought that the comparisons should concentrate more on measuring the outcome of the activities, for example, students' marks, qualifications for upper secondary school and national test results. Furthermore, they wanted more learning and experience exchanges at the operational level, not least to have a better understanding of outcome differences:

We should discuss the outcome measures, things that come out of the system: marks, qualifications and transition to upper secondary school (Education department manager 2).

According to the department managers, around the year 2003 the benchmarking project began to change from a cost comparison project managed by the controllers to a quality project driven by the department managers and the activity managers. These managers had tired of the one-sided focus on costs. They were also genuinely interested in outcome measures and in quality assurance outcome measures:

The benchmarking work is very important, but the spin-off effects are at least as important. (...) We have taken our own initiatives beyond those the controllers implemented (Education department manager 2).

However, the number of implemented outcome ratios was not large. A questionnaire was used to measure students' well-being (a measure of outcome).

Besides this, the only existing outcome measures were average marks for grades 7–9. Some activity managers also suspected that the use of different evaluation bases by the municipalities accounted for the differences in average marks. Therefore they favoured the initiation of process benchmarking. The criticism from the National Agency for Education about the lack of quality assurance in student evaluations was offered as support for this initiative. As a result, an active analysis group for grades 7–9 was established. In addition, department managers and activity managers formed analysis groups and began to work with process benchmarking using the labels ‘collegial evaluation’ and ‘critical neighbours’.

The education controllers were satisfied with this arrangement because they had learned that activity managers should be included in the work, especially in the development of the outcome measures. These controllers also agreed the activity managers should analyse the ratio comparisons in order to explain the performance gaps, and, not least, should begin to learn from each other. One controller said that as the activity managers felt more involved in the project, they participated more in the analysis groups.

Thus, local conditions and individual interests in some activities contributed to the enrolment of actors in the development of outcome ratios and in the process benchmarking activity. After a first unsuccessful attempt, when some controllers underestimated the sensitivities involved in the selection process for education outcome ratios, a general acceptance of outcome measures emerged. The department managers and some activity managers thought the use of outcome ratios meant the original idea had been translated to a quality project under their control. The advocacy of quality measures seemed, at least to some degree, a counter response to the Centres’ preference for cost comparisons.

On the whole, the education department managers did not see themselves as users of benchmarking information. However, one department manager disagreed. She said that in the spring of 2003, when she was rather new in her position, figures from a Network meeting indicated her department was inefficiently operated. It had the highest costs and its students showed the poorest performance (measured by average marks):

When I saw the numbers for the first time, I realized we had the highest costs and the poorest knowledge results. Then you had to ask yourself why (Education department manager 1).

She then appointed an analysis group consisting of the controller and the activity managers to make further analyses and to discuss how they could increase the department’s productivity and quality. The department also used the information to gradually alter its resource allocation model. One target was a teacher-to-student ratio no higher than the ratio in the other municipalities. The controller confirmed this story. After comparing both costs and outcomes (by average grades), the department manager took action. This ambitious action by a department manager as a consequence of benchmarking information was

an exception. This department manager, who was recently hired, did not feel threatened by the negative report. Because she was not responsible for past performance, she could use the report as the basis for making necessary changes.

The Centres and the chairpersons also used the benchmarking information in various ways in their negotiations about resources. In some municipalities, the Centres noted the high costs and argued for some departmental cutbacks. Two department managers were surprised their Centre and their chairpersons did not use the benchmarking information more. In other municipalities, the chairpersons, assisted by the department managers and controllers, used the cost ratios successfully to argue for more resources. One department manager was quite disappointed that his chairpersons did not use the ratios to request larger appropriations. Three controllers said they used the information in budget negotiations to show the Centres that the education activity needed more resources. One controller stated the comparisons could be used for self-interested purposes:

You use it [the comparison information] somewhat for your own benefit (Education department controller 2).

One education chairperson described how the use of the cost ratios to argue for internal reallocation of resources from pre-school to nine-year compulsory school. He also said that, as a result of the comparisons, a new resource allocation model was introduced in pre-school, and cutbacks were made in leisure activities:

When we saw that we had fewer children per employee in the leisure activities than other communities, we cut back on spending. When we were criticized by the opposition for this cutback, we could defend ourselves by showing that we were still better than other municipalities. We can point to the fact that we aren't the worst in [the region]. (. . .) Benchmarking is good for us politicians when we have to reallocate resources (Education chairperson 2).

Other chairpersons, although often unable to provide clear descriptions of how they used the information, were generally favourably disposed towards the comparisons.

When available, outcome (quality) indicators were useful in various actors' disputes about resources. Some department managers claimed the benchmarking project was of greater interest when it changed from a cost comparison project to a quality project. At this point, when the focus on outcomes gradually increased, process benchmarking became a viable path. However, two department managers said this change would not have occurred without the cost comparison. The cost information induced them to investigate whether differences in costs were the result of differences in quality. Thus, the cost comparisons apparently accelerated the work of developing outcome measures.

In contrast with these examples of the value of benchmarking, several actors thought benchmarking information was little used in their activities. For example, in a department with high costs and a high staff-to-client ratio, the controller

observed that the department manager declined to show the ratios to the chairpersons and the activity managers. The controller's assumption about this decision was that, because the information was 'threatening', the department manager wished to avoid annoying these actors. Additionally, some activity managers and some chairpersons were not convinced of the reliability of the ratios, despite improvements in the calculations. One chairperson still worried that the ratios were not correctly calculated. In his opinion, the differences in costs as reported by the various departments were mysteriously high:

What we see is a brief presentation in the committee. (...) It's easy to look at the information and then lay it aside. (...) Right costs must be compared. Sometimes a municipality is far lower and then you ask yourself: Is the measure calculated wrongly? (Education chairperson 3).

Nevertheless, in his department as well as in other departments, the information was used, although sometimes for different purposes. He admitted that the comparison reports were useful as information sources and in budget negotiations with the Centre. He thought the comparisons allowed the department to show it used its resources efficiently. Despite his doubts about the reliability of the ratios, in some cases they were still used.

DISCUSSION AND CONCLUSIONS

A large body of research concludes that PSOs that adopt benchmarking often face implementation and use problems. In this paper, we have used the translation approach and a case study to increase our understanding of these phenomena. The case illustrates that when benchmarking is adopted, a complex process involving many actors begins. As Preston et al. (1992, p. 582) write, 'the fate of a technology is in the hands of people other than the designers' (see also Briers and Chua, 2001). In this section we present the conclusions from our study on the implementation and use of benchmarking followed by our conclusions on MAIs in general.

Our first conclusion is that there is a link between benchmarking implementation problems and initiators' failure to build a strong network of benchmarking allies. Because some actors did not perceive performance benchmarking information as factual, they were unwilling to participate in the implementation work. They did not accept their roles as benchmarking co-workers and did not want to be a part of the benchmarking network. Because the initiators' failed to interest and enroll important actor groups in the implementation work, the resulting ratios were perceived as incomplete and unreliable.

Furthermore, if benchmarking is perceived as a cost containment project in the initiators' interest, other actors with other interests may resist commitment and involvement. When the purpose of benchmarking information is (or is assumed to be) tax revenue savings, in the interest of the Centre, some actors may be inclined to distance themselves from benchmarking implementation

work. For instance, the failure to implement outcome ratios was largely the result of some actors' reluctance to participate in the work. This reluctance was especially observed in activities and departments with poor performance records (high costs). In such situations, development and implementation of ratios perceived as accurate and relevant are inhibited. This is true even for actors who might have benefited from such ratios. As Briers and Chua (2001, p. 266) write in reference to an inscription device that failed: 'it was not made to be good (efficient)'. This is an illustration of the linkage between benchmarking as a technique and actors' interests.

Our second conclusion is that benchmarking implementation progresses when actors see they can make benchmarking more relevant and less focused on costs. Our study reveals that benchmarking implementation was influenced when ideas about quality comparisons and process benchmarking coincided with actors' wishes to decrease the focus on cost containment. In fact, in some cases this meant benchmarking developed into a more complete version (i.e., with more focus on outcome and process improvements). The gradually increasing interest in outcome ratios seemed to be a successful counter response that moved the focus away from the initiators' original interest in controlling costs. The original problematization – in this case, conspicuous cost differences – was hard to defend when challenged by problematization inspired by the loftier ideals of improved quality and efficiency (cf. Ezzamel, 1994; and Alcouffe et al., 2008). However, neither the CFOs nor the controllers seemed to object to the development. The controllers accepted the 'displacement' because they were persuaded by the 'counter intersement' from the department managers as well as by the necessity of forming an alliance with another actor group.

Our third conclusion is that 'successful' benchmarking implementation is a matter of creating a network of actors with mutual interests and favourable attitudes. Benchmarking – or more specifically, outcome ratios – is sometimes an unruly actant that is difficult to manage (cf. Lowe, 2001c). In one education activity outcome ratios seemed to have acquired the status of 'facts' (at least, they were 'factual enough'). In many other activities, such was not the case. In still more activities, there were no outcome ratios. This difference indicates that even in activities where actors support outcome ratios, the actant (the technique) need not yield to the interests of the actors in the network.

Our fourth conclusion is that the perception of implementation failure and success is heterogeneous and connected to various actors' adoption of the MAI. The chairpersons (with one exception) in this study seemed to view benchmarking mainly as a cost comparison project whereas department managers and controllers, although for different reasons, saw benchmarking as a project with a far broader range of design elements. The chairpersons did not experience implementation failure. For them, the gap between the adopted benchmarking and the implemented benchmarking was minimal. Presumably, their understanding of benchmarking as a concept was based on the written and oral reports the controllers had given them. Consequently, their understanding

of benchmarking was strongly influenced by the design elements actually implemented – that is, they adopted what was implemented, rather than vice versa. Because the chairpersons saw no discrepancy between the adoption and the implementation, they did not enroll as promoters of outcome ratios and process benchmarking. Accordingly, as an actor group, they were not a dynamic force for implementing further design elements in the benchmarking project.

Unlike the chairpersons, the controllers and the department managers did see a difference between the adopted benchmarking and the implemented benchmarking (e.g., the few outcome ratios, their doubts about the reliability of the cost ratios and the delay in advancing process benchmarking). Previous accounting research based on the translation approach has observed that heterogeneous interests are fuel for the translation process (Briers and Chua, 2001; Lowe, 2001b; Emsley, 2008; and Arnaboldi and Azzone, 2010). We observe that heterogeneous adoption of benchmarking (not necessarily a direct consequence of heterogeneous interests) also affects its implementation.

Our fifth conclusion is that, similar to benchmarking implementation, benchmarking use is linked to actors' interests, not least to the use of benchmarks in the struggle for resources. The chairpersons and department managers were interested in performance benchmarking when the ratios indicated success or when their use served their interests in various ways. In these instances, they used the ratios either externally at the Centres – to defend their costs or to request larger appropriations – or internally with their activity managers – to argue for cutbacks, to request resource reallocations and to justify the impossibility of additional allowances. Moreover, one department manager declined to reveal unfavourable ratios that presented a negative reality. Ezzamel (1994) has also observed how actors, when appropriate, use accounting information in budget negotiations. However, we found no convincing link between usefulness for resource negotiations and benchmarking use. Some department managers and controllers criticized the chairpersons for not using the ratios more in situations where more resources were needed.

Our sixth conclusion relates to the connection between the use of benchmarking and the perception of benchmarking information as 'factual' or 'factual enough'. The observation in accounting research that actors are unwilling to use unreliable measures is not novel. In benchmarking research, it is recognized there is uncertainty about whether performance gaps exist or whether unreliable comparison systems exist (see e.g., Northcott and Llewellyn, 2003; and Siverbo and Johansson, 2006). We found that the sparing use of benchmarking information by some actors connects to their doubts about the factuality of the benchmarking information. The failure to establish facts, and thus the failure to shape user groups' reality, frustrated the users of benchmarking. However, the perception of benchmarking information as 'factual' or 'factual enough' (and therefore usable) seemed to be as much a matter of whether the information served the actors' interests as of correct or incorrect ratios. Our study reveals that the chairperson who questioned the reliability of the cost

ratios used them in budget negotiations with the Centre. That is, the same figures were 'factual enough' (Chua, 1995, p. 138; see also Briers and Chua, 2001) in a convenient situation. This is a further indication of how the faith in a technology is interconnected with the interests of actors.

Our last conclusion relates to the increased use of benchmarking following its translation. We found that benchmarking use increased when actors other than the initiators complemented the original idea and 'counter-interested' the initiators and the controllers. The initiators accepted and even favoured this 'displacement'. Some activity managers' were motivated to participate when their department managers encouraged them to think of benchmarking as a quality project. For instance, the inclusion of outcome ratios in one activity within the education departments was an important step in the process that after a few years led to a partially new project trajectory. When some activity managers appeared to take charge of the process, the department controllers were rather pleased to be relieved of some responsibility even though 'their' cost and activity comparisons were no longer priorities.

This paper both supports and extends prior research on the implementation and use of MAIs from a translation perspective. First, it supports the view that accounting numbers are not passive reflections of an objective reality (Robson, 1992; Chua, 1995; and Lowe, 2001b and 2001c). Depending on the situation, accounting numbers sometimes constitute reality and sometimes not (Chua, 1995).

Second, and related, this paper confirms that accounting numbers may not be true or false in some objective sense. They are 'factual', or 'factual enough', depending on the content they reveal and on the interests of the organizational actors involved (Preston et al., 1992; Chua, 1995; and Briers and Chua, 2001).

Third, the translation process that is fuelled by the energy from various actor groups – especially those groups at the lower levels in organizations (Briers and Chua, 2001) – determines the success or failure of a MAI. The implementation and use of a MAI can follow a different trajectory than initiators intend. Displacements may be necessary in order to avoid the disassociation of actors from the MAI (Emsley, 2008).

Fourth, in the translation of MAIs, the adoption, implementation and use process is on-going. In this process – due to initiators' incomplete problematization and interessement – MAI implementation precedes adoption for some actor groups. Without further interessement efforts, these actor groups are unlikely to enrol in anything other than the status quo. In practice, they may not enrol as users of the implemented MAI. This heterogeneous adoption means that consensus about the failure (or success) of MAI implementation and use should not be expected.

A practical implication of this research is that traditional implementation strategies used to overcome resistance to change will not remedy implementation and use problems (from the initiators' perspective). Such problems must be dealt with through time-consuming interessement and enrolment in order to establish facts.

NOTES

- 1 This is not surprising since Callon (1986) was describing the decline in the scallop population in St. Brieuc Bay, Brittany, and not accounting changes.
- 2 All statements by the respondents were translated from Swedish.

REFERENCES

- Ahrens, T. and C.S. Chapman (2007), 'Doing Qualitative Field Research in Management Accounting: Positioning Data to Contribute to Theory', in C.S. Chapman, A.G. Hopwood and M.D. Shields (eds.), *Handbook of Management Accounting Research* (Elsevier Ltd, Oxford).
- Alcouffe, S., N. Berland and Y. Levant (2008), 'Actor-Networks and the Diffusion of Management Accounting Innovations: A Comparative Study', *Management Accounting Research*, Vol. 19, pp. 1–17.
- Alstete, J.W. (2008), 'Measurement Benchmarks or "Real" Benchmarking? An Examination of Current Perspectives', *Benchmarking: An International Journal*, Vol. 15, pp. 178–86.
- Ammons, D.N., C. Coe and M. Lombardo (2001), 'Performance-Comparison Projects in Local Government: Participants' Perspectives', *Public Administration Review*, Vol. 61, pp. 100–10.
- Andersen, B., B. Henriksen and I. Spjelkavik (2008), 'Benchmarking Applications in Public Sector Principal-Agent Relationships', *Benchmarking: An International Journal*, Vol. 15, pp. 723–41.
- Arnaboldi, M. and G. Azzone (2010), 'Constructing Performance Measurement in the Public Sector', *Critical Perspectives on Accounting*, Vol. 21, pp. 266–82.
- Askim, J. (2004), 'Benchmarking in Local Government Service Delivery: Window-Dressing or a Potent Driver for Improvement? Evidence from Norway', in S. Kuhlmann, J. Bogumil and H. Wollmann (eds.), *Leistungs- messung und -vergleich in Politik und Verwaltung* (VS Verlag, Wiesbaden).
- , Å. Johnsen and K. A. Christophersen (2008), 'Factors Behind Organizational Learning from Benchmarking: Experiences from Norwegian Municipal Benchmarking Networks', *Journal of Public Administration Research and Theory*, Vol. 18, pp. 297–320.
- Baxter, J. and W.F. Chua (2003), 'Alternative Management Accounting Research – Whence and Whither', *Accounting, Organizations and Society*, Vol. 28, pp. 97–126.
- Bhutta, K.S. and F. Huq (1999), 'Benchmarking – Best Practices: An Integrated Approach', *Benchmarking – An International Journal*, Vol. 6, pp. 254–68.
- Bjørnenak, T. and O. Olson (1999), 'Unbundling Management Accounting Innovations', *Management Accounting Research*, Vol. 10, pp. 325–38.
- Bogt, H. ter (2004), 'Politicians in Search of Performance Information? Survey Research on Dutch Aldermen's Use of Performance Information', *Financial Accountability & Management*, Vol. 20, pp. 221–52.
- (2008), 'Recent and Future Management Changes in Local Government: Continuing Focus on Rationality and Efficiency?' *Financial Accountability & Management*, Vol. 24, pp. 31–57.
- Bowerman, M. and A. Ball (2000), 'Great Expectations: Benchmarking for Best Value', *Public Money & Management*, Vol. 20, pp. 21–26.
- , G. Francis, A. Ball and J. Fry (2002), 'The Evolution of Benchmarking in UK Local Authorities', *Benchmarking: An International Journal*, Vol. 9, pp. 429–49.
- Briers, M. and W. F. Chua (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*, Vol. 26, pp. 237–69.
- Broadbent, J., K. Jacobs and R. Laughlin (1999), 'Comparing Schools in the U.K. and New Zealand: Individualizing and Socializing Accountabilities and Some Implications for Management Control', *Management Accounting Research*, Vol. 10, pp. 339–61.
- Bruijn, H. de and G. J. van Helden (2006), 'A Plea for Dialogue Driven Performance-Based Management Systems: Evidence from the Dutch Public Sector', *Financial Accountability & Management*, Vol. 22, pp. 405–23.
- Callon, M. (1986), 'Some Elements of a Sociology of Translation: Domestication of the Scallops and the Fishermen of St Brieuc Bay', in J. Law (ed.), *Power, Action and Belief: A New Sociology of Knowledge?* (Routledge, London).

- Cavalluzzo, K.S. and C. D. Ittner (2004), 'Implementing Performance Measurement Innovations: Evidence from Government', *Accounting, Organizations and Society*, Vol. 29, pp. 243–67.
- Chua, W. F. (1995), 'Experts, Networks and Inscriptions in the Fabrication of Accounting Images: A Story of the Representation of Three Public Hospitals', *Accounting, Organizations and Society*, Vol. 20, pp. 111–45.
- Dooren, W. van, (2005), 'What Makes Organisations Measure? Hypotheses on the Causes and Conditions for Performance Measurement', *Financial Accountability & Management*, Vol. 21, pp. 363–83.
- Dorsch, J. J. and M. M. Yasin (1998), 'A Framework for Benchmarking in the Public Sector. Literature Review and Directions for Future Research', *Benchmarking – An International Journal*, Vol. 11, pp. 91–115.
- Elnathan, D., T.W. Lin and S. M. Young (1996), 'Benchmarking and Management Accounting: A Framework for Research', *Journal of Management Accounting Research*, Vol. 8, pp. 37–54.
- Emsley, D. (2008), 'Different Interpretations of a "Fixed" Concept. Examining Juran's Cost of Quality from an Actor-Network Perspective', *Accounting, Auditing & Accountability Journal*, Vol. 21, pp. 375–97.
- Ezzamel, M. (1994), 'Organizational Change and Accounting: Understanding the Budgeting System in its Organizational Context', *Organization Studies*, Vol. 15, pp. 213–40.
- Francis, G. and J. Holloway (2007), 'What have we Learned? Themes from the Literature on Best-Practice Benchmarking', *International Journal of Management Reviews*, Vol. 9, pp. 171–89.
- Geiger, D. R. and C. D. Ittner (1996), 'The Influence of Funding Source and Legislative Requirements on Government Cost Accounting Practices', *Accounting, Organizations and Society*, Vol. 21, pp. 243–64.
- Gosselin, M. (1997), 'The Effect of Strategy and Organizational Structure on the Adoption and Implementation of Activity-Based Costing', *Accounting, Organizations and Society*, Vol. 20, pp. 105–22.
- Gunasekaran, A. (2005), 'Editorial. Benchmarking in Public Sector Organizations', *Benchmarking – An International Journal*, Vol. 12, pp. 289–92.
- Helden, G. J. van and S. Tillema (2005), 'In Search of a Benchmarking Theory for the Public Sector', *Financial Accountability & Management*, Vol. 21, pp. 337–61.
- Hood, C. (1995), 'The New Public Management in the 1980s: Variations on a Theme', *Accounting, Organizations and Society*, Vol. 20, pp. 93–109.
- Hoque, Z. (2008), 'Measuring and Reporting Public Sector Outputs/Outcomes: Exploratory Evidence from Australia', *International Journal of Public Sector Management*, Vol. 21, pp. 468–93.
- Jackson, A. and I. Lapsley (2003), 'The Diffusion of Accounting Practices in the New "Managerial" Public Sector"', *The International Journal of Public Management Sector Management*, Vol. 16, pp. 359–72.
- Johansson, T. and S. Siverbo (2009), 'Explaining the Utilization of Relative Performance Evaluation in Local Government: A Multi-Theoretical Study Using Data from Sweden', *Financial Accountability & Management*, Vol. 25, pp. 197–224.
- Johnsen, Å. (2005), 'What Does 25 Years of Experience Tell us About the State of Performance Measurement in Public Policy and Management?' *Public Money & Management*, Vol. 25, pp. 9–17.
- (2007), *Resultatstyrning i offentlig sektor* (Fagbokforlaget, Bergen).
- and J. Vakkuri (2006), 'Is there a Nordic Perspective on Public Sector Performance Measurement?' *Financial Accountability & Management*, Vol. 22, pp. 291–308.
- Kouzes, J. M. and P.R. Mico (1979), 'Domain Theory: An Introduction to Organizational Behavior in Human Service Organizations', *The Journal of Applied Behavioral Science*, Vol. 15, pp. 449–69.
- Lancer Julnes, P. de and M. Holzer (2001), 'Promoting the Utilization of Performance Measures in Public Organizations: An Empirical Study of Factors Affecting Adoption and Implementation', *Public Administration Review*, Vol. 61, pp. 693–708.
- Langfield-Smith, K. (2008), 'Strategic Management Accounting: How far Have We Come in 25 Years?' *Accounting, Auditing & Accountability Journal*, Vol. 21, pp. 204–28.
- Lapsley, I. and J. Pallot (2000), 'Accounting, Management and Organizational Change: A Comparative Study of Local Government', *Management Accounting Research*, Vol. 11, pp. 355–74.
- and E. Wright (2004), 'The Diffusion of Management Accounting Innovations in the Public Sector: A Research Agenda', *Management Accounting Research*, Vol. 15, pp. 355–74.

- Latour, B. (1986/1998), *Artefaktens återkomst. Ett möte mellan organisationsteori och tingens sociologi* (Nerenius & Santérus förlag, Stockholm).
- Lee, Y. P., S. Zailani and K. L. Soh (2006), 'Understanding Factors for Benchmarking Adoption', *Benchmarking – An International Journal*, Vol. 13, pp. 548–65.
- Leseure, M.J., J. Bauer, K. Birdi, A. Neely and D. Denyer (2004), 'Adoption of Promising Practices: A Systematic Review of the Evidence', *International Journal of Management Review*, Vol. 5/6, pp. 169–90.
- Llewellyn, S. and D. Northcott (2005), 'The Average Hospital', *Accounting, Organizations and Society*, Vol. 30, pp. 51–66.
- Longbottom, D. (2000), 'Benchmarking in the UK: An Empirical Study of Practitioners and Academics', *Benchmarking: An International Journal*, Vol. 7, pp. 98–117.
- Lounsbury, M. (2008), 'Institutional Rationality and Practice Variation: New Directions in the Institutional Analysis of Practice', *Accounting, Organizations and Society*, Vol. 33, No. 4–5, pp. 349–61.
- Lowe, A. (2001a), "Action at a Distance": Accounting Inscriptions and the Reporting of Episodes of Clinical Care', *Accounting Forum*, Vol. 25, pp. 1–25.
- (2001b), 'After ANT. An Illustrative Discussion of the Implications for Qualitative Accounting Case Research', *Accounting, Auditing & Accountability Journal*, Vol. 14, pp. 327–51.
- (2001c), 'Accounting Information Systems as Knowledge-Objects: Some Effects of Objectualization', *Management Accounting Research*, Vol. 12, pp. 75–100.
- and B. Koh (2007), 'Inscribing the Organization: Representations in Dispute Between Accounting and Production', *Critical Perspectives on Accounting*, Vol. 18, pp. 952–74.
- Magd, H. and A. Curry (2003), 'Benchmarking: Achieving Best Value in Public-Sector Organisations', *Benchmarking: An International Journal*, Vol. 10, pp. 261–86.
- Malmi, T. (1997), 'Towards Explaining Activity-Based Costing Failure: Accounting and Control in a Decentralized Organization', *Management Accounting Research*, Vol. 8, pp. 459–80.
- Mennicken, A. (2008), 'Connecting Worlds: The Translation of International Auditing Standards into Post-Soviet Audit Practice', *Accounting, Organizations and Society*, Vol. 33, pp. 384–414.
- Meyer, J.W. and B. Rowan (1977/1991), 'Institutionalized Organizations: Formal Structure as Myth and Ceremony', in W.W. Powell and P.J. DiMaggio (eds.), *The New Institutionalism in Organizational Analysis* (The University of Chicago Press, Chicago and London).
- Modell, S. (2007), 'Managing Accounting Change', in T. Hopper, D. Northcott and R. Scapens (eds.), *Issues in Management Accounting* (Prentice Hall, Harlow).
- (2009), 'Institutional Research in Performance Measurement and Management in the Public Sector Accounting Literature: A Review and Assessment', *Financial Accountability & Management*, Vol. 25, pp. 277–303.
- Moriarty, J.P. and C. Smallman (2009), 'En Route to a Theory of Benchmarking', *Benchmarking: An International Journal*, Vol. 16, pp. 484–503.
- Northcott, D. and S. Llewellyn (2003), 'The Ladder of Success in Healthcare: The UK National Reference Costing Index', *Management Accounting Research*, Vol. 14, pp. 51–66.
- (2005), 'Benchmarking in UK Health: A Gap Between Policy and Practice?' *Benchmarking: An International Journal*, Vol. 12, pp. 419–35.
- Östergren, K. (2006), 'The Institutional Construction of Consumerism: A Study of Implementing Quality Indicators', *Financial Accountability & Management*, Vol. 22, pp. 179–205.
- Otley, D. (2008), 'Did Kaplan and Johnson Get it Right?' *Accounting, Auditing & Accountability Journal*, Vol. 21, pp. 229–39.
- Pallot, J. (2001), 'A Decade in Review: New Zealand's Experience with Resource Accounting and Budgeting', *Financial Accountability & Management*, Vol. 17, pp. 271–90.
- Papadopoulos, T., Z. Radnor and Y. Merali (2011), 'The Role of Actor Associations in Understanding the Implementation of Lean Thinking in Healthcare', *International Journal of Operations & Production Management*, Vol. 31, No. 2, pp. 167–91.
- Pollit, C. (2006), 'Performance Information for Democracy: The Missing Link?' *Evaluation*, Vol. 12, pp. 38–55.
- Preston, A.M., D. Cooper and R. W. Coombs (1992), 'Fabrication Budgets: A Study of the Production of Management Budgeting in the National Health Service', *Accounting, Organizations and Society*, Vol. 17, pp. 561–93.
- Rigby, D. and B. Bilodeau (2007), *Management Tools and Trends 2007* (Bain & Company, Boston, MA).

- Robson K. (1992), 'Accounting Numbers as "Inscription": Action at a Distance and the Development of Accounting', *Accounting, Organizations and Society*, Vol. 17, pp. 685–708.
- Røvik, K. A. (1998/2000), *Moderna Organisa-tioner* (Liber, Malmö).
- Siverbo, S. and T. Johansson (2006), 'Relative Performance Evaluation in Swedish Local Government', *Financial Accountability & Management*, Vol. 22, pp. 271–90.
- Skærbæk, P. and P. Melander (2004), 'The Politics of the Changing Forms of Accounting. A Field Study of Strategy Translation in a Danish Government-Owned Company Under Privatisation', *Accounting Auditing & Accountability Journal*, Vol. 17, pp. 17–40.
- (2007), 'The Commodification of the Danish Defence Forces and the Troubled Identities of its Soldiers', *Financial Accountability & Management*, Vol. 23, pp. 243–68.
- Triantafillou, P. (2007), 'Benchmarking in the Public Sector: A Critical Conceptual Framework', *Public Administration*, Vol. 85, pp. 829–46.
- Wang, X. and E. Berman (2001), 'Hypotheses about Performance Measurement in Counties: Findings from a Survey', *Journal of Public Administration Research and Theory*, Vol. 11, pp. 403–28.
- Wolfram Cox, J., L. Mann and D. Samson (1997), 'Benchmarking as a Mixed Metaphor: Disentangling Assumptions of Competition and Collaboration', *Journal of Management Studies*, Vol. 32, pp. 285–314.
- Yin, R. K. (1984), *Case Study Research: Design and Methods* (Sage Publications, Beverly Hills, CA).