

# PUBLIC SECTOR ORGANIZATIONS' USE OF BENCHMARKING INFORMATION FOR PERFORMANCE IMPROVEMENT

Theoretical Analysis and Explorative Case Studies in Dutch Water Boards

SANDRA TILLEMA  
*University of Groningen*

**ABSTRACT:** *Public sector organizations frequently adopt benchmarking as a management tool to improve performance. This paper focuses on the question of whether it is likely that these organizations actually use the resulting benchmarking information to develop and implement performance improvement actions. It distinguishes various mechanisms through which specific public sector characteristics may stimulate or hinder the use of benchmarking information. The paper illustrates several mechanisms through four explorative case studies of Dutch water boards. A subsequent analysis reveals that the applicability of particular mechanisms depends on the specific circumstances, including stakeholders' involvement in the organization and the strength of the organization's managers.*

**KEYWORDS:** *benchmarking, new public management, performance management, public sector*

Recently, public sector organizations have introduced various private sector management styles and techniques, ranging from explicit performance measures and output controls to hands-on management and flexible reward systems (Hood, 1991, 1995). It is widely acknowledged, however, that public and private sector organizations differ in various important respects (Nutt & Backoff, 1993; Rainey, 1989; Rainey, Backoff, & Levine, 1976), which raises the question of whether the differences between the two sectors act as barriers to the transfer of management styles and techniques from the private to the public sector (Boyne, 2002). This paper focuses on one of the private sector techniques that has gained popularity

in the public sector recently: namely, the use of benchmarking as a management tool to improve performance. It examines whether the resulting benchmarking information is likely to be used for performance improvement in the public sector, given the specific characteristics of that sector.

Based on the differences that exist between public and private sector organizations, we might expect public sector benchmarking to be different from private sector benchmarking. So far, however, our knowledge of the relations between sector characteristics and the design and use of benchmarking is limited. Despite the emerging literature on public sector benchmarking (see Dorsch & Yasin, 1998), Bowerman, Francis, Ball, and Fry expressed concern about "the lack of understanding of the real nature of benchmarking in the [public] sector" (2002, p. 430). Based on extensive case and survey research, they argued that public sector benchmarking is different from private sector benchmarking in three respects. To be specific, they assert that—contrary to its private sector counterpart—public sector benchmarking may aim at an acceptable or average performance level rather than best practice (e.g., Llewellyn & Northcott, 2005), it is often imposed, and it frequently generates information that is disclosed outside the organization. Given these differences, Bowerman et al. questioned whether "the promise that benchmarking in the public sector will lead to enhanced organizational efficiency and effectiveness" can actually be fulfilled (pp. 434–435).

This paper takes the view that the differences between public and private sector benchmarking are the result of between-sector differences in organizational characteristics, and it investigates the relations between specific characteristics of public sector organizations and their use of benchmarking information for performance improvement. When investigating this topic, it should be recognized that there is no stereotypical public sector benchmarking project. That is, variations can be expected, not only between different benchmarking projects but also between different segments of the public sector. However, the first part of the paper (i.e., the theoretical analysis) ignores these complexities. It develops a theoretical framework that might be useful—to a greater or lesser extent—in investigating any public sector benchmarking project. The second part of the paper (i.e., the empirical analysis) investigates the applicability of the framework to a specific setting through four explorative case studies of Dutch water boards. The objectives of this empirical research are twofold: (a) to find illustrations of particular elements of the theoretical framework and (b) to find explanations for empirical observa-

---

The author thanks Henk ter Bogt, Jan van Helden, Bob Scapens, and two anonymous reviewers for their comments on and useful suggestions for revising earlier versions of this paper. Address correspondence to: Sandra Tillema, University of Groningen, Faculty of Economics, P.O. Box 800, 9700 AV Groningen, Netherlands; e-mail: s.tillema@rug.nl.

tions that are inconsistent with elements of this framework. These explanations improve our understanding of the circumstances that influence the applicability of certain elements of the theoretical framework.

In other words, the ultimate objective of this paper is to uncover the characteristics of public sector organizations that may affect those organizations' use of benchmarking information for performance improvement, to describe the mechanisms through which these characteristics function, to find illustrations of these mechanisms, and to explore under which circumstances certain mechanisms can or cannot be expected.

## **Theoretical Framework**

### **FORMS, TYPES, AND LEVELS OF BENCHMARKING**

In both the public and the private sector, various forms of benchmarking have been conducted for many years (Bowerman & Ball, 2000; Elnathan, Lin, & Young, 1996; Miller, De Meyer, & Nakane, 1992, chapter 2). Examples are data collected from member organizations by industry associations, and annual reviews of key statistics published by associations that focus on specific organizational functions. More recently, however, benchmarking has become a popular management tool to identify performance gaps and to improve performance. Using an overview provided by Spendolini (1992, pp. 9–10), I define this form of benchmarking as a continuous, systematic process for measuring, comparing, evaluating and understanding the products, services, functions or work processes of organizations (that are recognized as representing best practices)<sup>1</sup> for the purpose of organizational improvement. Therefore, a central characteristic of this form of benchmarking is that it aims at improving performance.

This form of benchmarking was popularized in the private sector, but it has been increasingly promoted in the public sector as well (Bowerman et al., 2002).<sup>2</sup> Today, many public sector organizations—ranging from central government departments and local government organizations to police forces and hospitals—are engaged in benchmarking projects that are aimed explicitly at performance improvement (e.g., Bowerman & Ball, 2000; Bowerman, Ball, & Francis, 2001; Dorsch & Yasin, 1998; de Groot, Goudriaan, Hoogwout, de Jong, & Poerstamper, 2004; Northcott & Llewellyn, 2003). However, performance improvement is not the sole objective of these projects. Other objectives may include meeting external requirements to provide comparative data, demonstrating or increasing accountability to the public for the use of resources, justifying or defending existing performance, and proving that the organization compares well against alternative service providers (Bowerman & Ball, 2000). These objectives may act as a barrier to the use of benchmarking as a management tool for performance improvement (Bowerman & Ball, 2000; Bowerman, Ball, & Francis 2001; Bowerman, Francis, Ball, &

Fry 2002). In other words, although I concentrate on public sector benchmarking projects that have performance improvement as a their main objective, I recognize that attempts made to attain this objective might be influenced by other objectives of these benchmarking projects.

At this point, it is also important to note that within the performance-improvement-oriented form of benchmarking, different types and levels can be distinguished. Fedor, Parsons, and Shalley (1996), for example, described four types (i.e., internal, industry, competitive, and best-in-class) and two levels of benchmarking (i.e., outcome and process). Basically, the remainder of this paper concentrates on the combination of industry benchmarking (involving comparisons between organizations that come from the same industry but that are not necessarily direct competitors) and benchmarking at the outcome level (involving a comparison of results rather than processes). This particular combination of benchmarking type and level is commonly referred to as *relative performance evaluation* (Northcott & Llewellyn, 2003, p. 52).

### **CHARACTERISTICS THAT INFLUENCE THE USE OF BENCHMARKING INFORMATION**

Although many organizations, including public sector organizations, are engaged in benchmarking projects that aim at performance improvement, they are not necessarily using the resulting benchmarking information to improve their performance. In this respect, several authors (e.g., Ammons, 1999; Elnathan et al., 1996; Fedor et al., 1996) have stressed the importance of a "proper mentality" or culture. Such a culture would consist of a recognition that one's own organization is probably not the best in all aspects of its operations, an eagerness to learn from other organizations, an openness in exchanging information, a willingness to adopt or adapt ideas developed elsewhere, and a resistance to tendencies to defend one's own (underperforming) operations. These conditions for the use of benchmarking information for performance improvement apply to organizations in general, that is, both public and private sector organizations. Public sector organizations, however, differ in various respects from private sector organizations (Boyne, 2002; Nutt, 2000; Nutt & Backoff, 1993; Rainey, 1989; Rainey et al., 1976).<sup>3</sup> Rainey et al. (1976), for instance, distinguished 25 differences between public and private sector organizations. I argue that six (combinations) of these differences are of particular interest when studying the use of benchmarking information for performance improvement in public sector organizations. In the following, I discuss the mechanisms through which the six public sector characteristics may influence such a use of benchmarking information.

#### ***Limited Market Exposure***

Public sector organizations are exposed to economic markets to a lesser extent than private sector organizations. In this context, it is important to note that economic

markets can serve as a source of revenues and as a source of information. As a source of revenues, markets enforce relatively automatic penalties and rewards. Consequently, public sector organizations are less likely to be subject to automatic incentives for cost reduction, operating efficiency, and effective performance (Rainey et al., 1976). The relative absence of these incentives might reduce their motivation to develop and implement action plans in response to benchmarking information. Van Helden and Tillema (2005), for instance, argued that public sector organizations have less need to conform to other, better performing organizations because nonconformance does not result in a direct threat to the organization's survival.

As a source of information, however, economic markets may have quite a different effect on the use of benchmarking information. That is, economic markets provide relatively clear, quantitative demand indicators, goals, and performance measures, such as prices, sales, and profits (Rainey et al., 1976). Their relative absence in the public sector means that public sector organizations do not have clear market signals about their success or failure (Nutt & Backoff, 1993). In such a situation, benchmarking information can be seen as an alternative to market signals (van Helden & Tillema, 2005, p. 339). That is, if public sector organizations have not formed (due the absence of clear market signals) an impression of their relative performance (i.e., their performance compared to similar organizations), benchmarking information might provide a first indication that they are not competitive in certain areas. Organizations may then use benchmarking information to develop and implement action plans to close the performance gaps revealed by the benchmarking.

Following from this argument, clearly a limited exposure to economic markets can have two opposite effects on the use of benchmarking information. On one hand, benchmarking can provide organizations with information about their relative performances that would otherwise be unavailable due to the absence of market signals. This information might stimulate these organizations to improve performance. On the other hand, due to the lack of market competition, the same organizations have no direct economic incentive to use that information for performance improvement. Under these circumstances, it might be important that organizations experience institutional pressures to improve their (relative) performance and to use benchmarking information for that purpose (van Helden & Tillema, 2005, pp. 340–341). Oversight bodies, for example, can raise considerations of efficiency and effectiveness (Nutt & Backoff, 1993; see also Northcott & Llewellyn, 2003, p. 57), which might stimulate organizations to use benchmarking information for performance improvement. The impact of institutional pressures is examined in the following discussion.

To some extent, public sector organizations may actually benefit from being exposed to economic markets to only a limited degree. More precisely, the sharing of information between organizations—which, to varying degrees, is part of the

benchmarking process—is often seen as problematic, especially for organizations within the same industry (Camp, 1995, p. 81; Drew, 1997, p. 431). Matters of confidentiality, sustaining competitive advantage, and antitrust laws may act as barriers to a within-industry information exchange and could complicate not only the process of finding benchmarking partners and gathering benchmarking data (Bowerman et al., 2002, p. 445) but also the process of investigating which business processes underpin superior benchmarking scores. In the absence of real market competition, these considerations are of only minor importance. As a result, public sector organizations may experience fewer difficulties in finding the means to close performance gaps.

### *Institutional Constraints*

Public sector organizations are confronted, to a much greater degree than private sector organizations, with legal mandates, other formal arrangements, and traditions that pose constraints on their purposes, methods, and domains of operation (Nutt & Backoff, 1993; Rainey et al., 1976; see also DiMaggio & Powell, 1983, pp. 150–151; Greenwood & Hinings, 1996, pp. 1027–1028; Meyer & Rowan, 1977, pp. 343–344). For instance, public sector organizations are often expected to provide a specific set of services to a particular type of customers who are located in a particular geographical area. With respect to benchmarking, this requirement may imply that these organizations are less autonomous and less flexible when developing and implementing action plans. In other words, compared with private sector organizations, public sector organizations have a smaller set of alternative actions that they could implement to close performance gaps.

### *Formal and Informal Influences*

In public sector organizations, the decision-making process is more complex due to a greater diversity and intensity of formal and informal influences from, for example, oversight bodies, governing bodies, individual politicians, interest groups, and public opinion (Rainey et al., 1976). Each of these stakeholders places demands and constraints on the organization, and, as a consequence, public sector organizations are often pushed and pulled in many different directions at the same time (Boyne, 2002). In such a situation, with multiple and diverse interests, compromises are needed to arrive at decisions that are acceptable to all the stakeholders involved. This need to compromise can apply equally to decisions about the action plans that could be developed and implemented in response to benchmarking results. As a consequence, public sector organizations may need more time to make decisions about the content of such action plans, or they may defer those decisions, and when they do make them, these decisions may be less radical than those of private sector organizations. The influence of multiple and diverse interests also results in more complex objectives.

### ***Public Scrutiny***

Public sector organizations are subject to greater public scrutiny than private sector organizations. They are able to keep fewer secrets and are subject to more outside monitoring (Rainey et al., 1976). Because of such issues as openness, accountability, and oversight, public sector organizations often must communicate their benchmarking scores to a wider public (Bowerman et al., 2002, p. 434). Actually, disclosure is a direct consequence of some of the main objectives of many public sector benchmarking projects. One might argue that the disclosure of benchmarking scores can result in institutional pressures being exerted by stakeholders for the development and implementation of action plans (van Helden & Tillema, 2005; cf. Bowerman et al., 2002, p. 445). However, in a number of papers, Bowerman (Bowerman & Ball, 2000; Bowerman, Ball, & Francis 2001; Bowerman, Francis, Ball, & Fry 2002) questioned whether benchmarking can enhance organizational efficiency and effectiveness when the results of benchmarking projects are published. Under such circumstances, organizations may regard benchmarking as a tool to demonstrate good performance and not as a tool to improve performance. As a result, when analyzing benchmarking scores, public sector organizations may not look for the inferior aspects of their business processes that might require improvements; instead, they search for factors that can be used to explain why their organization is not to blame for the inferior benchmarking scores. If such explanatory factors can be found, the organizations may not be committed to using the benchmarking information to improve performance.

### ***Complex Objectives***

Public sector organizations' mix of objectives is more complex in the sense that these objectives are multiple, diverse, and often conflicting, due to the influences of a variety of stakeholders (Rainey et al., 1976). A consequence of such complex and probably ambiguous objectives is that the outcomes of control processes depend on issues such as power structures, negotiation processes, the need to distribute scarce resources, vested interests, and conflicting values (Hofstede, 1981, pp. 197–198; see also Cyert & March, 1959, pp. 78–79; Simon, 1964; Wildavsky, 1989, pp. 275–280). As the benchmarking process can be regarded as a form of control, these issues may affect the development, content, and implementation of subsequent action plans. More specifically, decisions concerning such action plans may depend on the particular stakeholders that have a say in those decisions, their interests, and the problems and solutions that concern them. Together these elements influence the decisions that an organization makes as well as the decisions that it overlooks or deliberately avoids.

### ***Barriers to Innovations***

In public sector organizations, the cycle of elections and political appointments may result in frequent changes in policy and in interruptions in the implementation

of plans and projects. This—together with a weaker hierarchical authority and a lack of incentives for successful innovations—may explain why public sector organizations tend to be more cautious, less flexible, less innovative, and more short-sighted compared with private sector organizations (Boyne, 2002; Rainey et al., 1976). With respect to benchmarking, these characteristics may imply that compared with their private sector counterparts, public sector organizations are more likely to terminate the development or implementation of action plans and are more likely develop less radical plans. It should be noted, however, that benchmarking might also be a solution to some of the problems faced by public sector organizations that seek to generate alternative means of meeting expectations of increased efficiency and effectiveness (cf. Davies, 1998; Nutt, 2000). These problems include a lack of data describing emerging trends in service delivery, limited financial resources to uncover alternatives, and a fear of accusations that public money is being wasted. Under such circumstances, benchmarking could be seen as a time- and cost-saving means of learning and innovating. In other words, benchmarking is an alternative to research and development activities and may be used by public sector organizations to induce innovations. Still, the innovations discovered through benchmarking are generally imitative or incremental rather than radical in nature (cf. Drew, 1997, p. 429).

Thus, I argue that important characteristics of public sector organizations may influence these organizations' use of benchmarking information for performance improvement. Whereas some characteristics may influence this use of benchmarking information in a single way, other characteristics may influence it in various ways. I refer to a certain way in which a particular characteristic influences the use of benchmarking information as a *mechanism*. See Table 1 (left-hand side) for an overview of the 12 mechanisms. It is important to note that we should not expect all or even most of these mechanisms to be relevant in all situations. The applicability of the various mechanisms depends on the specific circumstances, including the characteristics of a particular benchmarking project, the segment of the public sector that undertakes the project, and the organization under consideration. To get a first impression of the circumstances in which particular mechanisms are (not) relevant, after an overview of research methods, I examine a public sector benchmarking project in a specific empirical context.

## **Research Method**

The empirical research focused on a benchmarking project that was conducted by Dutch water boards between 1999 and 2001. Water boards are governmental authorities that are responsible for the water management in a given area (Unie van Waterschappen, 1999). One of their main tasks is wastewater treatment, whereby wastewater from households and businesses is transported via the sewers to waste-

**Table 1. Overview of Public Sector Characteristics, the Mechanisms Through Which They Operate, the Direction of Their Influence, and the Case Studies' Results**

<i>Characteristics and Mechanisms</i>	<i>Direction</i>	<i>Water Boards</i>			
		<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>
Limited market exposure					
1a. Less subject to market incentives for performance improvement: fewer incentives to use benchmarking information.	-	C	C	C	S
1b. No clear market signals about success or failure: benchmarking information provides first indication of relative performance: incentive to use benchmarking information.	+	C	S	C	C
1c. Lower barriers to sharing information: easier to collect information about business processes that underpin good benchmarking scores.	+	C	C	C	C
Institutional constraints					
2. More constraints on purposes, methods, and domains of operation: less autonomy and less flexibility in developing and implementing action plans.	-	NR	NR	S	NR
Formal and informal influences					
3a. Stronger and more diverse influence of stakeholders: compromises must be reached: decisions on action plans need more time or are deferred.	-	C	C	C	C
3b. Stronger and more diverse influence of stakeholders: compromises must be reached: action plans are less radical.	-	C	C	C	C
Public scrutiny					
4a. Benchmarking scores are disclosed: stakeholders pressure organization to improve: incentive to use benchmark information.	+	C	C	C	C
4b. Benchmarking scores are disclosed: tendency to defend scores: fewer incentives to use benchmarking information.	-	C	C	C	S
Complex objectives					
5. Multiple, diverse, and conflicting objectives: initiation, content, and implementation of action plans depend on the stakeholders that are involved, their interests, relative power positions, and visions.	+/-	C	S	C	S

Barriers to innovation		NR	NR	NR	NR
6a. Cycle of elections and political appointments: frequent policy changes: development or implementation of action plans is terminated.	-	NR	NR	NR	NR
6b. Cycle of elections and political appointments: fear that implementation of action plans will be terminated after policy changes: action plans are less radical.	-	NR	NR	NR	NR
6c. Problems faced when generating means to improve: benchmarking is suitable means of learning and innovating.	+	C	C	C	C

*Notes:* S = Support: The case study results are consistent with the mechanism (i.e., the case study provides an illustration of this mechanism). C = Contradiction: The case study results contradict the expected outcomes of the mechanism (i.e., either the characteristics of the water board or the water board sector differ from the characteristics of public sector organizations that formed the basis for this mechanism, or the mechanism does not operate as expected. In practice, the distinction between these two situations is ambiguous. It was, therefore, decided to ignore this distinction in Table 7. NR = No results: The case study did not lead to results that can be linked to the mechanism. The mechanisms are numbered to facilitate cross-referencing.

water treatment plants and subsequently purified. The 25 water boards that were involved in wastewater treatment in 1999 carried out a benchmarking project, with the main objective of improving the water boards' performances.

The Association of Water Boards appointed external consultants to develop the performance measurement instrument and to collect and analyze the data. Together with representatives of certain water boards, these consultants developed an adapted version of the balanced scorecard as the measurement instrument (van Helden & Admiraal, 2002; see also Kaplan & Norton, 1992, 1996). This instrument comprised four perspectives: an operating perspective, a financial perspective, an environmental perspective, and an innovative perspective. Together, the four perspectives included 11 key performance indicators and a large number of other performance indicators. The benchmarking results were published in March 2001 (Arthur Andersen & Vertis, 2001). A subsequent survey among all participating water boards (with a response rate of 96 percent) revealed that the water boards very much appreciated the benchmarking information. For example, almost 85 percent of the water boards indicated that the benchmarking project had resulted in useful pointers to improve their performance (Tillema & van Helden, 2005).

The wastewater treatment benchmarking project was investigated using case research methods. Two water boards (labeled A and B) that had used the benchmarking information extensively and two water boards (labeled C and D) that had used this information only to a limited extent were selected. This selection of cases allows us to investigate the influence of various characteristics of public sector organizations on their decision to use benchmarking information for performance improvement.

The most important data source were semistructured interviews. These interviews took place in October 2002, about 18 months after the results of the benchmarking project had been published. Important issues in the interviews were the number, content, developmental stage, initiators, and (expected) effects of the action plans that had been developed as a result of the project; the reasons for developing or not developing action plans; the attitude toward the benchmarking project; the attitude toward performance gaps; the strategy of the water boards; and the influence of stakeholders. The persons interviewed in each water board were the chairperson, another member of the executive body, the secretary-director, the head of wastewater treatment, the controller, and the internal coordinator of the benchmarking project. The secretary-director of one of the water boards was not interviewed, however, because this person was not involved in the benchmarking project. The interviews lasted (with a few exceptions) between 45 and 90 minutes each and were tape-recorded. All interviews were summarized in interview reports, which were sent to the interviewees for approval. In addition to the interviews, document analysis, talks with persons in the water board sector, and group meetings were used to collect and verify the data.

## Results

### OVERVIEW

Table 1 summarizes the theoretical analysis. More precisely, the left-hand column comprises the public sector characteristics previously discussed and the mechanisms through which these characteristics may affect the use of benchmarking information for performance improvement. The signs in the middle column indicate that a particular mechanism is expected to result in a positive (+), negative (-) or indeterminate (+/-) influence on the use of benchmarking information. Table 1 also includes a summary of the empirical results. That is, the four right-hand columns reflect the extent to which the results from each of the four case studies are consistent with the theoretical framework.

Table 1 shows that the case research found illustrations of five mechanisms discussed in the theoretical framework. However, it also shows that the situation in the water boards that did not provide these illustrations actually contradicted four of these five mechanisms. Moreover, for five other mechanisms, the case research produced only contradictions.

### ILLUSTRATIONS OF THEORETICAL FRAMEWORK

#### *Mechanism 1a*

Mechanism 1a suggests that public sector organizations may be less motivated to use benchmarking information for performance improvement, as they are subject to fewer market incentives. This argument clearly applied to Water Board D. That is, although the benchmarking results suggested that its performance was not as good as expected, this water board decided to use the benchmarking information only to a limited extent for performance improvement. It argued that its performance in absolute terms is much more important than its performance compared with other organizations and that it was not required to match the performance of better-performing water boards. In other words, the water board argued that there was no need to be competitive with other water boards. So apparently, the rewards and penalties for (not) improving performance were too small to induce this water board to act on the benchmarking results.

#### *Mechanism 1b*

Similar to Water Board D, Water Board B stressed that it did not explicitly aim to improve its future benchmarking scores; improvements in absolute performance were regarded as more important than improvements in relative performance. However, in contrast to Water Board D, Water Board B indicated that it used the benchmarking information to judge the appropriateness of its performance targets in absolute terms. The wastewater treatment benchmarking project revealed

to Water Board B that its performance was much lower than that of many other water boards. As a reaction, this water board accelerated the implementation of its planned performance improvement actions. In other words, Water Board B regarded its benchmarking scores as an indication that its competitive position was rather weak, and this realization increased its sense of urgency about closing the gap between its targeted and actual performance. Thus, the water board used the benchmarking information as a substitute for market signals, which is consistent with Mechanism 1b.

### ***Mechanism 2***

Mechanism 2 suggests that because of institutional constraints, public sector organizations are less autonomous and less flexible when developing and implementing action plans following the results of a benchmarking project. Water Board C provides an illustration of such a situation. In this water board, there was a discussion of whether the organization was allowed to continue undertaking certain activities that had previously had a positive impact on its benchmarking scores in the financial perspective. More precisely, the water board was using excess capacity of its fermentation installations to generate electricity by processing liquid waste products from other industries. This electricity was being used by the water board itself and was thereby reducing the amount of electricity that it had to buy from power and gas companies. A further increase in these waste-processing activities could be achieved by increasing the capacity of fermentation installations to well beyond the size that was needed for the wastewater treatment activities. Such expansion could potentially lead to a situation in which the water board sells electricity (and heat for district heating) to power and gas companies. The proponents of this policy argued that such activities contribute not only to the water board's financial performance but also to a healthier environment. The opponents cast doubts on the desirability of competing with for-profit organizations and feared the risks that might be involved. In their view, organizations funded by public money, including the water boards, should not assume such risks. Although at the time of the case research a decision had not yet been reached, this example illustrates that institutional constraints may influence the content of public sector organizations' action plans.

### ***Mechanism 4b***

The results of the wastewater treatment benchmarking project were made public. Mechanism 4b suggests that under such circumstances, public sector organizations may regard benchmarking as a means to defend rather than to improve performance. Water Board D appeared to be most susceptible to the defensive mechanisms that can arise from the publication of benchmarking scores. That is, to avoid negative reactions from its external stakeholders, this water board's

first response to the benchmarking results was to organize a press conference and give an explanation for its disappointing scores. Subsequently, the benchmarking results were largely ignored, and some of the people interviewed in the course of the case study openly admitted that only these interviews had spurred them to dig up the benchmarking report. It appeared that once the water board had given an explanation for its benchmarking scores, the report was put away and not used to improve performance. The interviewees in this water board explained the water board's passivity also by referring to certain aspects of the design of the benchmarking project. They criticized the measurement instrument and questioned the comparability of the resulting performance information. In addition, they raised serious doubts about the fairness of the data provided by the other water boards.

### *Mechanism 5*

Mechanism 5 suggests that because public sector organizations often have complex objectives, the initiation, content, and implementation of their action plans depend on power structures and negotiation processes. The situation in Water Board B is in line with this expectation. Within this water board's governing bodies, there was disagreement about the relative importance of different aspects of performance. Some members argued that the water board should start with decisions on its wastewater treatment objectives (operating and environmental perspectives) and then determine what costs must be incurred and what pollution levies must be imposed to attain these objectives; others thought that the water board should start with decisions on acceptable levels of costs and levies (financial perspective) and then determine what level of wastewater treatment could be attained given these financial constraints. As a result of this disagreement, several interviewees found it hard to rank the different aspects of the water board's performance according to their relative importance. Moreover, no general guidelines existed that could be used to evaluate performance improvement actions. When discussing individual proposals for improvement, the interviewees often stressed that, ultimately, it is the governing bodies that must decide whether the resulting improvements in operating and environmental performance outweigh the costs of implementing those proposals. The difficulty that the water board had with setting priorities meant that it was hard to decide which action plans should be implemented first. According to the chairperson, the water board was swayed by the issues of the day; and it paid only limited attention to its future position. All in all, the situation at Water Board B seemed to hinder the rapid implementation of the proposed actions.

Also, Water Board D provided an illustrative example of the difficulties that may arise from having complex objectives. The development of one of the few actions that this water board proposed as a response to the benchmarking results (i.e., the implementation of an auditing instrument for sustainability) was in quite an early stage at the time of the interviews. Particularly the governing bodies

but also the Association of Water Boards exerted pressure to implement such an instrument. It was seen as a means of informing the outside world that the water board was undertaking numerous activities to preserve the environment. The water board's staff, however, strongly resisted the introduction of an auditing instrument. They argued that actions and environmentally friendly behavior are much more important than the paperwork that goes hand in hand with the use of an auditing instrument. Apparently, staff and governing bodies disagreed about the importance of the objective to demonstrate accountability to the outside world for the water board's environmental activities. The staff delayed the implementation of an auditing instrument by referring to practical issues, such as their involvement in other projects and a forthcoming merger. It seemed that the governing bodies were not sufficiently powerful to force the staff members to speed up the process. As a result, at the time of the interviews the water board had only a vague idea about implementing an auditing instrument.

### *Summary of the Results for Water Boards B and D*

These illustrations show that certain elements of the theoretical framework are useful in explaining the extent to which Water Boards B and D used the benchmarking information for performance improvement. Water Board B's situation can be summarized and compared as follows: Given the absence of clear market signals, it was only through the benchmarking that this water board found out that its relative performance was lagging behind; subsequently, it decided to accelerate the implementation of its existing action plans (Mechanism 1b). However, the presence of conflicting objectives acted as a barrier to a rapid implementation of these plans (Mechanism 5). In the case of Water Board D—due to a relative absence of market incentives (Mechanism 1a) in combination with a fear of criticism from external stakeholders following the publication of the benchmarking scores (Mechanism 4b)—this water board focused more on explaining its scores than on using the benchmarking information to develop and implement action plans. Although the water board did develop a few action plans, the implementation of one of them was hindered by a disagreement among some of its stakeholders about the relative importance of a particular objective (Mechanism 5).

### **EXPLANATION OF CONTRADICTIONS**

We now turn our attention to the circumstances that explain why certain elements of the theoretical framework are not relevant to all or some of the water boards. For the purpose of developing our understanding of the factors that influence the applicability of certain mechanisms to certain (groups of) organizations, I employ a cross-case analysis using variations on the tactics described in Eisenhardt (1989, pp. 540–541). The starting point for this analysis is the mechanisms for which the case studies did not provide any illustrations. The central questions are (a) which

characteristics do the four water boards have in common, (b) do these characteristics differ from the public sector characteristics that underpin our theoretical framework, and (c) can these divergent characteristics account for the irrelevance of particular mechanisms? Next, the analysis focuses on the mechanisms that are applicable to only one or two of the case water boards, while at the same time the other water boards produced only contradictory evidence. For each of these mechanisms, the water boards were divided into two groups: the water boards that did support this mechanism and the water boards that did not. Subsequently, the within-group similarities and across-group differences were listed, and I investigated whether these similarities and differences could explain the mixed empirical results. The use of these tactics resulted in a list of five explanatory factors. Some of these factors are closely related to concepts that are well-known from the literature (e.g., stakeholder involvement); other factors might be new to the field (e.g., ownership of the current performance).

Two factors appear to be important in explaining why certain mechanisms (i.e., Mechanisms 1c, 3a, 3b, 4a, and 6c) do not apply to all four water boards: stakeholder involvement and willingness to imitate. These sector-level factors refer to characteristics of the water board sector as a whole that limit the relevance of particular elements of the theoretical framework to this sector. Three other factors explain why certain mechanisms (Mechanisms 1a, 1b, 4b, and 5) are applicable to some but not all water boards: strength of managers, sector membership, and ownership of current performance. These organization level factors refer to characteristics of individual water boards that influence the extent to which particular elements of the theoretical framework are relevant to these organizations. Thus, given the characteristics of the water board sector and of the wastewater treatment benchmarking project, certain mechanisms might be relevant to all water boards, but the organization level factors influence the extent to which an individual water board is actually influenced by these mechanisms.

### *Stakeholder Involvement (Sector Level)*

In the wastewater treatment sector, stakeholder involvement—which is defined as the degree to which various stakeholders are concerned with the organization and try to influence its activities—is quite low. It appears that external stakeholders and internal stakeholders other than staff (i.e., governing bodies and members of these bodies) exert limited influence on the water boards' wastewater treatment decisions.

The low level of stakeholder involvement is related to specific characteristics of water boards in general and wastewater treatment in particular. Despite their often very long history, the water boards as a form of government are less known to the wider public than other forms of government (i.e., the central government and the provincial and municipal authorities). Moreover, the general public is less

familiar with the wastewater treatment task than with the other tasks of the water boards. This unfamiliarity might be due to two factors. First, the levies paid for wastewater treatment are relatively small compared to, for example, local taxes and charges for gas and electricity (Arthur Andersen & Vertis, 2001, p. 3). Second, no significant problems related to wastewater treatment have occurred recently. The governing bodies also seem to have relatively limited interest in wastewater treatment, compared with other water board tasks. This lack of interest may be related to this task's complex nature: It is often regarded as a rather technical task that can be understood only by professionals. Furthermore, members of the governing bodies view wastewater treatment as little more than managing sophisticated production facilities, which (in their opinion) requires no political choices. A final aspect that may explain the limited interest of the governing bodies is that, to a considerable extent, the objectives of wastewater treatment are uncontroversial as they are laid down in legal requirements. This narrowly defined task also reduces the need for intervention by the governing bodies.

The low level of stakeholder involvement in wastewater treatment explains why—in contrast to what we expected—the four water boards did not experience much pressure from their (external) stakeholders to use the benchmarking results to improve performance (Mechanism 4a), and why their decisions on action plans were not hindered by conflicts of interest between stakeholders (Mechanisms 3a and 3b). The low level of stakeholder involvement, therefore, had had negative as well as positive consequences for the water boards' use of benchmarking information.

#### *Willingness to Imitate (Sector Level)*

The water boards' willingness to imitate other organizations appeared to be rather limited. In general, the water boards seem to hesitate to consult better-performing organizations and imitate their business processes. This finding is somewhat surprising as the water board sector is highly interconnected. Managers and other employees from different water boards meet each other on various occasions, and individual water boards cooperate in several projects. Under these circumstances, we might expect that the barriers to exchanging information are not very high. However, the water boards appeared to have difficulty, not with acquiring information about better business processes, but with asking for—and subsequently using—such information. A factor that might explain this limited willingness to imitate is that the water boards do not experience much pressure to improve, which is related to the almost total lack of market competition (i.e., most customers have no alternatives to the services provided by their water board), and the low level of stakeholder involvement. This situation allows the water boards to operate relatively independently from similar organizations.

The limited willingness to imitate explains not only why, in contrast to what we expected, none of the four water boards had visited water boards with better

benchmarking scores to learn about their business processes (Mechanism 1c) but also why—despite its excellent benchmarking scores—Water Board A was not asked by any other water board to provide information about its business processes (Mechanism 1c). In addition, the limited willingness to imitate explains why the four water boards did not refer explicitly to benchmarking as a time- and cost-saving means of learning and innovating (Mechanism 6c). In contrast, the high cost of the benchmarking project was often brought up during the interviews. Thus, the case results suggest that the water boards did not explicitly use the relatively cheap and safe innovation opportunities that can become available through benchmarking. We can, therefore, conclude that the limited willingness to imitate had had a negative impact on the use of benchmarking information for performance improvement.

***Strength of Managers (Organization Level)***

An important difference between Water Boards A and C and Water Boards B and D is related to the strength of their managers. More precisely, the former two water boards appeared to have more powerful wastewater treatment managers than the latter two. In Water Board A, for example, the chairperson, the member of the executive body who was interviewed, and the secretary-director were all very positive about the water board's current performance and had great confidence in the expertise and motivation of the wastewater treatment managers. A comparable situation existed for Water Board C. The position of the wastewater treatment managers in Water Boards A and C seemed to minimize conflicts of interest. That is, these managers were able to balance the interests of different stakeholders (i.e., staff members and members of governing bodies with different backgrounds), to integrate these interests into a clear-cut strategy, and to apply this strategy. With respect to the benchmarking project, this situation implied that the various stakeholders agreed on the actions that had to be taken and on the sequence in which these actions had to be taken. A similar situation was clearly lacking in Water Boards B and D. The relative strength of their wastewater treatment managers, therefore, explains why, in contrast to Water Boards B and D, Water Boards A and C were not confronted with situations in which multiple, conflicting objectives hindered the rapid implementation of action plans (Mechanism 5).

***Sector Membership (Organization Level)***

A second difference between Water Boards A and C and Water Boards B and D concerns their degree of sector membership, a term that refers to the degree to which an organization identifies itself with sector-wide responsibilities. An important sector-wide issue in the water board sector concerns the future position of the water boards. In recent years, the best way to organize the water boards' activities has undergone considerable discussion. The alternatives that have been

suggested include placing the water boards under the direct supervision of the central government or dismantling the water boards and allocating their activities to other governmental organizations. These alternatives imply a serious threat to the water boards' autonomy and existence. The wastewater treatment benchmarking project was partly intended to deflect this threat. That is, through the benchmarking project, the water board sector wanted to demonstrate that it operates effectively and efficiently.

Water Boards A and C clearly act as "good" sector members who have collective responsibility to secure the legitimacy of the sector as a whole. Both these water boards were also strong advocates of the benchmarking project, even though they had already quite good information about their relative performance. They felt that the water board sector had an obligation to show to the outside world that it takes its performance seriously. For this reason, these water boards played an active role in the design phase of the benchmarking project. In addition, they took a serious look at their own benchmarking scores and followed up with necessary actions,<sup>4</sup> and they were very concerned that many other water boards had hardly acted on the benchmarking results. In contrast, the degree of sector membership in Water Boards B and D was much lower. That is, these water boards are normally not the driving forces behind sector-wide initiatives, including the wastewater treatment benchmarking project. Both water boards operate more or less as individual organizations that have their own responsibilities to meet the expectations of their stakeholders. Nevertheless, despite this similarity, Water Board B is much more open toward sector-wide developments than Water Board D. For example, whereas Water Board D strongly resisted the benchmarking project because it felt that it could not benefit from it, Water Board B had a much more positive attitude toward this project and was interested in what the organization could learn from it. In other words, whereas Water Board D to some extent prefers to isolate itself from sector-wide projects, Water Board B tries to benefit from them.

Water Boards A and C's high degree of sector membership, combined with the sector's legitimacy problem, explains why these water boards were motivated to use the benchmarking information to improve performance despite the absence of market incentives (Mechanism 1a) and stakeholder pressures to improve performance (Mechanism 4a). Moreover, as these water boards saw disclosure of the benchmarking scores as a means to gain legitimacy, it also partly explains why they did not try to defend their scores (Mechanism 4b). As such, the situation in Water Boards A and C is in sharp contrast with the situation in Water Board D, which provided clear illustrations of Mechanisms 1a and 4b. However, differences in the degree of sector membership did not result in differences in the use of benchmarking information between Water Boards A and C and Water Board B, as these three water boards all took their benchmarking scores fairly seriously.

### *Ownership of Current Performance (Organization Level)*

The differences between Water Boards B and D with respect to their use of the benchmarking information can be related to differences in their ownership of the current performance, a term that refers to the degree to which an organization's staff members regard themselves as being responsible for the organization's present performance level. Within Water Board B, this degree of ownership is relatively low. This water board was established only a few years before the benchmarking took place, and until the benchmarking results became available, the water board did not have a clear picture of its relative performance. In contrast, within Water Board D the degree of ownership of the current performance is rather high, which is due to the water board's long history (founded in the twelfth century) and large size. People were proud of being a member of such a large and important water board, and they felt that its status required it to have a position among the best-performing water boards. It was also their conviction that the water board's performance was quite good.

In Water Board B, the benchmarking project gave a first, rather disappointing, impression of the water board's relative performance, but it did not lead to defensive behavior because the low performance could be blamed on the water board's predecessors. It was argued that due to their (mis)management, the water board was now confronted with high levies and a limited capacity for treating wastewater. In response to the benchmarking information, the water board accelerated the implementation of its planned performance improvement actions, which is consistent with Mechanism 1b. Water Board D was also rather disappointed by its benchmarking scores, which seriously challenged the organization's perception of its performance. In this water board, however, the benchmarking scores led to the defensive behavior. Subsequently, an extensive use of the benchmarking information seemed unnecessary, which is consistent with Mechanism 4b. In other words, Water Board B regarded its disappointing benchmarking scores as a signal that its competitive position—which was regarded as a heritage from the past—had to be improved, whereas Water Board D regarded them as a false allegation of mismanagement. To some extent the interviewees in Water Board D were aware of the water board's defensive behavior. For instance, without explicitly referring to their own water board, some of them argued that benchmarking information may challenge complacency.

### *Summary of the Results for Water Boards A and C*

This analysis points at three factors that explain why the theoretical framework was of limited use in explaining Water Boards A and C's use of benchmarking information. First, because of a key characteristic of the Dutch water board sector (i.e., a low level of stakeholder involvement), Water Boards A and C's use of

benchmarking information (and that of other water boards) was not influenced to any great extent by external stakeholders. Second, the concerns about the water board sector's legitimacy stimulated Water Boards A and C to take the benchmarking information seriously. As such, these concerns partly replaced the roles that economic markets and an organization's stakeholders may have in imposing performance improvements. Finally, Water Boards A and C have powerful managers who are able to develop and implement a clear-cut strategy. Because of this characteristic, these water boards' internal stakeholders agreed on the action plans that had to be developed and implemented in response to the benchmarking information. As a result, the development and implementation of action plans was not hindered, despite the presence of multiple, conflicting objectives.

### **Conclusions and Discussion**

This paper examines whether—and if so, under which circumstances—public sector organizations are likely to use benchmarking information for performance improvement, focusing on the influence of specific characteristics of public sector organizations. A theoretical analysis shows that, although some of these characteristics may have a positive influence, many characteristics of public sector organizations can be expected to affect the use of benchmarking information negatively. Nevertheless, despite the numerous hindering characteristics, the empirical research reveals that information from a wastewater treatment benchmarking project was, to varying degrees, used by Dutch water boards. This result suggests that public sector organizations can overcome the barriers to the use of benchmarking information. However, we must acknowledge that the Dutch water board sector and its benchmarking project have specific characteristics, which probably have influenced the empirical results. Still, we can draw some conclusions from this empirical research that can be relevant in a broader context.

The first conclusion concerns the substitutes for market incentives that are available in the public sector. The empirical research shows that, with a low exposure to economic markets, public sector organizations may ignore information (including benchmarking information) that indicates that their relative performance is poor. The theoretical analysis discusses an alternative to market incentives, namely, that public scrutiny can result in a situation in which internal and external stakeholders pressure public sector organizations to improve their relative performance. In the Dutch water board sector, however, such pressures were not very strong. A more significant source of pressure was the pressure on the sector as a whole to demonstrate that water boards are the best mode of delivery of the services in question. This pressure may, in fact, influence the behavior of individual water boards but only if these organizations identify themselves with sector-wide responsibilities. In other segments of the public sector, however, stakeholders may be more closely

involved with individual organizations, and the pressures they exert to improve performance may be more severe. In such a setting, it might be less important that organizations assume a role as responsible sector members. Some segments of the public sector are actually (increasingly) exposed to economic markets, in the sense that their customers have the right to select their own suppliers. Public sector organizations then probably do not need stakeholder pressure or sector-wide legitimacy problems to have an incentive to maintain or improve their relative performance, because underperformance may have a direct negative impact on the organization's revenues.

The wastewater treatment benchmarking project also showed that public sector organizations may have difficulties imitating superior business processes. Due to these difficulties, they may not be able to realize the full potential that benchmarking may have, especially in a context in which clear market signals are missing. For this reason, public sector organizations might consider institutionalizing the imitation process, which would imply that cooperation among the benchmarking partners is not only embedded in the performance measurement and analysis stage of a benchmarking project (as it was in the wastewater treatment benchmarking project), but also in the performance improvement stage. Organizations might, for example, be stimulated to investigate which specific business processes underpin the better benchmarking scores.

The final conclusion is related to the role of managers and other staff. On the one hand, the empirical research suggests that staff members should focus on improving performance rather than on defending it. Such an improvement-oriented attitude requires developing a benchmarking culture. Implicitly, the empirical research referred to all aspects of a benchmarking culture, including a recognition that one's own organization is probably not the best in all aspects of operations and a resistance to tendencies to defend one's own (underperforming) operations. In other words, a benchmarking culture is of great importance, not only in private sector benchmarking but also in public sector benchmarking. On the other hand, the empirical research stresses the importance of having powerful managers, which is related to public sector organizations' multiple, diverse, and often conflicting objectives. In the Dutch water board sector, the presence of such objectives caused delays in the implementation of action plans, because various stakeholders did not agree on the question of which priority should be given to individual plans. Powerful managers were able to prevent such delays by developing, gaining support for, and applying a clear, guiding strategy that balances the various interests that are involved. However, it should be noted that the objectives of the wastewater treatment sector are rather simple. Roughly speaking, the effectiveness of wastewater treatment (i.e., the percentage of waste products that are removed from the wastewater) must be weighed against the costs of such a level of effectiveness. In other parts of the public sector, objectives are probably

more complex in the sense that the quality level of a particular service must be compared not only with the costs incurred but also with the quality levels of other, quite different services. Under these circumstances, it will be much more difficult to develop and apply a clear-cut strategy.

This study attempts to improve our understanding of the use of benchmarking information in a public sector context. It provides an overview of various mechanisms through which public sector characteristics might influence the use of benchmarking information, and it explores the circumstances under which certain mechanisms can or cannot be expected. However, because the empirical exploration is restricted to a rather specific setting, further research in the interesting area of public sector benchmarking is needed. It would be worthwhile to investigate the applicability of the theoretical framework in empirical settings with, for instance, a higher level or stakeholder involvement, a higher level of market exposure, or more complex objectives. Interesting areas include the local government sector, health care, and education. Hopefully, this study will be a valuable starting point for that type of research.

### Notes

1. Although some authors (e.g., Walgenbach & Hegele, 2001, pp. 124–126) have taken the view that the identification of “best practice” is the basic principle of benchmarking, Bowerman et al. (2002, pp. 433–434) questioned the significance of this principle in a public sector context.

2. However, Bowerman et al. (2002) noted that this form of benchmarking was being used in the public sector before it was popularized in the private sector.

3. For a critical evaluation of the empirical evidence of the differences between public and private sector organizations, see Boyne (2002).

4. After the publication of the benchmarking results, Water Board C carefully analyzed these results. Based on this analysis, the board decided that taking further improvement actions would be inconsistent with the water board’s low-cost strategy (i.e., a strategy to meet legal wastewater treatment standards at the lowest possible cost level). Therefore, this strategy explains why Water Board C used the benchmarking information only to a limited extent even though, at the same time, it attached great importance to the benchmarking project itself.

### References

- Ammons, D.N. (1999). A proper mentality for benchmarking. *Public Administration Review*, 59(2), 105–109.
- Arthur Andersen & Vertis. (2001). *Zuiver afvalwater: Bedrijfsvergelijking zuiveringsbeheer 1999* [Purified wastewater: Wastewater treatment benchmarking, 1999]. The Hague: Unie van Waterschappen.
- Bowerman, M., & Ball, A. (2000). Great expectations: Benchmarking for best value. *Public Money & Management*, 20(2), 21–26.
- Bowerman, M., Ball, A., & Francis, G. (2001). Benchmarking as a tool for the modernisation of local government. *Financial Accountability & Management*, 17(4), 321–329.
- Bowerman, M., Francis, G., Ball, A., & Fry, J. (2002). The evolution of benchmarking in U.K. local authorities. *Benchmarking: An International Journal*, 9(5), 429–449.

- Boyne, G.A. (2002). Public and private management: What's the difference. *Journal of Management Studies*, 39(1), 97–122.
- Camp, R.C. (1995). *Business process benchmarking: Finding and implementing best practices*. Milwaukee: ASQC Quality Press.
- Cyert, R.M., & March, J.G. (1959). A behavioral theory of organizational objectives. In M. Hair (Ed.), *Modern organization theory* (pp. 76–90). New York: Wiley.
- Davies, P. (1998). The burgeoning of benchmarking in British local government: The value of “learning by looking” in the public services. *Benchmarking for Quality Management and Technology*, 5(4), 260–270.
- 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.
- Dorsch, J.J., & Yasin, M.M. (1998). A framework for benchmarking in the public sector: Literature review and directions for future research. *International Journal of Public Sector Management*, 11(2/3), 91–115.
- Drew, S.A.W. (1997). From knowledge to action: The impact of benchmarking on organizational performance. *Long Range Planning*, 30(3), 427–441.
- Eisenhardt, K.M. (1989). Building theories from case study research. *Academy of Management Review*, 14(4), 532–550.
- Elnathan, D., Lin, T.W., & Young, S.M. (1996). Benchmarking and management accounting: A framework for research. *Journal of Management Accounting Research*, 8, 37–54.
- Fedor, D.B., Parsons, C.K., & Shalley, C.E. (1996). Organizational comparison processes: Investigating the adoption and impact of benchmarking-related activities. *Journal of Quality Management*, 1(2), 161–192.
- Greenwood, R., & Hinings, C.R. (1996). Understanding radical organizational change: Bringing together the old and the new institutionalism. *Academy of Management Review*, 21(4), 1022–1054.
- Groot, H. de, Goudriaan, R., Hoogwout, M., Jong, A. de, & Poerstamper, R. (2004). *Benchmarking in de publieke sector: Vergelijken van prestaties als management instrument* [Benchmarking in the public sector: Comparing performance as a management tool]. Part 6. The Hague: SDU Uitgevers.
- Helden, G.J. van, & Admiraal, R.J. (2002). Exploring the balancing concept in the design of a performance measurement system for waste water treatment in the Netherlands. In V. Montesinos & J.M. Vela (Eds.), *Innovations in governmental accounting* (pp. 247–263). Boston: Kluwer Academic.
- Helden, G.J. van, & Tillema, S. (2005). In search of a benchmarking theory for the public sector. *Financial Accountability & Management*, 21(3), 337–361.
- Hofstede, G. (1981). Management control of public and not-for-profit activities. *Accounting, Organizations and Society*, 6(3), 193–211.
- Hood, C. (1991). A public management for all seasons? *Public Administration*, 69(1), 3–19.
- Hood, C. (1995). The “new public management” in the 1980s: Variations on a theme. *Accounting, Organizations and Society*, 20(2/3), 93–109.
- Kaplan, R.S., & Norton, D.P. (1992). The balanced scorecard: Measures that drive performance. *Harvard Business Review*, 70(1), 71–79.
- Kaplan, R.S., & Norton, D.P. (1996). *The balanced scorecard: Translating strategy into action*. Boston: Harvard Business School Press.
- Llewellyn, S., & Northcott, D. (2005). The average hospital. *Accounting Organizations and Society*, 30(6), 555–583.
- Meyer, J.W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology*, 83(2), 340–363.

- Miller, J.G., De Meyer, A., & Nakane, J. (1992). *Benchmarking global manufacturing: Understanding international suppliers, customers, and competitors*. Homewood, IL: Business One Irwin.
- Northcott, D., & Llewellyn, S. (2003). The "ladder of success" in healthcare: The UK national reference costing index. *Management Accounting Research*, 14(1), 51–66.
- Nutt, P.C. (2000). Decision-making success in public, private and third sector organizations: Finding sector dependent best practice. *Journal of Management Studies*, 37(1), 77–108.
- Nutt, P.C., & Backoff, R.W. (1993). Organizational publicness and its implications for strategic management. *Journal of Public Administration Research and Theory*, 3(2), 209–231.
- Rainey, H.G. (1989). Public management: Recent research on the political context and managerial roles, structures, and behaviors. *Journal of Management*, 15(2), 229–250.
- Rainey, H.G., Backoff, R.W., & Levine, C.H. (1976). Comparing public and private organizations. *Public Administration Review*, 36(2), 233–244.
- Simon, H.A. (1964). On the concept of organizational goal. *Administrative Science Quarterly*, 9(1), 1–22.
- Spendolini, M.J. (1992). *The benchmarking book*. New York: AMACOM.
- Tillema, S., & Helden, G.J. van (2005). Appreciation, acceptance and use of multidimensional benchmarking information by public sector organizations. In A. Bourmistrov & F. Mellemvik (Eds.), *International trends and experiences in government accounting* (pp. 230–247). Oslo: Cappelen Akademisk Forlag.
- Unie van Waterschappen. (1999). *Water boards*. 7th ed. The Hague.
- Walgenbach, P., & Hegele, C. (2001). What can an apple learn from an orange? Or, what do companies use benchmarking for? *Organization*, 8(1), 121–144.
- Wildavsky, A.B. (1989). *Budgeting: A comparative theory of budgetary processes*. 2nd ed. New Brunswick, NJ: Transaction.

*Sandra Tillema is an assistant professor of finance and accounting at the University of Groningen. Her research interests are accounting change and public management.*