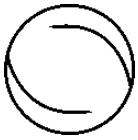

Article

Organization Studies
2015, Vol. 36(7) 953–978

DOI: 10.1177/0170840615580014
www.egosnet.org/os

Management Control, Results-Oriented Culture and Public Sector Performance: Empirical Evidence on New Public Management

Frank H. M. Verbeeten
VU University Amsterdam/Utrecht University School of Economics, The Netherlands

Roland F. Speklé
Nyenrode Business University, The Netherlands

Abstract

New Public Management (NPM) has been guiding public sector reform for over 25 years. Its position on the design of effective management control rests on three key ideas: (1) performance improvement requires a results-oriented culture that emphasizes outcomes rather than inputs or processes; (2) public sector organizations need to introduce performance management based on targets, monitoring and incentives; and (3) public sector organizations should decentralize decision rights and reduce their reliance on rules and procedures. Focusing on the particularly influential version of NPM as advocated by the OECD, we examine the validity of these ideas theoretically and empirically. We conclude that NPM's reform programme should be reconsidered. Although the evidence indicates that a results-oriented culture is positively associated with performance, we find little support for the assumed benefits of NPM-type performance contracting. In addition, the results suggest that both the effects of decentralization and the reliance on rules and procedures are opposite to NPM's expectations.

Keywords

agency theory, civil society, corporate governance, nonprofit organizations, organizational control, structural equation modelling

Introduction

Since the 1980s, the public sector has been subjected to reform initiatives in the spirit of New Public Management (NPM) to address concerns about its performance. The term NPM refers to a broad set of beliefs, doctrines and codified experiences that collectively serve as a frame of

Corresponding author:
Frank H.M. Verbeeten, Utrecht University School of Economics, Faculty of Law, Economics and Governance, PO Box 80125, 3508 TC Utrecht, The Netherlands.
Email: f.h.m.verbeeten@uu.nl; f.h.m.verbeeten@vu.nl

reference in the evaluation and redesign of the public sector. This frame of reference varies over time and across countries (Hood, 1995; Pollitt, 2002) and its enactment is mediated by context and human agency (Thomas & Davies, 2005). Local and temporal variety notwithstanding, NPM is characterized by a quest for economic rationality and a strong trust in market-based discipline, managerial control and hands-on management (Arellano-Gault & Lepore, 2011; O'Reilly & Reed, 2011; Townley, Cooper & Oakes, 2003).

In this paper, we examine the validity of a specific branch of NPM, i.e. the part that is concerned with the design of management controls to support the effectiveness and efficiency of public sector organizations. To circumvent the problem of variety in NPM, we rely on OECD reports to identify the basic and common components of this branch. These reports draw on the views and experiences of a large number of member states and reflect a consensus on NPM thinking among OECD members that have been engaged in the reform process. We shall refer to this version as OECD-NPM. The gist of OECD-NPM thinking is sometimes expressed in two characteristic slogans: 'let the manager manage' and 'make the manager manage' (OECD, 1997). OECD-NPM believes that managers in the public sector should be freed from traditional bureaucratic constraints and input controls to allow them to manage more effectively ('let the manager manage'; Laegreid, Roness & Verhoest, 2011; Nielsen, 2013). These constraints, it is argued, should be replaced with a performance contracting system in which targets and incentives ensure that the newly-freed managers continue to act in the organization's interests ('make the manager manage'; Burgess & Ratto, 2003; Dewatripont, Jewitt & Tirole, 1999). Rooted in a managerialist notion of control (O'Reilly & Reed, 2011), OECD-NPM sees performance measurement systems as vital catalysts for performance improvement (OECD, 1994, 1997, 2002). However, the relationship is not direct. Rather, performance measurement systems are held to support the emergence of a 'results-oriented culture' which, in turn, should contribute to performance (OECD, 1994, 1997; see also Hood, 1991; Parker & Bradley, 2000; Laegreid et al., 2011).

After almost three decades of reforms, the effect of NPM on public sector performance is still controversial (Bejerot & Hasselbladh, 2013). Even the OECD itself admits that results of the reform programmes remain unclear, recognizing that 'many important public management changes that have been presented as best practices later appeared to develop undesired and unanticipated consequences' (OECD, 2002, p. 27). Formal evidence on the performance effects of organizational design choices is still limited (Bejerot & Hasselbladh, 2013) and larger-scale empirical studies are especially scarce (Nielsen, 2013; Taylor, 2011; Van Helden, 2005). Moreover, the empirical literature only addresses parts of NPM's approach to internal governance effectiveness (Diefenbach, 2009), ignoring NPM's view on internal deregulation, i.e. the role of bureaucratic constraints, rules and procedures, and decentralization of decision rights. Furthermore, even though OECD-NPM emphasizes the accountability and incentive role of performance measurement, it explicitly allows performance information to serve a broader set of purposes, in particular, to support learning. This other role has received scant attention in empirical studies (Frey, Homberg & Osterloh, 2013; Moynihan & Pandey, 2005). Finally, the literature has modelled only the direct effect of performance measurement systems on performance, neglecting the mediating role of a results-oriented culture as suggested by OECD-NPM.

In our study, we fill this gap by developing a richer model in which we account for NPM ideas on both performance measurement and internal deregulation. In addition, we follow OECD-NPM and acknowledge that the role of performance measurement extends beyond accountability and incentive provision. Finally, we take into account the claims as to the mediating effect of a results-oriented culture. We confront this model with the relevant literature to assess its consistency with theory and empirical findings, and to provide a theoretical context to help interpret the results. We test the model empirically, using data from a survey of financial managers in 96 municipalities in

the Netherlands. Since the late 1980s, Dutch municipalities have been under constant pressure to adopt NPM-based thinking in management and control, while retaining sufficient autonomy to resist or mould this pressure. Therefore, this sample spans the spectrum from true NPM adopters to organizations that preserved traditional modes of internal governance, making it uniquely informative for our purposes.

Our study casts doubt on the validity of the OECD-NPM approach to management control. Based on the confrontation with the literature, we conclude that this approach lacks theoretical substantiation and that many of its propositions are inconsistent with extant theory and prior empirical work. Our own empirical findings do not support OECD-NPM either. Although we find a positive association between the degree of results orientation and performance, the use of performance measurement systems for accountability and incentives does not appear to affect such a culture. There is a positive effect of the use of performance measures to communicate direction, but this particular use is not part of OECD-NPM's control repertoire. In addition, whereas OECD-NPM assumes a negative association between the reliance on rules and procedures and a results-oriented culture, our results indicate that this effect is in fact positive. Finally, we find that even though the association between decentralization and results-oriented culture is broadly consistent with OECD-NPM, the ultimate net effect on performance is negative. The overall conclusion, therefore, is that the OECD-NPM reform agenda should be reconsidered.

The remainder of this paper is structured as follows. In the second section, we review the relevant OECD documents to specify the OECD-NPM working theory of effective management control in the public sector, and we confront the resulting model with the literature. Next, we provide details on our methods and operationalize the model. Then we report on the empirical analysis, and the final section presents conclusions and a discussion of the findings.

Literature Review

We are interested in a theoretical and empirical assessment of the operative set of NPM-related ideas that have actually played a role in the reform and redesign of public sector organizations, and that have had a real impact on organizational practice. To identify the key components of this set of ideas, we rely on OECD documents. Since the emergence of NPM, the OECD has been a strong supporter of the associated reforms. From the late 1980s onwards, the OECD organized a series of conferences and seminars to bring together experts and senior government representatives from its member states involved in internal governance reform. The purpose of these meetings was to share experiences and 'best practices' to help member states implement and refine their reform programmes. In addition, the OECD commissioned several comparative studies of internal governance across member states, also with the aim of promoting 'good governance practice'. These studies and the reports of the expert meetings have no regulatory status, but are nevertheless meant to provide authoritative guidance. The documents are quite explicit about the how and why of ongoing reform programmes and the expected effects of various control choices, allowing us to specify the common threads in NPM-type thinking and to explicate a working theory of management control design as it has influenced actual public sector reforms.

NPM's working theory as it emerges from the OECD sources ('OECD-NPM') is a normative theory. Rather than trying to explain the design and functioning of internal governance, it seeks to change it. The proposed model rests on three related pillars (Laegreid et al., 2013) that form an integrated package of reforms (Arellano-Gault & Lepore, 2011; Brunsson & Sahlin-Andersson, 2000). First, performance improvement requires a cultural change. A new results-oriented culture should replace the preoccupation with rules and procedures that hitherto prevailed, and should ensure a focus on outcomes and results. Second, public sector employees need to be made

accountable for performance to encourage them to think in terms of results. This requires the use of performance information to define outcome targets, to monitor operations and policy implementation, and to provide incentives that are linked to performance. Third, public sector organizations should accomplish internal deregulation to empower managers to take appropriate action in the pursuit of performance.

In the following sections, we discuss these pillars more fully and specify the key elements of OECD-NPM in the form of testable propositions. We then proceed to examine these propositions, relying on insights from academic literature.

OECD-NPM as a 'working theory' of management control

Results-oriented culture and organizational performance. According to the OECD, the first step towards performance improvement is to create a 'results-oriented culture' within public sector organizations (OECD, 1994, 1997, 2002, 2005). This step is considered essential, and the difficulties in achieving such a culture are seen as the main reason why progress of improvement programs has been slow (OECD, 2005). The OECD reports provide no real definition of a results-oriented culture, but hold several clues as to its meaning. For instance, the notion of results orientation is associated with managers taking responsibility for results that are valued by society (OECD, 1994, 1997), a focus on 'the benefits that arise from one's actions rather than on processes or what one does' (OECD, 2002, p. 4), an 'awareness of costs, deadlines, quality' (OECD, 1994, p. 76), and an atmosphere in which people are willing to question their action choices, 'asking themselves what the benefits of their activities are supposed to be …, and what they can do to be more effective' (OECD, 2002, p. 13). A results-oriented culture is contrasted with the traditional bureaucratic model, which values compliance to pre-set rules and regulations (OECD, 1994, 1997, 2002). According to OECD-NPM, this traditional model is no longer tenable and should be replaced with a focus on 'big picture outcomes' (OECD, 2002, p. 4). Capturing this OECD-NPM idea in a proposition:

P1: A results-oriented culture contributes positively to organizational performance.

Performance measurement systems and results-oriented culture. In the OECD-NPM perspective, performance measurement systems are powerful catalysts for the aspired cultural change (OECD, 1994, 1997, 2002). Managers need to be made accountable for performance to encourage them to think in terms of results and to make them realize that mere compliance with rules and procedures is no longer sufficient (OECD, 1994, p. 19). This, it is argued, requires the use of accounting information to define performance targets, to monitor and evaluate the efficiency and effectiveness of operations and programmes against these targets, and to reward (penalize) good (bad) performance (OECD, 1994, 1997). OECD-NPM thus emphasizes the role of accounting in creating an internal system of quasi-contracting in which explicit and measurable pre-set performance targets should guide civil servants' efforts towards the achievement of organizational objectives. In fact, this role has become part of the very definition of performance management within OECD-NPM, i.e. 'a set of procedures for defining performance, measuring it, and linking it to incentives or sanctions' (OECD, 2007, p. 8).

Even though professional pride and the intrinsic desire to meet performance expectations are mentioned by the OECD as potential motivators (OECD, 1994, 2002, 2007), the importance of extrinsic incentives in its thinking can hardly be overstated: 'performance measures are only relevant if there are consequences for over- or underperformance against targets' (OECD, 1997, p. 22), and 'effective control cannot be exercised without a system of rewards and penalties' (OECD, 1994, p. 85). The incentives can take different forms, including pay-for-performance compensation

schemes and performance-based promotion structures (OECD, 2007). OECD-NPM acknowledges that strong incentives may lead to gaming, goal displacement and data manipulation (OECD, 2002, 2005, 2007). These dysfunctional consequences, however, are not sufficient to abandon the idea of performance-related pay and promotion. Rather, OECD-NPM takes them to indicate that current practice needs further refinement to reduce the negative effects, while retaining 'the benefit of well-structured performance management arrangements' (OECD, 2007, p. 13). Casting these views in the form of testable propositions:

P2: The use of performance measurement systems for accountability purposes supports a results-oriented culture.

P3: To achieve a results-oriented culture, performance information should be tied to managers' pay and promotion prospects.

Despite the emphasis on incentives and accountability, OECD-NPM does not entirely ignore other roles of performance information. Particularly, there are occasional references to a learning support role of performance measurement. In this role, performance information is used to provide feedback on the consequences of specific action choices to acquire a deeper understanding of the causes of failure and success, and to identify ways to improve performance (OECD, 2002, pp. 15–16). Relatedly, performance measures may help to focus attention on programs and policy areas that are in need of revision (OECD, 2002: 23). Therefore:

P4: The use of performance measurement systems to support learning contributes to a results-oriented culture.

Internal deregulation and results-oriented culture. Up to this point, the discussion has centred on the apparatus to 'make the manager manage'. However, to create the results-oriented culture needed for performance improvement, it is equally important to 'let the manager manage' (OECD, 1997). According to OECD-NPM, this requires public sector organizations to relax the bureaucratic constraints and the rules and regulations that typify traditional modes of governance. These action controls are inimical to an orientation on results: 'it would be very difficult to talk about results and performance with managers whose agencies are rigidly governed by rules and regulations' (OECD, 1994, p. 21). Accordingly:

P5: A reliance on rules and procedures decreases the organization's orientation on results (i.e., subtracts from a results-oriented culture).

The removal of rules and regulation, however, is not sufficient for managerial empowerment. Internal deregulation also requires decentralization of decision rights to ensure that managers have not just the tools, but also the power, to achieve what is required (OECD, 1994, p. 56). Benefits of decentralization include better access to local, specialized knowledge about the impact of action choices, higher responsiveness to local conditions and client expectations, increased motivation at lower organizational levels, and economizing on senior management time (OECD, 1994, p. 59).

OECD-NPM distinguishes between operational and strategic autonomy.[1] Operational autonomy refers to decentralization of decision rights in operational matters, human resource management and internal organization. Strategic autonomy denotes definition of goals and priority setting in policy affairs and the choice of service delivery methods (OECD, 1994, p. 61). Operational autonomy is desirable in OECD-NPM thinking, as it should provide increased flexibility and

productivity gains (OECD, 1994, pp. 61–2; OECD, 1997). The OECD's position on strategic autonomy, however, is ambivalent. On the one hand, decentralized input in policy development may be valuable to capture experience from the field. But the involvement of local operating units in policy development would imply the introduction of a political dimension in these units' operations. This would erode the orientation on results, as 'efforts are redirected towards influencing the selection of performance targets rather than achieving them' (OECD, 2007, p. 11). OECD-NPM suggests that this downside more than offsets the benefits associated with the better use of local knowledge in policy development. Thus:

P6: Decentralization of decision rights in operational matters contributes to a results-oriented culture.

P7: Decentralization of decision rights in strategic matters diminishes the organization's orientation on results (i.e. subtracts from a results-oriented culture).

The OECD-NPM model. Figure 1 illustrates NPM's working theory of effective public sector management control as it emerges from the OECD documents.

Confronting the OECD-NPM model with theory

The model depicted in Figure 1 represents a practitioners' framework rather than an academically well-grounded theory. In the present section, we confront this model with the relevant literature to assess whether it is consistent with current theorizing and extant empirical work.

Results-oriented culture and organizational performance. OECD-NPM's view of organizational culture is rather instrumental and suggests that culture can actively be managed to arrive at some desired end. This view may well be naïve. The more usual perspective in the literature is that culture is resilient and stable, and difficult to change in purposive ways (e.g. Rodrigues, 2006; Schein, 1984). In a field study of cultural change programmes in public sector organizations, Parker and Bradley (2000) found evidence that there are indeed limits to the manageability of culture and that organizational culture is quite robust, even if the change agenda enjoys strong management support.

The link between organizational culture and performance has mostly been studied in a normative research tradition (Garnett, Marlowe & Pandey, 2008), and empirical evidence on this relationship is limited (Wilderom, Glunk & Maslowski, 2000). Garnett et al. (2008) find a positive relation between a results-oriented culture and organizational performance, but show that this relationship is mediated by task instructions, feedback information, and communication about performance problems. Thus, whereas OECD-NPM positions culture as a mediator in the relationship between performance information and performance, the results of Garnett et al. (2008) suggest that performance information mediates the relationship between culture and performance. Moynihan and Pandey (2005) find that organizations with a developmental culture (i.e. a culture that focuses on the needs of the organization and the ability to change to meet new demands) are likely to achieve higher levels of effectiveness in the eyes of their employees. More generally, it is plausible to assume that different cultures may emphasize different dimensions of effectiveness (Cameron & Freeman, 1985). Therefore, it is ultimately an empirical question whether a results-oriented culture will increase organizational effectiveness.

Performance measurement systems and results-oriented culture. The unconditional belief of OECD-NPM in the power of performance contracting meets uneasily with the literature. Both the behavioural and economics literature suggest that explicit performance contracting is only viable in

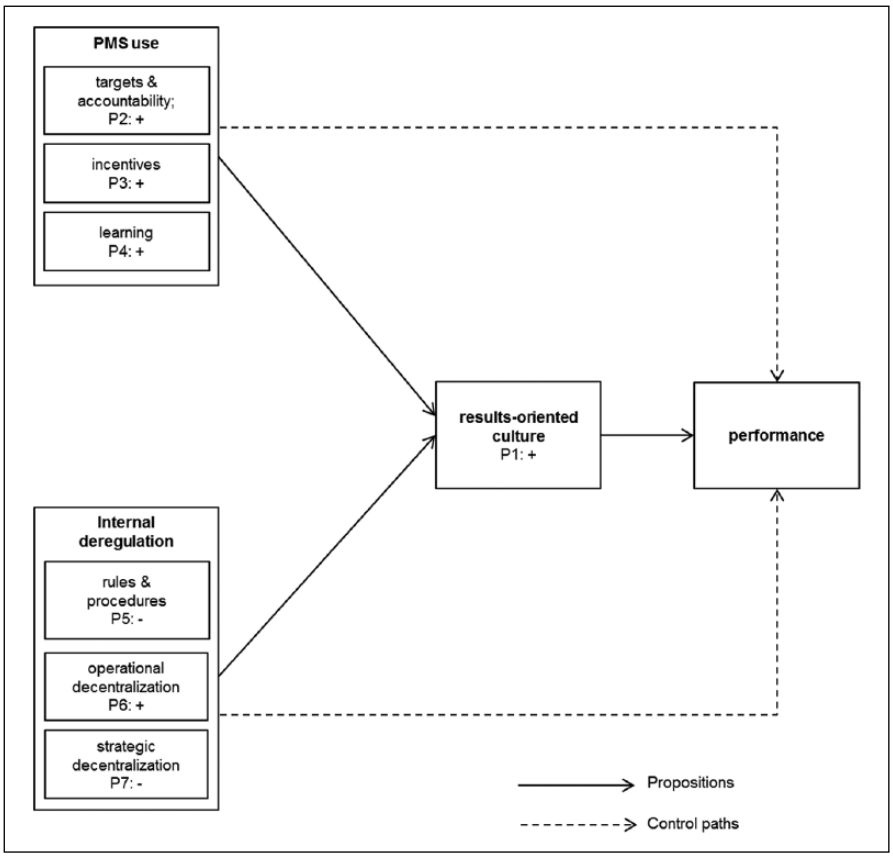

Figure 1. The OECD-NPM theory of management control.

specific situations. Hofstede (1981) argues that performance contracting requires unambiguous objectives, measureable outputs, repetitive activities and prior knowledge of the effects of management interventions. Economists take a similar position, emphasizing the importance of contractibility and output measurability for performance contracting (Burgess & Ratto, 2003; Propper & Wilson, 2003; Tirole, 1994). Overall, the theoretical consensus is that performance contracting requires that: (1) goals can be specified unambiguously in advance; (2) the organization is able to select undistorted performance measures, i.e. metrics that provide incentives that are adequately aligned with the organization's ultimate objectives; and (3) organizational actors know and control the production function that transforms efforts into results, and are able to predict the likely outcomes of alternative courses of action (Feltham & Xie, 1994; Gibbons, 1998; Hofstede, 1981; Speklé & Verbeeten, 2014). Empirical studies have in fact shown that a strong emphasis on targets and incentives may result in dysfunctional behaviour when goals are ambiguous and difficult to measure (e.g. Bevan & Hood, 2006; Newberry & Pallot, 2004; Speklé & Verbeeten, 2014). Ambiguous goals are common in the public sector (Burgess & Ratto, 2003; Dixit, 2002; Tirole, 1994), which hinders selection of appropriate performance measures (Cavalluzzo & Ittner, 2004). The public sector cannot revert to an uncontested and timeless performance measure similar to profits or shareholder value in private firms, and performance management in the public sector is never fully devoid of politics (Arellano-Gault, Demourtain, Rouillard & Thoening, 2013).

OECD-NPM is well aware of these problems (see OECD, 2002;,2005, 2007), but apparently does not see them as prohibitive. It should, however, be noted that whereas the academic literature has mostly examined the direct effects of the use of performance measurement systems on managerial behaviour and organizational outcomes (see Franco-Santos, Luncianetti & Bourne, 2012 for an overview), OECD-NPM postulates an indirect effect in which a results-oriented culture mediates the relationship between performance measurement system use and performance. Therefore, even though the use of performance information for target-setting, accountability and incentives may result in gaming and sub-optimization (as most of the literature suggests), it is still possible that the positive effects of an increase in results orientation outweigh the costs associated with the unintended side effects. This claim has no precedent in the academic literature. The only study we know of that comes close is Abernethy and Vagnoni (2004), who explore the effects of performance information on cost consciousness (a potential element of a results-oriented culture) of physician managers in Italian hospitals. Their findings, however, indicate that the use of performance information for control purposes has no effect on cost consciousness.

OECD-NPM's recognition that performance information may be used for various different purposes is consistent with a growing literature that studies the different roles of performance measurement systems (e.g. Hansen & Van der Stede, 2004; Henri, 2006). These studies rely on different classifications to operationalize these roles, but they all build on the classical distinction between decision-facilitating and decision-influencing uses of performance measurement (Demski & Feltham, 1976). The decision-influencing role refers to the use of information for motivating and controlling managers and employees (i.e. the primary focus of NPM), while the decision-facilitating role refers to the provision of information to guide decisions and managerial action (Grafton, Lillis & Widener, 2010; Van Veen-Dirks, 2010). Using the performance measurement system for decision-facilitating purposes leads to a shared frame of reference as to what goals should be achieved, what constitutes satisfactory performance, and what input is required to achieve the strategic goals (Burchell et al., 1980; Mundy, 2010; Simons, 1995). The decision-facilitating role includes the learning support function as identified by OECD-NPM (Frey et al., 2013), and also includes functions such as strategy formation and communication of goals (Hansen & Van der Stede, 2004; Henri, 2006). These latter functions are not part of current OECD-NPM, but may be key to a sound understanding of the effects of performance measurement system use in public sector organizations.

Empirical studies have focused on the determinants of the roles of the performance measurement system, rather than on the effects of the different uses on (intermediate variables leading to) performance (Hansen & Van der Stede, 2004; Henri, 2006). A general conclusion of this literature is that the decision-facilitating role is especially important when objectives are complex and ambiguous, when adaptation to changing circumstances is important, and when the organization cannot fall back on routines to achieve its goals. Because public sector organizations often operate in these conditions, a decision-facilitating use of performance measurement systems may be expected to have value in this sector (Frey et al., 2013). Speklé and Verbeeten (2014) provide empirical evidence to support this expectation. Their findings indicate that an exploratory, decision-facilitating use of performance information is associated with higher performance in the public sector. As OECD-NPM does not recognize this type of use in its repertoire, it may have incorrectly attributed the positive effects of performance measurement systems to target setting, accountability or incentives when, in fact, these results should be ascribed to the broader decision-facilitating use of the system.

Internal deregulation and results-oriented culture. The attitude of OECD-NPM towards rules and procedures is markedly disapproving. In this view, rules and procedures create bureaucratic

rigidity, frustrate responsiveness, stifle creativity and are inimical to a focus on results (Laegreid et al., 2011; Moynihan & Pandey, 2005; Nielsen, 2013). This unfavourable assessment, however, seems to overlook that controls that constrain behaviour may also have positive outcomes in terms of organizational goal achievement (Adler & Borys, 1996; Mundy, 2010). For example, part of the purpose of rules and regulations is to inform employees on the acceptable domain of opportunity-seeking behaviour, stimulating employees to direct their efforts towards areas where they are most valuable (Simons, 1995). Furthermore, rules and standard operating procedures may be the result of years of experimentation and may communicate best practices that have surfaced over time. Accordingly, rules and procedures may facilitate task performance rather than create bureaucratic red tape (Adler & Borys, 1996), and NPM's position may well be too negative.

OECD-NPM's plea for operational decentralization rests on three implicit assumptions: (1) that managers in the public sector possess unique and valuable local knowledge that is not available at higher levels in the organizational hierarchy and that cannot be transferred at acceptable costs; (2) that the quality of performance measurement is sufficient to ensure goal-consistent behaviour; and (3) that coordination requirements do not interfere. These assumptions may be unwarranted. Information asymmetries and differences in relative expertise (Jensen & Meckling, 1992) are likely to exist, but are not necessarily large enough to favour decentralization. Decentralization may result in opportunistic behaviour and additional coordination costs. Therefore, decentralization requires the organization to invest in a performance measurement system to allow monitoring, to secure coordination and to provide incentives to ensure that managers' actions are consistent with the goals of the organization (Abernethy, Bouwens & Van Lent, 2004; Jensen & Meckling, 1992). We have already argued that this is difficult in the public sector. Furthermore, NPM has been criticized for ignoring the cost of coordination and its failure to account for the interdependencies between agencies and departments (e.g. Mascarenhas, 1993). Thus, OECD-NPM may well overstate the benefits of decentralization.

Intermediate conclusion. Overall, our conclusion is that the OECD-NPM approach finds only mixed support in the academic literature. Even though its propositions can claim some basis in the literature, none of them is fully consistent with current theorizing in the relevant fields of study, and none can fall back on truly solid empirical support from previous research. For the purposes of our study, however, we temporarily ignore the discrepancies and frictions, and first confront the unamended OECD-NPM model with the empirical data, allowing OECD-NPM to speak for itself. We will come back to the alternative perspectives and counter-arguments from the literature when interpreting and discussing the empirical findings.

Research Design, Sample and Measurement

Institutional setting[2]

This study is based on survey data collected from senior financial managers in municipalities in the Netherlands. The Dutch governmental system is organized as three different layers: central government, provinces and municipalities. The system can be characterized as a decentralized unitary state. This means that the relations between the layers are hierarchical: municipalities have to adhere to the policies of the central government and the province. At the same time, municipalities have considerable freedom in implementing policies, which is an important source of variation between municipalities (Ter Bogt, 2008).

Municipalities obtain their funds partially from central government. Some of these funds are earmarked for specific goals but most of the funds are generic. Democratically elected councils

decide upon the allocation of the generic funds. Central funding is mostly based on structural characteristics (number of inhabitants, kilometres of road length, etc.) and is independent of performance. There is no link between efficiency and central funding, and less efficient municipalities need to find funding for their deficits within their own budget, or need to cut back on service levels. In addition to the central funds, municipalities obtain their resources from local taxes and revenues from the provision of goods and services (passports, building permits, etc.).

Municipalities in the Netherlands have considerable freedom in the design and use of their control system, and they have been experimenting with NPM reforms since the 1980s. Since 1985, municipalities have been required to apply accrual accounting rules. In the 1990s, many municipalities introduced decentralized organizational structures as well as output-oriented budgeting procedures, including the use of output budgets, interim reports and (non-)financial performance indicators. Since 2004, the new Governments Accounts Act requires Dutch municipalities to distinguish between policy-based outcome budgets (general programme budgets for the municipal council) and output or product budgets (more detailed budgets for internal use).

Many reform initiatives have been voluntary and grew out of organizational or political agendas at the individual municipalities. Even in the case of mandatory adoption, Dutch municipalities have always enjoyed considerable autonomy. They have not been 'passive recipients of the discourses of change' (Thomas & Davies, 2005, p. 683), but have been allowed to mould the reform to suit their specific preferences – or even to resist it and to resort to purely ceremonial adoption. Consequently, our empirical setting ensures significant variety among organizations, spanning the continuum from almost prototypical NPM adopters at one end to organizations that have largely preserved the traditional modes of internal governance at the other.

Survey design and sample selection

In building the questionnaire, we relied as much as possible on instruments validated in previous studies, sometimes with slight modification to suit our research context. The survey was pre-tested by three experts: two (previous) managers of nonprofit organizations and one survey expert. This led to some minor adjustments in the wording of the questions. We approached financial managers because their functional position and seniority ensures a good overview of the functioning of (the controls of) the organization. A total of 96 useable surveys were returned, which is 21.7% of the total population (all 443 municipalities in the Netherlands). The distribution of the number of citizens in our sample is not significantly different from the population. Respondents included the highest ranked general managers (12%), heads of the financial department (42%), controllers (28%) and other respondents (18%, including finance employees). On average, respondents had been working for 12 years in their organization (median 7 years), and have been employed in their current function for about 6 years (median 5 years). This suggests that respondents are well informed about control practices in their organization.

Measurement of variables

Each variable has been measured using multiple items. Most items are measured on a fully anchored 5-point semantic scale. The Appendix reproduces the relevant parts of the survey and reports factor analysis results and item-level descriptive statistics. All variables for each measurement instrument load on a single factor and Cronbach's alpha for each measure is above the threshold of 0.7 (Hair, Anderson, Tatham & Black, 1998), with only one exception.[3]

Because we rely on data from self-report questionnaires to measure both the dependent and the independent variables, common method bias (CMB) is a concern (Podsakoff & Organ, 1986).

However, we are able to validate the dependent variable with archival data (see our discussion below). Furthermore, the specification of the model alleviates the CMB problem. Because our model includes mediator variables and a rather broad set of predictors, it is unlikely that the relationships in our model are part of respondents' cognitive maps (Chang, Van Witteloostuijn & Eden, 2010). This mitigates some of the most powerful causes of CMB, i.e. implicit theories, consistency motif and social desirability (Podsakoff, MacKenzie, Lee & Podsakoff, 2003). To assess the extent of remaining CMB, we run a Harman's one-factor test on the survey items that we used to form the constructs in our model. The factor solution (details not reported) yields nine factors with eigenvalues larger than one, equalling the number of variables in our model. The first factor explains 32% of total variance. In addition to this exploratory factor analysis, we perform Harman's test using confirmatory factor analysis. Craighead, Ketchen, Dunn and Hult (2011) argue that this approach is more robust because differences between a one-factor model and the multifactor model can be tested by comparing the fit statistics. The fit statistics are considerably better for the multifactor model relative to the one-factor model, including χ^2 (2298.6, d.f. 819 for the one-factor model versus 1114.4, d.f. 781 for the multifactor model); CFI (0.482 versus 0.880); and RMSEA (0.139 versus 0.067). Overall, these results support the absence of significant single-source bias.

Organizational performance. Organizational performance is measured with a well-established instrument developed by Van de Ven and Ferry (1980) to assess performance in public sector organizations. This instrument has also been used by, for instance, Dunk and Lysons (1997) and Verbeeten (2008). To validate this measure, we gather additional information on the financial performance of municipalities. The Association of Dutch Municipalities (VNG) proposes three key metrics to evaluate long-term financial performance of municipalities (Van der Lei, 2013): net debt relative to the total budget, net debt per inhabitant, and the change in net debt per inhabitant. An increase in each of these items signals that the municipality is likely to run into financial problems in the long run (Van der Lei, 2013). A factor analysis indicates that these three items load on a single construct (Cronbach's alpha 0.74). We expect that this financial performance measure is negatively associated with our organizational performance construct; even though organizational performance captures multiple dimensions, we assume that our respondents consider long term financial viability when evaluating organizational performance. Consistent with our expectation, we find a negative and significant correlation ($\rho = -1.15$, $p < 0.10$) between the archival measure and our organizational performance construct, adding to the credibility of the survey-based organizational performance measure.

Results-oriented culture. The instrument for results-oriented culture was purposely developed for this project and captures the extent to which respondents agree on a series of four statements related to result accountability of middle managers, acknowledgement of individuals' contributions to performance (both positive and negative), and top management's commitment to goal achievement – all of which are key to result-oriented organizations as described by the OECD (2002, 2005).

Performance measurement systems. To capture the way in which performance measures are used, we adapt an instrument from Henri (2006). Building on various existing classifications, Henri (2006) developed a taxonomy of performance measurement system use that includes four different roles: monitoring (to track progress towards goals), attention-focusing (to provide direction to the organization), strategic decision-making (to uncover cause-and-effect relationships) and legitimization (to validate current and future actions; see Henri, 2006).[4] The attention-focusing, strategic decision-making and legitimization roles are different dimensions of the decision-facilitating use of control systems (Demski & Feltham, 1976; Speklé & Verbeeten, 2014).

Because we expect the different uses to correlate, we use factor analysis with oblique rotation to examine the underlying structure. This analysis reveals three components that (after removing two items with high cross-loadings) reflect the monitoring and attention-focusing roles as per Henri (2006), and a combined strategic decision-making/legitimization use.[5] We retain these three factors. Monitoring covers part of the use for target-setting, accountability and incentives as it features in NPM-thinking (proposition P2). Strategic decision-making captures the use of performance measurement systems to gain a deeper understanding of cause-and-effect relationships, to verify and adjust assumptions, and to question currently held beliefs; we use it as proxy for the learning-support use of performance information as identified by OECD-NPM (proposition P4). The attention-focusing use is not specifically referred to in NPM, but is included in the analysis as a control variable, being an essential element in the more exploratory, decision-facilitating use of the performance measurement system (Frey et al., 2013; Speklé & Verbeeten, 2014).

Henri's instrument does not explicitly pick up on the use of performance measures for incentive purposes (proposition P3). For this reason, we included four additional questions in the survey to capture this type of use. These questions address the role of information on outputs and outcomes in compensation and/or career decisions.

Internal deregulation. Based on Gordon and Narayanan (1984), we asked respondents to indicate the degree to which the organization relies on a manual that describes the administrative processes in the organization, and the degree to which employees of the organization rely on guidelines and procedures. Both items load on a single factor, which we use to measure the reliance on rules and procedures.

We measure decentralization of decision rights with an instrument based on Gordon and Narayanan (1984) and used, among others, by Abernethy et al. (2004) and Verbeeten (2008). Factor analysis returns two constructs that mirror the conceptual distinction made by OECD-NPM between strategic and operational decentralization. The first component captures strategic decentralization, including policy decision rights and investment decisions. The second component refers to operational decentralization and includes internal process design, personnel management and outsourcing decisions.

Results

We test the model using partial least squares analysis (PLS).[6] PLS is a non-parametric component-based structural equation modelling (SEM) technique. As opposed to covariance-based SEM methods (e.g. AMOS and LISREL), PLS puts low demands on sample size. This characteristic makes PLS quite suitable for our project, which combines a rather complex model with a relatively small sample size. The analysis proceeds in two stages. First, we examine the reliability and validity of the measurement model. Second, we evaluate the structural model, examining the relationships between the constructs. OECD-NPM argues that the effects of control structure choices on performance are mediated by the orientation on results, but we include possible direct effects on performance as well to examine whether the suggested mediator effect does in fact exists, and whether it adds explanatory power over and beyond the main effects.

Measurement model, descriptive statistics and correlations

All composite variables are modeled as reflective constructs (Jarvis, MacKenzie & Podsakoff, 2003).[7] Indicator loadings are reported in the Appendix. Tables 1 and 2 provide descriptive

Table 1. Descriptive statistics.

	Theoretical range	Actual range	Mean	Standard deviation	Composite reliability	AVE
Performance	1–5	1.43–4.57	3.39	0.54	0.88	0.51
Results-oriented culture	1–5	1.25–5.00	3.42	0.83	0.92	0.74
PMS use						
• Monitoring	1–5	1.00–4.50	2.71	0.83	0.95	0.84
• Incentives	1–5	1.00–5.00	1.98	0.78	0.94	0.79
• Strategic decision-making	1–5	1.00–4.44	2.71	0.69	0.95	0.67
• Attention-focusing	1–5	1.00–4.20	2.59	0.80	0.94	0.74
Internal deregulation						
• Operational decentralization	1–5	1.00–5.00	3.03	0.91	0.84	0.52
• Strategic decentralization	1–5	1.00–5.00	1.91	0.83	0.85	0.74
• Rules and procedures	1–5	1.00–4.50	2.76	0.76	0.87	0.77

N = 96. Range, mean and standard deviation statistics are based on construct scores calculated as the unweighted average of underlying item scores. In the subsequent PLS analysis, the weights of the indicators are estimated in the measurement model.

statistics of the constructs and bivariate correlations. To assess the reliability of the measurement model, we evaluate the composite reliability (CR) and the average variance extracted (AVE) of the constructs (Werts, Linn & Jöreskog, 1974; see Table 1). The interpretation of CR is similar to that of Cronbach's alpha, but CR does not assume that indicators receive equal weight in the calculation of the construct (Henseler, Ringle & Sinkovics, 2009). The data in Table 1 indicate that composite reliability is quite good, with all scores above 0.80 (Henseler et al. 2009; Hulland, 1999). Convergent validity is also adequate, with all AVE scores exceeding 0.50 (Henseler et al., 2009; Hulland, 1999). In addition, the square root of the AVE scores of each variable is higher than the bivariate correlations between that variable and the other exogenous constructs in the model (see Table 2), demonstrating discriminant validity (Fornell & Larcker, 1981).

The correlations reported in Table 2 between the independent variables are generally low enough not to prompt serious multicollinearity concerns.[8] It should, however, be noted that the correlations between the different uses of performance information are quite high, indicating that they overlap. This finding is not surprising: the roles of the performance measurement system are not mutually exclusive, and the bivariate correlations suggest that if an organization intensifies its use of the performance measurement system for a specific purpose, it is likely to intensify its use for the other purposes as well. This observation might reflect differences in management styles between municipalities, e.g. a 'numbers-based' versus an 'intuitive' style. It is also possible that these correlations indicate the existence of 'natural' spill-over effects. For instance, if performance information becomes more salient because senior management increases its reliance on this information to focus employee attention, this may spur a more intense use of the information in decision-making or monitoring as well.

The bivariate analysis offers some support for the OECD-NPM view of effective management control. The various uses of the performance measurement system are positively and significantly associated with both a results-oriented culture and performance. Also, there appears to be a positive relation between the degree of results-orientation and performance. However, contrary to OECD-NPM, a reliance on rules and procedures contributes positively to the organization's results orientation and to performance, while decentralization is not correlated with either of these outcome variables.

Table 2. Correlations.

	1	2	3	4	5	6	7	8	9
1: Performance	0.71								
2: Results-oriented culture	0.35	0.86							
3: Monitoring	0.37	0.38	0.92						
4: Incentives	0.23	0.25	0.41	0.89					
5: Strategic decision-making	0.31	0.28	0.69	0.49	0.82				
6: Attention-focusing	0.40	0.32	0.72	0.42	0.77	0.86			
7: Operational decentralization	–0.10	0.05	0.03	–0.06	–0.12	–0.05	0.72		
8: Strategic decentralization	0.08	–0.14	–0.11	–0.05	–0.06	0.04	0.36	0.86	
9: Rules and procedures	0.33	0.43	0.37	0.35	0.29	0.26	0.03	0.01	0.88

N = 96. Off-diagonal elements are correlations, diagonal numbers (underscored) report the square root of AVE. Correlations with an absolute value greater than 0.20 and 0.26 are significant at the 0.05 and 0.01 level, respectively (two-tailed).

Table 3. PLS path coefficients.

Proposition	Path from:	Path to:	
		Results-oriented culture	Performance
P1: +	Results-oriented culture	–	0.21 (1.56/0.06)*
P2: +	Monitoring	0.14 (1.16/0.12)	0.14 (1.29/0.10)*
P3: +	Incentives	0.04 (0.59/0.28)	0.01 (0.11/0.46)
P4: +	Strategic decision-making	–0.07 (0.60/0.27)	–0.10 (0.83/0.20)
Not predicted	Attention-focusing	0.18 (1.35/0.09)*	0.25 (1.64/0.05)**
P5: -	Rules and procedures	0.33 (3.10/0.00)***	0.15 (1.42/0.08)*
P6: +	Operational decentralization	0.11 (1.12/0.13)	–0.17 (1.81/0.04)**
P7: -	Strategic decentralization	–0.18 (1.84/0.03)**	0.17 (1.59/0.06)*
		R^2 = 0.27	R^2 = 0.27

N = 96. Absolute t-values/*p*-values between parentheses.
Reported *p*-values are one-tailed: *$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

Structural model results

The PLS estimates of the structural model are reported in Table 3 and illustrated in Figure 2. To evaluate the significance of the path coefficients, we run a bootstrapping procedure (5,000 subsamples with replacement[9]). The outcomes of this procedure are also reported in Table 3.

Consistent with OECD-NPM proposition P1, the results indicate that the degree of results orientation is positively associated with organizational performance. However, we find no support for the key OECD-NPM claim that the use of performance measurement systems for accountability and incentive purposes (propositions P2 and P3) contributes to a results-oriented culture. We do, however, find a small and marginally significant positive direct effect of monitoring on organizational performance. There is no evidence that the use of the performance measurement system for learning (strategic decision-making) has an effect (P4), but we do observe a positive association with both results orientation and performance for an attention-focusing role of the system. This role, however, is not part of the OECD-NPM repertoire. Moreover, we find that the reliance on rules contributes positively to a results-oriented culture, contradicting proposition P5. This finding

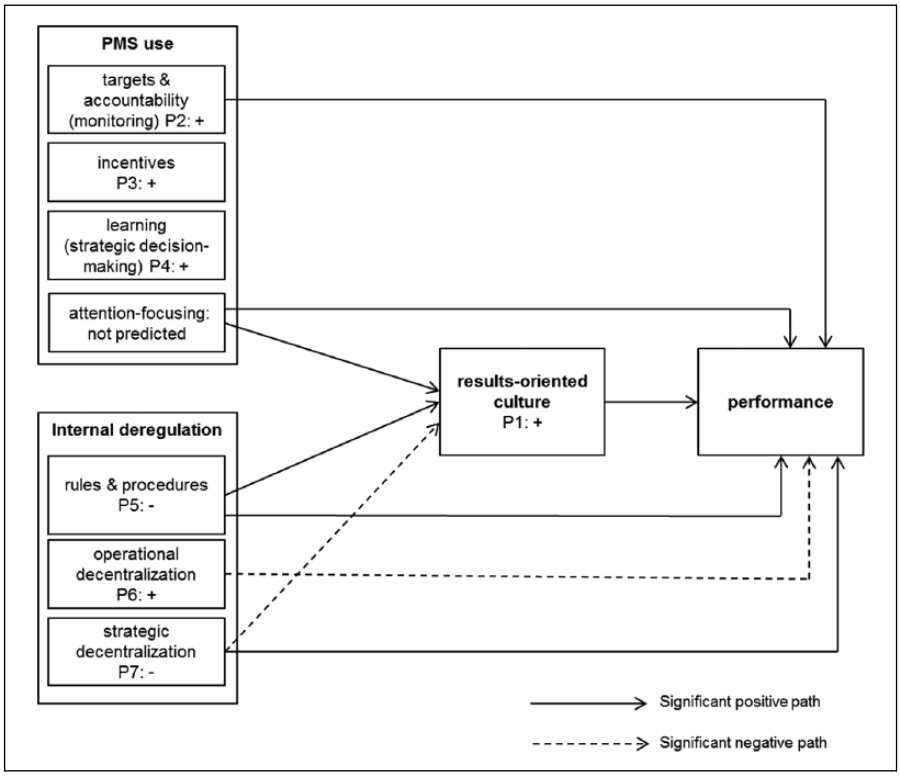

Figure 2. Significant paths.

suggests that OECD-NPM's position on rules and procedures is too negative and that these controls may actually facilitate task performance.

At first glance, the results seem partly to support OECD-NPM's position on decentralization. That is, we do not find the significant effect of operational decentralization on the degree of results orientation as predicted in proposition P6, but find a negative effect of strategic decentralization. This latter result is consistent with NPM's concern that decentralization of strategic decision rights may introduce a political dimension in operational affairs, drawing the organization away from its operational tasks (proposition P7). However, operational and strategic decentralization also affect performance directly. The direct effect of operational decentralization on performance is negative and significant, consistent with the suggestion in the literature that decentralization may be associated with additional coordination costs or private goal pursuit at the expense of organizational performance. Because this negative effect is not compensated for by the indirect effect through results orientation (which is insignificant), we conclude that OECD-NPM's suggestion to increase operational decentralization is potentially dysfunctional.

The effects of strategic decentralization are opposite of those observed for operational decentralization. Decentralization of strategic decision rights negatively affects the degree of results orientation, but it has a positive direct effect on performance. This positive direct effect may indicate that decentralized policy-making leads to more effective government programmes that better take into account local circumstances (OECD, 1994). The results suggest that the positive effect of strategic decentralization outweighs the cost of 'politicization' of performance measurement,

Table 4. PLS path coefficients trimmed model.

Path from:	Path to	
	Results-oriented culture	Performance
Results-oriented culture		0.21 (1.64/0.05)*
Attention-focusing	0.23 (2.32/0.01)**	0.28 (2.58/0.00)***
Rules and procedures	0.37 (4.01/0.00)***	0.16 (1.66/0.05)**
Operational decentralization		–0.16 (1.78/0.04)**
Strategic decentralization	–0.16 (1.82/0.03)**	0.13 (1.30/0.10)*
	R^2 = 0.25	R^2 = 0.27

N = 96. Absolute t-values/*p*-values between parentheses.
Reported *p*-values are one-tailed: *$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

because the direct effect on performance is much larger than the (negative) indirect effect through results orientation. In fact, a further analysis[10] reveals that the indirect performance effect is not significant ($Z = -1.10$, $p = 0.14$, one-tailed), and we conclude that the advice of OECD-NPM on the allocation of strategic policy decision rights appears ill-founded.

Additional analysis

Our sample size is equal to 12 times the largest number of paths to the individual constructs in the structural model, which exceeds the threshold for robust PLS path modelling (Henseler et al., 2009; Ringle, Sarstedt & Straub, 2012). However, the analysis may lack statistical power because of the limited number of observations and could fail to detect significant effects of smaller size (Henseler et al., 2009). Therefore, we also analyse a trimmed model. We arrive at this model in a step-wise procedure in which we eliminate the insignificant paths one at a time, starting with the path with the lowest t-value. Between two steps, we re-estimate the model to identify the next path to eliminate. Ultimately, we arrive at a parsimonious model containing only significant paths. The trimmed model contains only 8 structural paths (as opposed to 15 paths in the original model), and the ratio of observations to the largest number of paths to an endogenous construct improves to 19. The results of the trimmed model (presented in Table 4) are substantively similar to the results of the original analysis[11] and reinforce our earlier inferences.

As a further analysis, we also ran the original model controlling for three additional variables that – according to academic literature – may affect the relations in the model, i.e. the quality of the performance measurement system, size, and clarity of goals (see Cavaluzzo & Ittner, 2004; Chun & Rainey, 2005; Moynihan & Pandey, 2005). The (untabulated) results of this more elaborate model confirm our earlier findings and lead to qualitatively similar conclusions.

Conclusions and Discussion

This paper is one of the first empirical studies to evaluate the full package of NPM thinking on effective management control. Our study addresses both performance measurement systems and internal deregulation as key components of the OECD-NPM-approach. We also take into account the contention that the effects of management control choices on performance are indirect, operating through their impact on result-oriented culture rather than affecting performance directly (Hood, 1991). We test these OECD-NPM ideas in a sample of 96 Dutch municipalities, an environment characterized by significant variation in the adoption of NPM-based initiatives.

Our findings provide little support for the OECD-NPM working theory of management control, and even suggest that NPM-consistent reforms may have had a negative impact on public sector performance. First, we find that the assumed benefits of internal deregulation ('letting the managers manage') are illusory in our sample; instead, we observe a positive effect of rules and procedures on both a results-oriented culture and on performance, and a negative effect of operational decentralization on performance. In addition, our data indicate that the net performance effect of strategic decentralization is positive, casting doubt on the benefits OECD-NPM ascribes to the separation of policy development and policy implementation. Our findings, however, are well in line with theoretical insights from the literature, i.e. that internal deregulation leads to higher coordination costs and interdependencies between organizational parts (Mascarenhas, 1993), excessive opportunity-seeking behaviour (Jensen & Meckling, 1992; Simons, 1995) and a failure to codify 'best practices' (Adler & Borys, 1996). With its exclusive focus on the downside of bureaucracy, OECD-NPM appears to ignore that the real-world complexities of the public sector may actually necessitate a reliance on internal regulation. Subsequent studies could further explore the trade-off between the downside of internal regulation and its potential benefits, and may investigate under what specific (organizational) circumstances different aspects of internal deregulation help to improve the achievement of a results-oriented culture or organizational performance.

Second, we find no empirical support for NPM's other key tenet either, i.e. that the use of performance information for purposes of accountability and incentives ('making the managers manage') improves performance. We find insignificant results for these variables in our model, suggesting that results orientation and performance are not associated with these uses of performance information. However, the reliance on incentives is quite low in our sample (see Table 1),[12] and it may be the case that the assumed benefits of incentives only materialize above a certain threshold. In addition, we note that our data on the use of performance information relate to civil service managers, not to political decision-makers. Ter Bogt (2004) observes that these politicians are not heavy users of formal performance measurement systems, but rely on informal, verbal communication as their preferred source of information. This apparent disconnect between the formal performance measurement system on which public sector managers are supposed to rely and the informational preferences of their political bosses is likely to affect the organizational status and impact of formal control systems, potentially undermining their effectiveness. For these reasons, it may be that our study underestimates the true potential of accountability and incentive-oriented uses of performance measurement systems. But this is speculative and we do not have the data to support these contentions. We do, however, find evidence to suggest that performance measurement matters if it is being used for attention-focusing, i.e. to provide guidance to the organization as to top management's view of the organization, key success factors and critical uncertainties (see Frey et al., 2013; Speklé & Verbeeten, 2014). This way of using the system, however, is not part of the OECD-NPM approach to public sector management control effectiveness.

Taken together, the null finding for incentive-oriented performance measurement system use and the unanticipated positive effect of an attention-focusing role indicate the need to take a broad perspective on the potential roles of the performance measurement system (Frey et al., 2013; Speklé & Verbeeten, 2014). Because the various roles of the system are associated (see Table 2), ignoring individual roles may result in correlated omitted variable problems and, hence, in misleading conclusions. In fact, it is quite possible that previously reported reform benefits (OECD, 1994, 1997) do not come from target-based incentives, but originate from an exploratory, decision-facilitating use of the performance measurement system. Future studies should consider these multiple roles, and may investigate how specific public sector organizations succeed (or fail) in using the performance measurement system effectively (e.g. Townley et al., 2003). Such studies could

also seek to investigate the relationship between control and organizational culture in more detail. In our empirical analysis we follow OECD-NPM and model culture as an (intermediate) outcome of control system features. However, previous literature (e.g. Henri, 2006) suggests that the relationship is the other way around, positioning culture as a determinant rather than a consequence of control practices. A better understanding of the exact nature of this relationship is clearly needed.

Finally, our empirical findings indicate the need to explore the relationship between the components of the control system. OECD-NPM implies that target-based performance management and internal deregulation are complementary, where the benefits of performance contracting require operational decentralization and loosening of rules. Because of sample size limitations, we could only include the main effects of these control components in our analysis and had to ignore potential interaction effects. However, in the bivariate analysis (see Table 2), we find a positive association between the reliance on rules and procedures and all four uses of the performance measurement system – including incentive-oriented use. This finding is consistent with a very different logic than the one advanced in OECD-NPM, and suggests that strict rules and regulations are necessary to control excessive opportunity-seeking behaviour that may result from a strong reliance on performance measurement. A further examination of such interdependencies between elements of the control structure is an important avenue for future research.

The results of this study should be interpreted with care. First, our data come exclusively from Dutch municipalities. Such a single-sector single-country sample provides indirect control over potentially confounding effects of sector and institutional factors, but limits the generalizability of the findings. Because Dutch municipalities differ substantially in the adoption and implementation of NPM ideas, our findings may be relevant for the study of performance effects of control systems in the public sector more broadly. However, the response to control practices may differ depending on institutional setting and cultural tradition, and some of our results may be affected by the context in which they were obtained. A factor to consider here is the fit between the control structure and the broader administrative culture. Administrative cultures differ significantly across the globe (Van de Walle, Sterck, Van Dooren, Bouckaert & Pommer, 2004) and may affect the behavioural impact of internal deregulation or the response to result-oriented incentives. Moreover, the administrative context needs to be factored in to the analysis to acquire an adequate understanding of how control practices that appear very similar from the outside may still be quite heterogeneous in reality. For example, the low reliance on incentive contracting we observe in our data is characteristic of the public sector in general. In OECD member states, the variable performance-dependent component of public sector compensation is typically less than 5%; the Dutch figure is about 3% (OECD, 2007). These figures, however, exclude in-range salary increases, which may be linked to performance. This is in fact the case in the Netherlands, but only formally. In actual practice, such increases are seen as a right that can only be denied if performance is exceptionally bad, thus further weakening the link between pay and performance. In other countries, this may well be different. Even though formal evidence is hard to find (Pollitt, 2002), many would agree that the Dutch setting would be comparable with the Nordic nations, and that this cluster of northwestern countries holds a sort of midway position between the Anglo-Saxon public interest model on the one hand and the more legalistic and bureaucratic models of France and the Mediterranean states on the other (Pollitt, 2002; Van de Walle et al., 2004). Further research is required to assess whether the results can be replicated in a broader public sector setting or in other institutional settings.

Another caveat is that our study relies on cross-sectional data; the evidence pertains to associations rather than causal relationships. In addition, the results presented here are based on perceptions rather than 'hard' measures, and our proxies may misrepresent some of the key concepts in the analyses because of inappropriate measures or inadequate interpretation of the survey. Although our reliance on previously validated instruments, the pre-test of the questionnaire and the

correlation of our organizational performance measure with objective financial performance data should alleviate these concerns, some apprehension may remain.

Despite these limitations, our study has important implications for public policy. In a moment of introspection, participants in one of the OECD expert meetings commented on the virtual absence of sound evaluations on the effects of reform initiatives. They challenged advocates of NPM-type reforms to 'take their own medicine' and to evaluate the impact of their results-based reform initiatives, just as they are urging government agencies to evaluate their own policies and programmes (OECD, 2002, p. 27). Our study underscores the need for such evaluations and suggests that the validity of the NPM approach to effective management control may have been taken for granted too long by those involved in public sector reforms (Bejerot & Hasselbladh, 2013).

It is beyond the scope of this paper to make firm suggestions as to the required changes in the OECD reform agenda. A few tentative remarks based on or triggered by our findings may nevertheless be appropriate. First, the unconditional belief of OECD-NPM in the beneficial effects of incentives in public sector organizations is unwarranted. This belief finds no basis in theory – rather on the contrary – and cannot be buttressed empirically either. In our study, we found no effect of incentives. Our study does, however, suggest that the future reform agenda should consider a broader set of uses of performance measurement systems, particularly the decision-facilitating, attention-focusing role.

Second, future reform initiatives require a more nuanced view of the role and contribution of rules and procedures. Our study shows that a reliance on such controls is actually positively associated with organizational performance, and indeed, a well-functioning public sector may require extensive regulation for purposes of coordination, codification of experience, or to counteract all too strong or all too selective incentives associated with the use of target-based performance contracting. Regulation plays an essential role in ensuring a fair and equitable treatment of citizens; this important role seems to have dropped from NPM's efficiency-dominated reform agenda entirely. Internal governance, however, is not just about economic goals, but can address the full set of organizational objectives, including political and ideological ones. Many would agree that the present emphasis on efficiency does not adequately capture the essential purposes of the public sector. Finding a satisfactory balance between managerial notions of results on the one hand and democratic values that require due process on the other is one of the main challenges facing public administration (Arellano-Gault et al., 2013).

Finally, in our examination of the OECD reports, we found no substantial reference to the political context in which public sector organizations operate. This context, however, has a considerable impact on the internal functioning of public sector organizations (Arellano-Gault et al., 2013), posing unique challenges to leadership, for example. An implicit assumption in NPM is that public sector organizations enjoy significant sovereignty in matters of policy formation. The reality, however, is that policy-making requires commitment from competing stakeholder groups and broad public engagement. Managing the associated tensions requires a type of leadership that is collaborative and inclusive (Oborn, Barrett & Dawson, 2013), which may be difficult to reconcile with the managerialism that dominates the NPM discourse (O'Reilly & Reed, 2011). Nevertheless, such matters should be incorporated in any aspiring theory of public sector management. They should also be part of the normative agenda for reform to ensure a sufficient connection to the complexities of the world it seeks to change.

Acknowledgements

We thank David Arellano-Gault (the senior editor) and three anonymous reviewers for their constructive comments. A previous version of this paper has benefited from comments and suggestions by Anne-Marie Kruis, Brendan O'Dwyer, Henk ter Bogt, Jan van Helden, Paula van Veen-Dirks, Sally Widener and Raymond

Zaal. We also acknowledge valuable feedback from workshop participants at the EIASM conference on Accounting, Auditing and Management in Public Sector Reforms, the EIASM Conference on Performance Measurement and Management Control, and the AAA Management Accounting Section (especially Anne Lillis). Special thanks to Ramon Reijpert for research assistance.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Notes

1. The OECD uses the term 'autonomy' as a synonym for decentralization. We follow the OECD terms when discussing OECD documents.
2. This section is based upon Pollitt and Bouckaert, 2004; Speklé and Verbeeten, 2014; Ter Bogt, 2008; and Van Helden, 1998.
3. Cronbach's alpha for strategic decentralization is 0.66. Its composite reliability, however, is quite good, with a score of 0.85 (see Table 1).
4. Henri's use of the term 'legitimization' differs from the concept of legitimization in institutional theory, where it is associated with impression management and (ceremonial) submission to outside pressure. Henri (2006) also implies such notions in his definition of the term, but his subsequent operationalization positions the construct exclusively in the context of economically rational decision-making.
5. The observation that legitimization and strategic decision-making cannot be separated empirically is not surprising. Legitimization as construed by Henri (2006) includes the validation of current and future actions, and most of the relevant questionnaire items are in fact associated with justification of current decisions, verification of assumptions and reinforcement of beliefs (see the Appendix for the wording of the items). But all this is also closely related to the theoretical notion of strategic decision-making; apparently, our respondents appreciate them as elements of this latter notion.
6. We use SmartPLS 2.0 (Ringle, Wende & Will, 2005).
7. Factor analysis confirmed unidimensionality of the constructs.
8. This is also confirmed in a series of two (untabulated) OLS regressions with performance and results-oriented culture as the dependent variables. In these regressions, the highest variance inflation factor is 3.2, well below the threshold values of 6 or 10 recommended by Cohen, Cohen, West and Aiken (2003) and Hair et al. (1998), respectively.
9. We correct for arbitrary sign chances by reversing the sign of the paths in bootstrapping subsamples if these signs differ from the original path estimates (Henseler et al., 2009).
10. In this analysis, we use the 'product of coefficient strategy' described by Preacher and Hayes (2008).
11. This also holds for the measurement model.
12. Only twelve respondents (12.5% of the sample) report an incentive-oriented use of the performance measurement system at or above the midpoint of the measurement scale.

References

Abernethy, M. A., Bouwens, J., & Van Lent, L. (2004). Determinants of control system design in divisionalized firms. *Accounting Review*, *7*, 545–570.

Abernethy, M. A., & Vagnoni, E. (2004). Power, organization design and managerial behaviour. *Accounting, Organizations and Society*, *29*, 207–225.

Adler, P. S., & Borys, B. (1996). Two types of bureaucracy: Enabling and coercive. *Administrative Science Quarterly*, *41*, 61–89.

Arellano-Gault, D., Demourtain, D., Rouillard, C., & Thoening, J.C. (2013). Bringing public organization and organizing back in. *Organization Studies*, *34*, 145–167.

Arellano-Gault, D., & Lepore, W. (2011). Transparency reforms in the public sector: Beyond the new economics of organization. *Organization Studies*, *32*, 1029–1050.

Bejerot, E., & Hasselbladh, H. (2013). Forms of intervention in public sector organizations: Generic traits in public sector reforms. *Organization Studies*, *34*, 1357–1380.

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

Brunsson, N., & Sahlin-Andersson, K. (2000). Constructing organizations: The example of public sector reform. *Organization Studies*, *21*, 721–746.

Burgess, S., & Ratto, M. (2003). The role of incentives in the public sector: Issues and evidence. *Oxford Review of Economic Policy*, *19*, 285–299.

Cameron, K. S., & Freeman, S. J. (1985). *Cultural congruence, strength, and type: relationships to effectiveness*. School of Business Administration, University of Michigan.

Cavalluzzo, K. S., & Ittner, C. D. (2004). Implementing performance measurement innovations: Evidence from government. *Accounting, Organizations and Society*, *29*, 243–267.

Chang, S., Van Witteloostuijn, A., & Eden, L. (2010). Common method variance in international business research. *Journal of International Business Studies*, *41*, 178–184.

Chun, Y. H., & Rainey, G. H. (2005). Goal ambiguity and organizational performance in US federal agencies. *Journal of Public Administration Research and Theory*, *15*, 529–557.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*, 3rd edition. Mahwah, NJ: Lawrence Erlbaum Associates.

Craighead, C. W., Ketchen, D. J., Dunn, K. S., & Hult, G. T. M. (2011). Addressing common method variance: Guidelines for survey research on information technology, operations, and supply chain management. *IEEE Transactions on Engineering Management*, *58*, 578–588.

Demski, J. S., & Feltham, G. A. (1976). *Cost determination: a conceptual approach*. Ames: Iowa State University Press.

Dewatripont, M., Jewitt, I., & Tirole, J. (1999). The economics of career concerns, part II: Application to missions and accountability of government agencies. *Review of Economic Studies*, *66*, 199–217.

Diefenbach, T. (2009). New Public Management in public sector organizations: The dark sides of managerialistic 'enlightenment'. *Public Administration*, *87*, 892–909.

Dixit, A. (2002). Incentives and organizations in the public sector: An interpretive review. *Journal of Human Resources*, *37*, 696–718.

Dunk, A. S., & Lysons, A. F. (1997). An analysis of departmental effectiveness, participative budgetary control processes and environmental dimensionality within the competing values framework: A public sector study. *Financial Accountability & Management*, *13*, 1–15.

Feltham, G. A., & Xie, J. (1994). Performance measure congruity and diversity in multi-task principal/agent relations. *Accounting Review*, *69*, 429–453.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(February), 39–50.

Franco-Santos, M., Lucianetti, L., & Bourne, M. (2012). Contemporary performance measurement systems: A review of their consequences and a framework for research. *Management Accounting Research*, *23*, 79–119.

Frey, B. S., Homberg, F., & Osterloh, M. (2013). Organizational control systems and pay-for-performance in the public service. *Organization Studies*, *34*, 949–972.

Garnett, J. L., Marlowe, J., & Pandey, S. K. (2008). Penetrating the performance predicament: Communication as a mediator or moderator of organizational culture's impact on public organizational performance. *Public Administration Review*, *68*, 266–281.

Gibbons, R. (1998). Incentives in organizations. *Journal of Economic Perspectives*, *12*(4), 115–132.

Gordon, L. A., & Narayanan, V. K. (1984). Management accounting systems, perceived environmental uncertainty and organization structure: An empirical investigation. *Accounting, Organizations and Society*, *9*, 33–47.

Grafton, J., Lillis, A. M., & Widener, S. K. (2010). The role of performance measurement and evaluation in building organizational capabilities and performance. *Accounting, Organizations and Society*, *35*, 689–706.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis*, 5th edition. Upper Saddle River, NJ: Prentice Hall.

Hansen, S. C., & Van der Stede, W. A. (2004). Multiple facets of budgeting: An exploratory analysis. *Management Accounting Research*, *15*, 415–439.

Henri, J. F. (2006). Organizational culture and performance measurement systems. *Accounting, Organizations and Society*, *31*, 77–103.

Henseler, J., Ringle, C. M., & Sinkovics, R. R. (2009). The use of partial least squares path modeling in international marketing. *Advances in International Marketing*, *20*, 277–319.

Hofstede, G. (1981). Management control of public and not-for-profit activities. *Accounting, Organizations and Society*, *6*, 193–211.

Hood, C. (1991). A public management for all seasons? *Public Administration*, *69*, 3–19.

Hood, C. (1995). The 'new public management' in the 1980s: Variations on a theme. *Accounting, Organizations and Society*, *20*, 93–109.

Hulland, J. (1999). Use of partial least squares (PLS) in strategic management research: A review of four recent studies. *Strategic Management Journal*, 20, 195–204.

Jarvis, C. B., MacKenzie, S. B., & Podsakoff, P. M. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research*, *30*, 199–218.

Jensen, M. C., & Meckling, W. H. (1992). Specific and general knowledge, and organization structure. In L. Werin & H. Wijkander (Eds.), *Contract economics* (pp. 251–274). Oxford: Basil Blackwell.

Laegreid, P., Roness, P. G., & Verhoest, K. (2011). Explaining the innovative culture and activities of state agencies. *Organization Studies*, *32*, 1321–1347.

Mascarenhas, R .C. (1993). Building an enterprise culture in the public sector: Reform of the public sector in Australia, Britain, and New Zealand. *Public Administration Review*, *53*, 319–328.

Moynihan, D. P., & Pandey, S. J. (2005). Testing how management matters in an era of government by performance management. *Journal of Public Administration Research and Theory*, *15*, 421–439.

Mundy, J. (2010). Creating dynamic tensions through a balanced use of management control systems. *Accounting, Organizations and Society*, *35*, 499–523.

Newberry, S., & Pallot, J. (2004). Freedom or coercion? NPM incentives in New Zealand central government departments. *Management Accounting Research*, *15*, 247–266.

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

Oborn, E., Barrett, M., & Dawson, S. (2013). Distributed leadership in policy formulation: A sociomaterial perspective. *Organization Studies*, *34*, 253–276.

OECD (1994). *Performance management in government: performance measurement and results-oriented management*. PUMA occasional papers 3.

OECD (1997). *In search of results: performance management practices*. PUMA.

OECD (2002). (Perrin, B.) *Implementing the vision: addressing challenges to results-focused management and budgeting*.

OECD (2005). (Mayne, J.) *Using performance information for managing and budgeting: challenges, lessons and opportunities*. Working party of senior budget officials.

OECD (2007). (Ketelaar, A., Manning, N., & Turkisch, E.). *Performance-based arrangements for senior civil servants: OECD and other country experiences*. OECD working papers on public governance (2007)/5.

O'Reilly, D., & Reed, M. (2011). The grit in the oyster: professionalism, managerialism and leaderism as discourses of UK public services modernization. *Organization Studies*, *32*, 1079–1101.

Parker, R., & Bradley, L. (2000). Organisational culture in the public sector: Evidence from six organisations. *International Journal of Public Sector Management*, *13*, 125–141

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*, 879–903.

Podsakoff, P. M., & Organ, D. W. (1986). Self-reports in organizational research: Problems and prospects. *Journal of Management*, *12*, 531–544.

Pollitt, C. (2002). Clarifying convergence: Striking similarities and durable differences in public management reform. *Public Management Review*, *4*, 471–492.

Preacher, K. J., & Hayes, A. F. (2008). Contemporary approaches to assessing mediation in communication research. In A. F. Hayes, M. D. Slater, & L. B. Snyder (Eds.), *The Sage sourcebook of advanced data analysis methods for communication research* (pp. 13–54). Thousand Oaks, CA: SAGE Publications.

Propper, C., & Wilson, D. (2003). The use and usefulness of performance measures in the public sector. *Oxford Review of Economic Policy*, *19*, 250–267.
Ringle, C. M., Sarstedt, M., & Straub, D. W. (2012). A critical look at the use of PLS-SEM in MIS Quarterly. *MIS Quarterly*, *36*, iii–xiv.
Ringle, C., Wende, S., & Will, A. (2005). *SmartPLS*. http://www.smartpls.de
Rodrigues, S. B. (2006). The political dynamics of organizational culture in an institutionalized environment. *Organization Studies*, *27*, 537–557.
Schein, E. H. (1984). Coming to a new awareness of organizational culture. *Sloan Management Review*, *25*(2), 3–16.
Simons, R. (1995). *Levers of control: how managers use innovative control systems to drive strategic renewal*. Boston, MA: Harvard Business School Press.
Speklé, R., & Verbeeten, F. H. M. (2014). The use of performance measurement systems in the public sector: Effects on performance. *Management Accounting Research*, *25*, 131–146.
Taylor, J. (2011). Strengthening the link between performance measurement and decision making. *Public Administration*, *89*, 860–878.
Ter Bogt, H. J. (2004). Politicians in search of performance information? Survey research on Dutch aldermen's use of performance information. *Financial Accountability & Management*, *20*, 221–252.
Ter Bogt, H. J. (2008). Management accounting change and New Public Management in local government: A reassessment of ambitions and results – An institutionalist approach to accounting change in the Dutch public sector. *Financial Accountability & Management*, *24*, 209–241.
Thomas, R., & Davies, A. (2005). Theorizing the micro-politics of resistance: New Public Management and managerial identities in the UK public services. *Organization Studies*, *26*, 683–706.
Tirole, J. (1994). The internal organization of government. *Oxford Economic Papers*, *46*, 1–29.
Townley, B., Cooper, D. J., & Oakes, L. (2003). Performance measures and the rationalization of organizations. *Organization Studies*, *24*, 1045–1071.
Van de Ven, A. H., & Ferry, D. L. (1980). *Measuring and assessing organizations*. New York: John Wiley.
Van de Walle, S., Sterck, M., Van Dooren, W., Bouckaert, G., & Pommer, E. (2004). Public Management. In: *Public sector performance: an international comparison of education, health care, law and order and public administration* (pp. 234–270). The Hague: SCP Social and Cultural Planning Office.
Van Helden, G. J. (1998). A review of the policy and management instruments project for municipalities in the Netherlands. *Financial Accountability & Management*, *14*, 85–104.
Van Helden, G. J. (2005). Researching public sector transformation: The role of management accounting. *Financial Accountability & Management*, *21*, 99–133.
Van Veen-Dirks, P. (2010). Different uses of performance measures: The evaluation versus reward of production managers. *Accounting, Organizations and Society*, *35*, 141–164.
Van der Lei, J. (2013). *Houdbare Gemeentefinanciën*. Den Haag: Vereniging Nederlandse Gemeenten (VNG).
Verbeeten, F. H. M. (2008). Performance management practices in public sector organizations: Impact on performance. *Accounting, Auditing and Accountability Journal*, *21*, 427–454.
Werts, C. E., Linn, R. L., & Jöreskog, K. G. (1974). Intraclass reliability estimates: Testing structural assumptions. *Educational and Psychological Measurement*, *34*, 25–33.
Wilderom, C. P. M., Glunk, U., & Maslowski, R. (2000). Organizational culture as a predictor of organizational performance. In N. M. Ashkanasy, C. P. M. Wilderom, & M. Peterson (Eds.), *Handbook of organizational culture and climate* (pp. 193–209). Thousand Oaks, CA: SAGE Publications.

Author biographies

Frank H. M. Verbeeten is Professor of Accounting at Utrecht University School of Economics, and is also associated with the EMFC/ARCA programmes at the VU University Amsterdam. His research focuses on management control practices in organizations, with particular interests in public sector organizations and intangible-intensive firms. His publications have appeared in several accounting and management journals.

Roland Speklé is Professor of Management Accounting & Control at Nyenrode Business University. His research interests are in the field of the economics of organizational control, the design of effective

management control structures, and performance measurement systems. His work has been published in *Accounting, Organizations and Society, Management Accounting Research* and *European Accounting Review*.

Appendix: Survey questions, item-level descriptives, factor analysis results and indicator loadings

Performance measurement system use

Please indicate how often top management uses the performance measurement system for the following purposes (1 = not at all, 5 = always):

	mean	s.d.	Pattern loadings			PLS indicator loadings
			Monitoring	**Attention focusing**	**Strategic decision-making**	
To track progress towards goals	2.56	0.90	0.990	0.017	–0.170	0.882
To review important goals	2.72	0.95	0.908	0.033	0.022	0.947
To monitor results	2.81	0.93	0.873	0.041	0.069	0.948
To compare outcomes to expectations	2.75	0.84	0.676	0.086	0.222	0.887
To tie the organization together (dropped)						
To enable a focus on common issues	2.56	0.93	0.166	0.724	0.058	0.882
To enable a focus on critical success factors	2.58	0.97	0.267	0.720	–0.024	0.895
To emphasize a joint culture within the organization	2.63	1.00	0.006	0.947	–0.068	0.894
To provide a common view of the organization	2.60	0.89	0.019	0.842	0.034	0.879
To enable discussion of policy assumptions and results	2.55	0.86	–0.117	0.659	0.248	0.753
To enable discussion between superiors, subordinates, and peers (dropped)						
To facilitate strategic decision-making	2.73	0.90	0.022	0.293	0.608	0.810
To make decisions when it is difficult to differentiate between alternative solutions	2.53	0.75	0.016	–0.007	0.815	0.815
To reach decisions in unstructured problems	2.43	0.82	–0.090	0.097	0.781	0.798
To facilitate strategic and policy discussions	2.76	0.89	0.118	–0.130	0.875	0.851
To justify decisions	2.92	0.88	0.293	–0.205	0.798	0.834
To verify policy assumptions	2.64	0.78	–0.204	0.238	0.772	0.800
To reinforce a vision	2.72	0.87	–0.144	0.097	0.862	0.831
To (re) confirm existing policies	2.80	0.88	0.109	–0.057	0.774	0.804
To support actions	2.83	0.82	0.064	0.246	0.598	0.832

Incentives

Please indicate how often top management uses specific information (output and outcome measures) from the performance measurement system for the following purposes (1 = not at all, 5 = always):

	mean	s.d.	factor loadings	PLS indicator loadings
Output measures are used to determine employee bonuses	2.02	0.97	0.849	0.851
Output measures are used to make career decisions regarding individual employees	2.03	0.80	0.882	0.892
Outcome measures are used to determine employee bonuses	1.93	0.92	0.909	0.903
Outcome measures are used to make career decisions regarding individual employees	1.95	0.84	0.904	0.899

Results-oriented culture

Please indicate the extent to which you agree with the following statements (1 = strongly disagree, 5 = strongly agree):

	mean	s.d.	factor loadings	PLS indicator loadings
Higher management of my organization is strongly committed to achieving the formulated objectives	3.55	1.01	0.827	0.839
Managers in my organization are being held responsible for the results they achieve	3.40	0.98	0.909	0.915
Managers in my organization are confronted when they do not succeed in realizing their targets	3.36	0.91	0.889	0.894
Employees in my organization receive recognition when they help to achieve the objectives of my organization	3.35	0.96	0.825	0.795

Rules and procedures

Please indicate the extent to which you agree with the following statements (1 = strongly disagree, 5 = strongly agree):

	mean	s.d.	factor loadings	PLS indicator loading
My organization relies heavily on a manual holding specific procedural requirements	2.41	0.84	0.878	0.870
Employees can fall back on written rules and procedures	3.11	0.89	0.878	0.885

Decentralization of decision rights

Please indicate the extent to which decentralized units have the authority to make the following decisions (1 = all decisions are made at the central level, 5 = all decisions are made at the local/decentralized level):

	mean	s.d.	Pattern loading		PLS indicator loading
			Strategic decentralization	**Operational decentralization**	
Policy and strategic decisions	1.74	0.89	0.893	–0.070	0.806
Investment decisions	2.07	1.03	0.836	0.025	0.914
Decisions regarding internal processes	3.27	1.01	0.060	0.773	0.453
Outsourcing decisions	2.57	1.18	0.195	0.731	0.823
Human resource management decisions	2.51	1.07	0.238	0.632	0.611
Determining procedures and work process descriptions	3.47	1.17	–0.087	0.891	0.857
Decisions on reallocation of funds	3.32	1.17	–0.205	0.935	0.794

Performance

How do you compare the performance of your organization to other, comparable organizations on the following dimensions (1=far below average, 5=far above average):

	mean	s.d.	factor loadings	PLS indicator loading
The quantity or amount of work produced	3.63	0.65	0.673	0.638
The quality or accuracy of work produced	3.46	0.81	0.768	0.736
The number of innovations or new ideas by the unit	3.18	0.81	0.640	0.645
Reputation of 'work excellence'	3.13	0.78	0.645	0.677
Attainment of unit production or service goals	3.41	0.68	0.742	0.793
Efficiency of unit operations	3.23	0.78	0.776	0.770
Morale of unit personnel	3.66	0.81	0.769	0.722