This is the accepted manuscript (post-print version) of the article.
The content in the accepted manuscript version is identical to the final published version, although typography and layout may differ.

How to cite this publication
Please cite the final published version:


Publication metadata

Title: The effects of performance systems and management concepts on organisational performance in the public sector in Denmark.
Author(s): Nielsen, Steen ; Nielsen, Christian.
DOI/Link: 10.1504/IJBSR.2020.106277
Document version: Accepted manuscript (post-print)
License:
The effects of performance systems and management concepts on organizational performance in the public sector in Denmark

Abstract
This paper explores how different types and groups of performance systems and management concepts explain the cross-sectional variations of organizational performance within the public sector, by applying a Structural Equation Modelling approach to survey data from 318 respondents in the Danish municipal sector. We argue that links exist between different types of performance systems and a number of performance outcomes. More specific, we explore how different performance measures such as benchmarking, employee significance, innovation, accounting models and systems, strategy execution and cost allocation directly or indirectly affect performance. In particular, our data suggests that the association between adoption of accounting systems and performance is fully mediated through both strategy execution and cost allocation. Practitioners can improve empirical specifications for examining the effects of different measures and gain an understanding of the mediating effects of accounting and management concepts. These findings add to our understanding of the functioning of performance management systems within the public sector in theory and in practice.

Keywords: Accounting systems, measures, management concepts, new public management, private sector, organizational performance, structural equation modelling, and survey.

JEL Classification: H83, M41, M48
**Introduction**

In the wake of the latter years’ performance management wave in public sector organizations (Van Dooren *et al.*, 2015), a growing body of academic literature has started examining the adoption and use of private accounting systems and tools. For example, different financial and non-financial metrics such as customer satisfaction, employee satisfaction, and the implementation of different accounting-based concepts and frameworks within the public sector (cf. Barzelay and Thompson, 2006). A common feature of these studies is that they are able only to provide weak statistical relations between a few non-financial performance measures and performance. Despite the academic focus on public sector performance, Schmidle (2011) and Taylor (2014) both argue that managers in the public sector in fact not readily use performance information. The empirical tests in these studies typically assume a uniform linear relationship between the non-financial, forward-looking performance measures and a specific measure of future performance. This study posits that such relationships are more complex and often work through intermediate and multiple linear constructs, as suggested by for example, Baekgaard (2015).

Sajay *et al.* (2008) highlight the importance of constructing performance metrics that contribute to improved administrative and political decisions, because, put simply, public organizations operate some of the world’s biggest spending budgets. One avenue for attaining this is by improving the quality of public sector performance information by developing performance systems in the public sector that can handle multiple performance indicators (van Thiel and Leeuw, 2016). New Public Management (NPM) is often cited as a means of achieving such better *Value for Money* in terms of improved resource management (Bouckaert and Halligan, 2008; Nielsen, 2016) and greater accountability of managerial decisions and actions in the public sector (Upping and Oliver, 2011).

NPM has been suggested as a ‘right way of reorganizing and structuring public sector organizations around the globe’, although recently there have been some suggestions that NPM are not as flawless as they may seem at first sight (Bouckart and Van Dooren, 2009; Lapsley, 2009; Charbonneau *et al.*, 2015). Consequently, public managers should not seek the one magic performance measure. Instead, they need to think seriously about the managerial purposes to which performance measurement might contribute and how they may ensure that changes in performance measures are coordinated with the organizational structures and related incentive structures (Thibodeau *et al.*, 2007; Northcott and Taulapapa, 2012). Hence, we posit the following research question:

*What are the effects of performance systems and their measures and management concepts on the performance of public-sector organisations in Denmark?*
We use Structural Equation Modelling to verify the effects of different types of measures and models for performance. For example, related to cost allocation; to strategy; the use of different management accounting systems; the use of benchmarking; the importance of employees; innovation and learning. Based on our literature review of performance and performance measurement in the public sector, we argue that the way these measures are interrelated and bundled may be seen as a very important factor for their actual effect on performance. In particular, we consider whether these relationships are direct or mediated by one another. Our model provides a rich setting that illustrates relationships that are much more complicated than previous research has been capable of showing.

A framework for measures and performance

The ability to allocate costs

The underlying rationale for the path entitled ‘the ability to allocate costs’ is that conscious allocation or assignment of costs creates transparency in resource use, and in the end, should, therefore, also affect both performance and learning in the organization (Verbeeten, 2011). This notion can be directly coupled to taxpayers’ demands for efficiency and effectiveness in the use of resources (Hood, 1995). In addition, transparency with regards to cost allocation has been shown to have positive effects on citizen attitudes towards public services (Baekgaard, 2015). Cost allocation models are built upon arguments such as creating transparency around decision-making (e.g. outsourcing decisions and benchmarking with other organizations) and legitimization of the organization’s activities vis-à-vis external stakeholders (Verbeeten, 2011; Pollitt, 2013). Diefenbach (2009) indicated that costs were important because a rising focus on cost-reduction, downsizing, competitive tendering, outsourcing, and privatization of services has led to the need for precise cost-information in the public sector.

Empirical evidence, relating to the diffusion of management accounting techniques also provides arguments for full cost allocation through systems such as ABC (Activity-based Costing), strategic cost management and resource management (Duncan et al., 2009; Lapsley and Wright, 2004). Cost management information is useful for general managers, operational managers or program managers, as it also provides insight on how to reduce costs and increase efficiency (Arnaboldi and Lapsley, 2005). Based on the summary of these insights, we posit that cost allocation is a critical precursor to the role of performance in the public sector, leading to the following hypothesis:

H1: There is a positive association between the ability to allocate costs and performance
Non-financial information and strategy execution

An increasing number of contributions suggest that non-financial measures linked to strategy can be applied to ensure that public sector managers continue to act in the best interests of society (Newberry and Pallot, 2004). In addition to this, empirical research indicates that even softer concepts such as mission statements (Bart et al., 2001), communication (Garnett et al., 2008) and the introduction of results-oriented organizational cultures make positive contributions to performance within the public sector (see, e.g., Hood, 1995; Merchant and Van der Stede, 2007).

These non-financial performance measures may affect performance directly or indirectly by supporting strategy execution (Groot and Selto, 2013; Merchant and Van der Stede, 2007). A specific complication in the public sector is that public organizations must reflect outcomes for multiple constituencies, leading to goal ambiguity (Kanter and Summers, 1987), and this is not always at odds with the often traditionally trust-based systems of the political-administrative context (Christensen et al., 2006). Due to these multiple constituencies, managers in the public sector have to include long-term measures (outcomes) and short-term measures (processes and activities performed), while at the same time focusing on efficiency and effectiveness.

Therefore, we examine whether non-financial performance measures are significant drivers of performance in the public sector. Prior literature suggests a positive correlation between strategy and performance measures when these are explicitly derived from the strategy and mission (Bart et al., 2001). However, other contributions have argued that the overall performance of public sector organizations actually decreases (Staw and Epstein 2000; Butterfield et al. 2004) as a reaction to widespread efficiency- and measurement-fever (Brignall and Modell 2000) and compliance focus (Power, 1999). Because of these seemingly contradictory arguments, the hypothesis is formulated as:

H2: There is a positive association between management control/strategy execution and performance.

Adoption of management accounting systems

The literature identifies two general objectives of introducing new accounting systems and models: first to increase the performance of the organization (Diefenbach, 2009; Van Helden, 2005) and second, to increase the transparency of the cost base (Verbeeten, 2011). Furthermore, Brignall and Modell (2000) argue that caution should be taken with direct transfer of private sector concepts to the public domain, because introducing new accounting systems also means introducing new values, which do not necessarily harmonize with the existing value set of the organization (see also Nacschold and Daley, 1999).
The literature suggests that management accounting systems and concepts carry with them a metaphor of rationalization and efficiency to the organization, but in addition, also one of customer service orientation and orientation towards the satisfaction of citizens. While the adoption of management accounting systems is a prerequisite for any form of effect on the organization, such adoptions may have negative consequences too (Brewer et al., 2003). The literature here suggests (see e.g. Maiga et al. 2014) that the adoption of accounting systems and concepts may have an effect on both cost allocation and on performance, leading to the following hypotheses:

H3a: There is a positive association between managers’ adoption of accounting systems and the ability to allocate costs.

H3b: There is a positive association between managers’ adoption of accounting systems and performance.

**Effect of benchmarking**

Benchmarking and NPM are often conjoined (Hood, 1991) as illustrated by Ammons and Rivenbark (2008). The benchmarking literature emphasizes the use of performance measures as an important and continuous source of information for evaluation of services against the best competitors or peers (Kouzmin et al., 1999) thus providing motivational and managerial effects (Behn, 2012; Hood, 2012).

This variable comes from an increasing interest in minimizing sub-optimization and ensuring goal congruence between employees and the organization and therefore also an expectation of increasing performance (Bowerman et al., 2002; Hope and Fraser, 2003). However, the obvious risk with benchmarking is that the ‘average’ rather than the ‘best practice’ is promoted as the norm (Behn, 2012; Llewellyn and Northcott, 2005). Benchmarking may also be used to focus attention on cost rationalizations and improvements in the long-term, even though problems exist in estimating both fixed and variable costs (Knutsson et al., 2012).

Even though our literature review reveals deficiencies concerning 1) analytical and empirically-grounded approaches to benchmarking, 2) frameworks for assessing benchmarking in terms of bottom-line improvements and 3) a distinct theory of benchmarking for the public sector (Dorsch and Yasin, 1998; van Helden and Tillema, 2005), some empirical evidence of the effect of benchmarking can be found. Benchmarking is applied for fulfilling different goals, for example, to reinforce economic pressures using benchmarking as targets, but also as an indicator for innovation and quality (Kouzmin et al., 1999). Together, these insights lead us to posit whether benchmarking of financial and non-financial measures is a critical precursor of innovation and learning, performance and the ability to allocate costs. This leads to the following hypotheses:
H4a: There is a positive association between benchmarking of financial and non-financial measures and innovation and learning.

H4b: There is a positive association between benchmarking of financial and non-financial measures and performance.

H4c: There is a positive association between benchmarking of financial and non-financial measures and the ability to allocate costs.

The significance of employees

In the public sector, where services are the main output, employees naturally play a major role in the value creation for the citizens (see for example Vargo et al., 2008). If an organization can show commitment to its employees by actively focusing on improving employee satisfaction and building their competences, this is expected to have a positive effect on the result (Huselid, 1995). Of course, such notions must be translated into better organizational performance through manifest-type variables such as the employees’ degree of understanding of organizational goals (Covaleski and Dirsmith, 1981) and how they individually contribute, as well as employee commitment to organizational goals (Latham et al., 2008).

These dependencies on employees suggest a positive sign between the perceived significance of the employee and organizational performance:

H5: There is a positive association between employee significance and performance.

Innovation and learning

Two important variables in the inovation management literature are innovation and learning and their importance for measuring performance. In a seminal study of innovation and learning, Hurley and Hult (1998) found that high levels of innovativeness were generally associated with organizational cultures that emphasize learning. A recent study by Naranjo-Valencia et al. (2011) showed that both innovation and learning contribute positively to business performance, while organizational learning as a separate construct affects innovation. Similarly, Oviedo-García et al. (2014) support the view that learning capability, especially knowledge flows and stocks, has a positive impact on the performance of public institutions. This relationship was also found by Calantone et al. (2002), whose results show that learning orientation and firm innovativeness affects firm performance.

For Behn (2012), fostering improvement is the primary reason for introducing performance measurement to organizations. An example of a managerial activity put in place to enhance innovation and learning is mentoring.
Hall and Smith (2009) showed that focus on mentoring programs in the organization helps to improve learning and reduces employee turnover. However, Diefenbach (2009), who argued that employees enrolled in mentoring programs induce negative consequences for organizational performance because such programs hinder the employees’ focus on increased efficiency, productivity and performance and instead encourage them to engage in knowledge-sharing and social activities that do not create value, found contradictory evidence.

Diefenbach (2009), who argued that employees enrolled in mentoring programs induce negative consequences for organizational performance because such programs hinder the employees’ focus on increased efficiency, productivity and performance and instead encourage them to engage in knowledge-sharing and social activities that do not create value, found contradictory evidence.

We choose to investigate innovation and learning as an aggregated measure consisting of different associations between citizens and employees. This is because it is often treated in this way in the management accounting and performance measurement literature (see, e.g., Hoque and James 2000). According to Henri (2006), innovation and learning reflects the capabilities of the organization where Mundy (2010) suggested that the ability to create dynamic tension between innovation, learning and controlling in itself becomes an enabling capability. Accordingly, we posit the following hypothesis in relation to innovation and learning:

H6: There is a positive association between innovation and learning and performance.

**Performance**

Performance is not a an unproblematic concept in the public sector. First of all this is so because of its ambiguous meaning in different environments and second, because of difficulties in measuring it as an output variable in different circumstances and cultures (Behn 2003; Theil and Leeuw, 2012). In the discussion of performance in relation to accounting systems and measures, the estimation of possible relations among different performance measures is often problematic because if a single metric is changed it will also affect the remainder of the metrics in the framework (Kaplan and Norton, 2001). In this study we use the word ‘performance’ as a broad outcome variable related to the effects of defining goals, selecting strategies, and rewarding performance (Ittner and Larcker, 2001; Kaplan and Norton, 2001).

Diefenbach (2009) also argues that performance in public sector organisations must invariably be associated with customer orientation and thereby the quality of service delivery as seen from the citizens’ perspective. In addition to service quality, increased organizational efficiency, effectiveness, and productivity defined and measured in technological terms are also important notions of public sector performance. In this study, we apply a compound variable consisting of two related types of elements: 1) quality of the services delivered to the citizen and 2) efficiency and effectiveness of core value-creating activities performed by the organization (Pollit 2013).

The conceptual model is illustrated in figure 1.
We use SmartPLS (Partial Least Squares Latent Variable Modelling Approach) Version 2.0-M2 (Ringle et al., 2005) to test our hypotheses and to examine measurement properties.

Research Methodology

Sample and characteristics

Respondents ranged from head of the budgeting and financial department, to controllers and to lower level employees and managers. On average, the respondents’ duration of employment in the organizations was 11.8 years (median 8 years) with a minimum of 1 year to a maximum of 40 years, and a standard deviation of 10.3 years. These data, therefore, suggest that respondents are well informed on their area of responsibility of the organization. This added up to approximately 1,000 potential respondents in Denmark. We followed the survey methodology proposed by Dillman (2000), who argues that because individuals are more likely to comply with the requests of familiar and authoritative sources, support of a survey may increase the response rate.

A pilot test was first undertaken with three representative respondents to refine the initial questionnaire. Information was gathered manually and each respondent was contacted personally in order to get feedback for the final structure and wording of the questionnaire. We received a total of 349 responses, a response rate of 35%. Eight of these were blanks, in that an electronic response from the particular respondent was registered but no answers were given. Twenty-three responses were deemed unusable due to an excessive proportion of answers in the ‘don’t know’ category, resulting in 318 usable responses and a usable response rate of about 32%, which seems to be on an acceptable level, because response rates seem to have decreased over the years (Baruch, 1999; Van der Stede et al., 2005). The questionnaire was developed based on literature review covering not only NPM but also performance management systems and measurement in general within the public sector.

Measurement

All questions, except the background and control questions, were measured using multi-item scales derived from existing studies whenever possible and formulated as statements on a 7-digit Likert scale ranging from 1 (‘Full disagreement’) to 7 (‘Full agreement’). Only items above the threshold of 0.60 are included (Gefen and Straub, 2005; Hair et al., 2010).
Cost allocation

The instrument for cost allocation was developed from Groot and Budding’s (2004) using different types of cost drivers. The factor analysis showed a five-item scale measuring the ability of cost allocations. In PLS, weights were assigned to variables that constitute a construct based on the overall relationship in the nomological model (Chin, 1998). The construct included allocation of labour, allocation of indirect costs across services, price changes based on cost allocations, allocation of variable costs, and calculation of costs associated with different services.

Non-financial information and strategy execution

This construct aimed to uncover whether non-financial performance measures related to the strategy execution of the organization is related to performance (Merchant and Van der Stede, 2007). Such performance measures and their relationships have been argued to be an important factor in firm performance and strategy (Hill and Snell, 2006; Langfield-Smith, 1998). The construct illustrates different types of links between the short and long run. The analysis supported the use of six items.

Accounting systems

This variable is related to an increasing interest in using new financial and accounting systems and concepts, inspired by surveys showing the used in the private company systems used in the public sector (Borins, 2001; Lapsley and Wright, 2004). The variable included Activity-based Costing, the Balanced Scorecard, Intellectual Capital, Business Excellence, and TQM (Total Quality Management).

Benchmarking

Benchmarking is one of the most used tools within the public sector (Behn, 2003). We used this variable to find out whether financial and non-financial measures used in the organization could be compared to a reference group for the sake of managing the organizational unit. The literature on beyond budgeting (Hope and Fraser, 2003) and reforming the public sector through private sector models (Brignall and Modell, 2000) advocate the advantages using benchmarking as a relative evaluation criterion. The variable included benchmarking of costs, of employees, of citizen, and of processes.

Employee significance

This variable is inspired by different measures used in the private service profit chain literature (Heskett et al., 1994, Loveman, 1998) and in the private theory of intellectual capital (Mouritsen and Larsen, 2005). In NPM we found evidence that employee attitudes to, for example, job satisfaction and organizational commitment are
closely linked to individual performance and thus to employee efforts to enhance attitudinal outcomes also resulting in important performance benefits (Noblet and Rodwell, 2009). This construct consists of five items, structured staff development interviews, employees’ satisfaction, employees’ training, and employees’ satisfaction with management.

**Innovation and learning**

Following Argyris (1976) and discussions about double-loop strategic learning and feedback (Kaplan and Norton, 2007), we measured this construct as a combination of, service levels and targets, citizen satisfaction, user satisfaction, quality of services, employee education, and the use of IT. We were interested in knowing whether the organizational unit had metrics that add value to performance. If the feedback result is unfavorable, the strategic learning should enable the organization unit to modify its strategy to reflect real-time learning (Kaplan and Norton, 2007).

**Performance**

Following previous calls for broader perceptions of performance (Dunk and Lysons, 1997; Verbeeten, 2008), this variable is constructed as a combination of expected factors and output metrics related to public sector organizations. This included aspects of efficiency as well as service effectiveness and quality towards affected citizens. The measure of overall performance consisted of a seven-item scale measuring the extent to which the organization is able to reach an acceptable level of outcome from its different input resources.

**Analysis**

A confirmative factor analysis with oblique rotation showed the items loaded on seven factors with eigenvalue greater than one. The measures for each construct were generated from previous research but also modified to fit the current research. We examined the dimensionality and subsequently eliminated all items below 0.60 (Barclay et al., 1995; Hair et al., 2010). All item loadings - except for one - were statistically significant at the 0.001 level. The appendix contains a full overview of the definitions and indicators used in the questionnaire for each construct (italic indicating eliminated items).

Table 1 shows the results supporting the inclusion of each of the constructs in the overall model.
We then used partial least squares structural equation modelling (PLS-SEM) to validate the measurement model and test our hypotheses. PLS-SEM is well-suited in research areas where no strong theory exists (Wold, 1980; Hulland 1999). PLS permits a two-stage assessment of the data. First, the measurement model, relating manifest variables to their latent variables. Second, the structural model, relating endogenous latent variables to other latent variables (Hair et al. 2010). The second stage is done through the use of bootstrapping (Efron, 1994).

The result of the PLS analysis path model is shown in Figure 2.

The measurement model

All item loadings were significant at the 1% level in relation to their construct (see Table 1). AVE (Average Variance Extracted) reflects the average communality for each latent construct and is used to establish convergent validity. In an adequate model, AVE should be greater than 0.5 (Chin, 1998; Höck and Ringle, 2006) which is the case here, except for the construct of ‘strategy execution’. Even though strategy execution and its importance is specifically discussed in the Balanced Scorecard literature (Kaplan, 2001) it seems that this effect has not entered into the public sector domain. The constructs of ‘benchmarking of financial and non-financial values’ and ‘employee significance’ and the underlying variables also seem to have caught the construct well (Garson, 2010). CR (Composite Reliability), a measure of reliability that is often preferred to Cronbach’s alpha, is above the acceptable limit of 0.7 for all constructs (Chin, 1998; Nunnally, 1978), indicating the degree of consistency (Hair et al., 2010).

As Table 2 shows, the square roots of AVE are all greater than the inter-construct validity (Fornell and Larcker, 1981). We used reflective measures for all constructs except for the control variables (budget control, annual appropriation, and the number of employees) which were viewed as formative measures. To secure non-multicollinearity, the correlations between variables in relation to the other constructs should be low enough, which is also the case here except for ‘performance’. This might not be surprising, because several of the performance measures within performance and performance management (and within management accounting) are not mutually exclusive (Chenhall, 2005).

In summary, all reliability and validity measures (convergent and discriminant) seem to be acceptable, supporting the hypothesized model.
The structural model

To test the structural model that includes the hypotheses, we run a PLS bootstrapping procedure (with 500 subsamples with replacement) and evaluate the path coefficients and their significance (Efron, 1994). The results are shown in Table 3 together with the $t$-statistics. We used path coefficients and $R^2$ to evaluate the model. Figure 2 shows that the $R^2$ for the dependent construct ‘performance’ has the predictive ability of 0.64, so that the model as a whole can be seen as a relatively good model (Gefen and Straub 2005; Hair et al. 2010).

The results also indicate that 35% of ‘innovation and learning’, 27% of ‘employee significance’, and 23% of ‘cost allocation ability’ can be explained by these constructs, whereas only 11% of ‘adoption of accounting systems’, 4% of ‘strategy execution’, and 3% percent of ‘benchmarking’ can be explained by these constructs.

Discussion and concluding remarks

This article set out to explore how different performance measures and management concepts affect performance in the public sector. Based on our statistical analysis, we have shown how different measures and accounting systems can be bundled to influence both the direct as well as the indirect effects of performance outcome. Our investigation of performance measures and their statistical associations in public sector organizations resulted in a number of interesting observations. Initially, our exploratory data analysis identified six constructs, which influence the performance of public sector organizations. These were benchmarking; innovation and learning; employees; accounting systems; strategy; and cost allocation.

Our data show a number of interesting paths. First, the paths of the model indicate a full mediation effect associated with benchmarking and performance, with indirect effects acting through employee significance, innovation and learning, and cost allocations. Contrary to existing literature (Kouzmin et al., 1999), the association between benchmarking and innovation and learning is, however, only partially mediated through cost allocation. We show that the benchmarking dimension should include both financial and non-financial measures relating to different aspects such as costs, employees, citizens, and processes, despite existing empirical research on benchmarking has almost exclusively considering the use of financial benchmarking (Tillema, 2005). Benchmarking seems to be a very relevant instrument because it serves a dual competitive and comparative purpose (Behn, 2012) while simultaneously supporting the sharing of best practice and thereby promoting learning (Northcott and Llewellyn, 2005; Knutsson et al., 2012). The full mediating effect on performance through both innovation and learning and cost allocation ability supports other authors’ considerations of the multi-oriented
problems with benchmarking, for example the rewarding of ‘average’ behavior and not best practice (Otley, 1999; Knutsson et al., 2012).

Second, our data also shows that the association between adoption of accounting systems and performance is fully mediated through strategy execution and cost allocation, but only partially mediated on cost allocation alone through strategy execution. The reason for the negative correlation may be that the accounting systems are not fully implemented such that instead of creating value and better decisions they create frustration and uncertainty. Evidence from existing research, however, suggests that accounting systems are thought to be helpful and therefore to improve performance (Sajay et al., 2008; Brignall and Modell, 2000). However, we could not prove any direct effect from such systems on performance. Instead, these accounting systems turned out to have a negative effect on cost allocation. One reason might be that the diffusion and implementation of accounting systems cause confusion rather than positive momentum. Brickley et al. (1997) also note that the introduction of accounting systems and concepts also risks leaving expectations unfulfilled because they are marketed too aggressively or because they remove focus from other (important) aspects. A third reason might be that this bundle of accounting concepts comprise a heterogeneous set of systems, that seen in combination have a negatively effect on performance. You either have these cost-accounting system and techniques, or not.

This result supports earlier findings that showed that the introduction of such accounting systems may create unintended effects and tensions and too many promises (Lapsley, 2009) such that decision-makers get the models mixed-up and thereby lose track of purpose and benefits (Behn, 2003) or simply do not obtain the desired results because these systems are too bureaucratic and inflexible (Pollit, 2013). The negative effects of accounting systems on both benchmarking and strategy execution are interesting and warrant future studies to determine whether this may be – or may not be - due to poor implementation cultures in these organizations or simply because they do not work in public organizations.

Third, our data also show that the employee significance construct, which included aspects such as employee satisfaction, qualifications, and training, turned out to be an important element for developing public sector performance, which in many respects may not be surprising. Recently, the importance of employee commitment in the public sector has been accentuated by Su et al. (2013). Our study contributes by showing, empirically, that this dimension is indeed an important element in the development of the public sector and it provides a unique contribution by illustrating that the employee dimension is an extremely important driver of improved future performance. Future research could, therefore, concentrate on a better understanding of the role of employees in value creation in the public sector or the issue that Hood (2012) calls his ‘second hypothesis: “how performance depends on the culture in which management by numbers takes place”.'
Fourth, we find a positive association between innovation and learning and performance and thus our analysis supports Mundy (2010) who views innovation and learning as an important prerequisite for performance. Aspects such as customer satisfaction and the quality of services provided proved to be important factors. Existing empirical research has tended to focus only on internal factors such as knowledge sharing and organizational memory (Jiménez-Jiménez and Sanz-Valle, 2011). Our results illustrate that the ability to allocate or assign costs and to focus on benchmarking performance provides a platform for innovation and learning, much in accordance with Behn (2012) and Henri (2006).

Fifth, our data indicate that the strategy execution dimension has no direct effect on performance, but our analysis illustrates that the negative impact of accounting systems on strategy is a sensitive one. However, we did find that strategy execution had a statistically significant positive effect on cost allocation ability. The strategy construct incorporates the links between service goals and the economic framework, measures of the overall goals, global challenges, and the importance of non-financial measures in general. Our results contradict the view that attention to strategy and mission within the public sector is important as the driving force of performance as proposed by Kaplan and Norton (2001) and Newberry and Pallot, (2004) and that the formulation of a mission has a positive effect on performance even though the impact might not be comprehensive (Bart et al., 2001). This might be due to the fact that decision-making in the public sector is often politically driven. This does, however, merit further investigation, as several recent studies focusing on productivity seem to claim that strategic alignment is a predefined goal.

Finally, our study supports the view that different types of accounting measures and dimensions can be used to improve and increase performance, although the effect seems to be limited (Ittner & Larcker, 1998; Cavalluzzo and Ittner, 2004). Critics have argued that NPM, and performance measurement ideologies in general, do not fit with the culture and mission of the public sector (Lapsley, 2009) and that managers in the public sector do not allocate sufficient resources into checking performance data (Hood, 2006). Others have highlighted the failure of NPM in one of the global frontiers, New Zealand, and have focused on how this country has moved towards a ‘post-NPM era’ system (Zafra-Gómez et al., 2009).

In conclusion, this article argues that the core of the problem of improving performance of public sector organizations might be found in the accounting and reporting systems and in the attitudes within the public sector. We believe that our findings are also interesting for practitioners. Public organizations around the world continue to struggle with the design and use of their performance systems and the structure and the associations between different measures and tools for decisions (Paulsson, 2012). Not only in the traditional way, but probably much more through today’s unprecedented “information explosion” as they address increasingly intertwined public
issues. Therefore, and as proposed by IBM, a historic opportunity exists to accelerate desired outcomes by embracing analytics as a core management competency (Messatfa et al., 2011).

This leads to the suggestion that these two factors are not independent and that changes in the number of measures might call for the knowledge of how these measures are grouped and interrelated at the same time. A change in the level of a single measure may affect a number of other measures and in the end also performance. This study provides evidence that improvements and increased performance are not achieved merely by implementing more concepts, but rather by carefully considering how these concepts interact with each other and the existing systems as well as with the corporate culture of learning and innovation. In fact, the introduction of advanced reporting systems, and the simultaneous use of too many accounting systems and concepts, will, therefore, often have a negative effect on performance.

Limitations

Research on the public sector is in an emergent state, which is why we have conducted our research mostly as an exploratory study and, therefore, have accepted the limitations to such studies.

First, the results of the path analysis only represent statistical associations based on a cross sectional survey generated by an iterative combination of principal components analysis that relates measures to constructs (and several ‘rules-of-thumbs’) which means that they are not necessarily sufficient to prove causal relationships (Chin, 1998). A specific problem may exist in this respect because outcome performance measures may normally be lagged measures (Kaplan and Norton, 2007), meaning that a dynamic test model could have been used instead (e.g. time-series-analysis or System Dynamics). Furthermore, while the evaluation of the measurement model in Smart-PLS provided some indication that our measures exhibit sufficient reliability and validity, it should also be taken into account that we did not allow for nonlinearities.

Second, using self-reported and perceived questions – instead of real data - in combination with the PLS-SEM methodology should be considered when drawing firm conclusions (Van der Stede et al., 2005; Hair et al., 2010; Ringle et al., 2005). Donaldson and Grant-Vallone, (2002), therefore, suggest that a conceptual framework for understanding factors that influence the motivation of an employee to bias his or her responses to questions (for example the so-called ‘battery effect’ for a construct) should be included by researchers. Such a framework has, however, not be used here.

Third, the use of buzzwords such as BSC, ABC or TQM to measure global challenges and the relation between citizens and employees might face the risk of a ‘modernity bias’, because as a public-sector manager you are expected to apply these techniques in order to avoid the risk of being accused of being old-fashioned. In
performance management and NPM, you have believers and non-believers. This may be a very strong underlying factor for the relations between factors.

Finally, although evaluation of the measurement model in Smart-PLS provided some indication to the effect that our measures exhibit sufficient reliability and validity, more research is needed to further establish these assumptions. Therefore, it is important to remember that no model is unique in the level of fit achieved and that for any model with an acceptable level of fits there often exists a number of alternative or equivalent models and competing ‘theories’ (Bagozzi and Philipps, 1982; Chin, 1998; Hair et al., 2010). Clearly, future research should address these limitations and try to establish whether our findings hold in other settings and with alternative methods and measures. A specific problem arising in the public sector is how to evaluate output performance. In this paper, performance is constructed as a combination of expected factors and output metrics based on six different items. However, a more specific performance measure – related to specific areas in the public sector – would be preferable.

In the words of Hibbert et al. (2014), generating theory in early-stage explorative settings may be seen as a relationally reflexive practice either by engaging alternative views outside of the prime community or by questioning existing methodological practices.
References


Appendix: Definitions and indicators by constructs

1. **Background information and control variables**
   - Job title and position
   - Professional training and seniority
   - Number of employees
   - Employee affiliation
   - Annual grant for the organization
   - Number of budget controls

2. **Questions regarding management and strategy execution (construct B)**
   The purpose of these questions is to uncover the link between financial management and strategic direction of the organization, including the degree of decentralization and frame control.
   - *The budgeting process is clearly defined and without any degree of freedom*
   - Linking of service goals and the economic framework control
   - Freedom to make budgets for each area
   - Measurements coupled to an overall goal of the organization
   - Measurements coupled with our global challenges
   - Importance of non-economic measurements in relation to the objectives and framework

3. **Questions regarding the degree of cost allocation in the organization: (construct A)**
   In the following we would like to know more about the concepts of cost distributions within the organization and thus the degree to which you judge that such distributions take place.
   - Allocate direct labour and direct material consumption for our services
   - Allocate all indirect costs across services (e.g. administrative costs)
   - Make price changes in the current financial system
   - Allocate all variable costs to the organization's services
   - Calculate costs associated with each type of service / product

4. **Measurements of the employee significance: (Construct E)**
   This category refers to the spread/ dissemination of different measures/metrics about the employees: it can be both financial and non-financial metrics.
   - Using structured staff development interviews with employees for improvements
   - Conducting regular employee satisfaction surveys
   - Employees and their qualifications
   - Measurement of employee training
   - Measurement of employee satisfaction with management

5. **Measurements relating to innovation and learning (construct F)**
   We are interested in knowing whether you have metrics related to its 'customers' and their satisfaction with the offered services and output.
   - We have measures for our actual use of information technology in our daily work
   - Service levels are key targets for our organization
   - Relation between citizens / users mix and demographics
   - Relation between citizen / user satisfaction services
   - Relation between citizens and employees through the use of Information Technology
   - Relation between quality of the services provided
   - Relation between performance and power on citizens' perception of service quality
6. **Measurements related processes and projects (construct G)**

In this category, we are interested in learning more about how you measure performance and outcome of your work. Such measurements may e.g. relate to the desire for workflows that are more efficient and the increase of outcome.

- *Measurement of project management in our organization*
- Measurement of all work
- Measuring efficiency changes in the processes as a result of managerial adjustments
- Measuring to the project level
- Using non-financial metrics
- Measuring the outcome of innovation and learning
- Assessing metrics related to quality

7. **Benchmarking of financial and non-financial metrics (construct D)**

In this category we are interested in whether the measurements/metrics used in the management of the organization, both financial and non-financial, are used in compared to a reference group.

- Benchmark costs and cost structure with a reference group of like-minded organizations
- Benchmark results from our employee surveys with like-minded organizations
- Benchmark results from our citizen surveys with like-minded organizations
- Benchmark results from our process measurements with like-minded organizations in the private sector

8. **Questions regarding the adoption of accounting systems (construct C)**

In the following, a number of systems/concepts from recent management accounting are mentioned. We are thus interested in whether these systems and concepts are adopted and included in your daily financial work or ongoing projects in the organization.

- Adoption and use of ABC
- Adoption and use of BSC
- Adoption and use of Intellectual Capital
- Adoption and use of Business Excellence and / or Commonwealth Asset Framework model (CAF)
- Adoption and use of Total Quality Management (TQM)
Figure 1: The study framework and path diagram
Table 1

Individual item loadings, eigenvalue, composite reliability, average variance extracted (AVE)

Panel A: The ability to allocate costs (composite reliability = 0.93, AVE = 0.72)

<table>
<thead>
<tr>
<th>Description</th>
<th>Loading</th>
<th>t-statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Able to allocate direct labor and direct material consumption for our services</td>
<td>0.85</td>
<td>14.58</td>
</tr>
<tr>
<td>Able to allocate all indirect costs across services (e.g. administrative costs)</td>
<td>0.87</td>
<td>19.62</td>
</tr>
<tr>
<td>Able to make price changes in accordance with our goals in current financial system</td>
<td>0.74</td>
<td>12.09</td>
</tr>
<tr>
<td>Able to assign all variable costs to the organization's services</td>
<td>0.90</td>
<td>27.43</td>
</tr>
<tr>
<td>Able to calculating costs associated with each type of service/product</td>
<td>0.87</td>
<td>16.87</td>
</tr>
</tbody>
</table>

Panel B: Linking management control and strategy execution (composite reliability = 0.81, AVE = 0.47)

<table>
<thead>
<tr>
<th>Description</th>
<th>Loading</th>
<th>t-statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Links between the organization's service goals and the economic framework control</td>
<td>0.63</td>
<td>3.67</td>
</tr>
<tr>
<td>Within the frame control we have freedom to make budgets for each area</td>
<td>0.71</td>
<td>2.90</td>
</tr>
<tr>
<td>Measurements are generally coupled to an overall goal of the organization</td>
<td>0.71</td>
<td>7.30</td>
</tr>
<tr>
<td>Measurements are generally coupled with our global challenges</td>
<td>0.69</td>
<td>6.91</td>
</tr>
<tr>
<td>Non-economic measurements are at least as important as economic measurements</td>
<td>0.66</td>
<td>6.28</td>
</tr>
</tbody>
</table>

Panel C: The adoption of accounting systems (composite reliability = 0.88, AVE = 0.60)

<table>
<thead>
<tr>
<th>Description</th>
<th>Loading</th>
<th>t-statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activity-Based Costing (ABC)</td>
<td>0.61</td>
<td>11.37</td>
</tr>
<tr>
<td>Balanced Scorecard (BSC)</td>
<td>0.82</td>
<td>35.54</td>
</tr>
<tr>
<td>Intellectual Capital</td>
<td>0.80</td>
<td>27.11</td>
</tr>
<tr>
<td>Business Excellence</td>
<td>0.78</td>
<td>23.04</td>
</tr>
<tr>
<td>Total Quality Management</td>
<td>0.83</td>
<td>34.97</td>
</tr>
</tbody>
</table>

Panel D: The effect of benchmarking (composite reliability = 0.94, AVE = 0.79)

<table>
<thead>
<tr>
<th>Description</th>
<th>Loading</th>
<th>t-statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benchmarking all costs and cost structure with a reference group of like-minded organizations</td>
<td>0.81</td>
<td>12.60</td>
</tr>
<tr>
<td>Benchmarking all the results from our employee surveys with like-minded organizations</td>
<td>0.94</td>
<td>46.80</td>
</tr>
<tr>
<td>Benchmarking all the results from our citizen surveys with like-minded organizations</td>
<td>0.92</td>
<td>30.19</td>
</tr>
<tr>
<td>Benchmarking all the results from our process measurements with like-minded organizations in the private sector</td>
<td>0.88</td>
<td>21.75</td>
</tr>
</tbody>
</table>

Panel E: Employee significance (composite reliability = 0.95, AVE = 0.82)

<table>
<thead>
<tr>
<th>Description</th>
<th>Loading</th>
<th>t-statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conducting regular employee satisfaction surveys in the organization</td>
<td>0.83</td>
<td>5.08</td>
</tr>
<tr>
<td>Measurement of our employees and their qualifications (e.g. Education, courses, etc.)</td>
<td>0.94</td>
<td>19.69</td>
</tr>
<tr>
<td>Measurement of employee training</td>
<td>0.93</td>
<td>7.40</td>
</tr>
<tr>
<td>Measurement of employees' satisfaction with management</td>
<td>0.92</td>
<td>5.11</td>
</tr>
</tbody>
</table>

Panel F: Measuring innovation & learning (composite reliability = 0.91, AVE = 0.66)

<table>
<thead>
<tr>
<th>Description</th>
<th>Loading</th>
<th>t-statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measuring relations between citizens/users’ mix and demographics</td>
<td>0.75</td>
<td>11.27</td>
</tr>
<tr>
<td>Measuring relations between citizen/user satisfaction with the services they receive</td>
<td>0.84</td>
<td>15.85</td>
</tr>
<tr>
<td>Accessing polls on citizens’ and employees through IT and opportunities</td>
<td>0.74</td>
<td>9.87</td>
</tr>
<tr>
<td>Measuring the quality of the services provided</td>
<td>0.86</td>
<td>18.46</td>
</tr>
<tr>
<td>Measuring performance and power on citizens’ perception of service quality</td>
<td>0.85</td>
<td>18.51</td>
</tr>
</tbody>
</table>

Panel G: Performance (composite reliability = 0.93, AVE = 0.68)

<table>
<thead>
<tr>
<th>Description</th>
<th>Loading</th>
<th>t-statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measuring the outcome of all our work</td>
<td>0.85</td>
<td>10.96</td>
</tr>
<tr>
<td>Measuring the outcome of efficiency changes</td>
<td>0.83</td>
<td>11.80</td>
</tr>
<tr>
<td>Measuring the outcome of projects</td>
<td>0.70</td>
<td>9.93</td>
</tr>
<tr>
<td>Measuring the outcome by using non-financial metrics</td>
<td>0.88</td>
<td>18.21</td>
</tr>
<tr>
<td>Measuring the outcome of innovation and learning</td>
<td>0.80</td>
<td>14.93</td>
</tr>
<tr>
<td>Measuring the outcome of quality by the use of metrics</td>
<td>0.87</td>
<td>18.78</td>
</tr>
</tbody>
</table>

*All t-statistics reported in this table are statistically significant at p < 0.001, except one statement in the Management Control construct (2.9) with p<0.01.
*t-statistic > 1.64 is significant at p < 0.10 level
**t-statistic > 1.96 is significant at p < 0.05 level
***t-statistic > 2.58 is significant at p < 0.01 level
****t-statistic > 3.29 is significant at p < 0.001 level

Figure 2: Results of the PLS path model estimation
Table 2
Inter-construct correlation and square root of average variance extracted statistics

<table>
<thead>
<tr>
<th>Constructs with reflective items</th>
<th>Manifest variables with formative items</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Benchmarking</td>
<td>(1) (2) (3) (4) (5) (6) (7)</td>
</tr>
<tr>
<td>(2) Cost allocation</td>
<td>0.89</td>
</tr>
<tr>
<td>(3) Employee significance</td>
<td>0.51 0.12 0.91</td>
</tr>
<tr>
<td>(4) Innovation and learning</td>
<td>0.55 0.42 0.34 0.81</td>
</tr>
<tr>
<td>(5) Accounting systems</td>
<td>-0.18 -0.21 -0.15 -0.17 0.77</td>
</tr>
<tr>
<td>(6) Performance</td>
<td>0.62 0.42 0.64 0.62 -0.13 0.82</td>
</tr>
<tr>
<td>(7) Strategy execution</td>
<td>0.32 0.38 0.31 0.36 -0.20 0.39 0.69</td>
</tr>
<tr>
<td>(8) Number of budget controls</td>
<td>-0.03 -0.08 -0.01 -0.04 0.13 -0.05 -0.12</td>
</tr>
<tr>
<td>(9) Amount of annual grant</td>
<td>-0.05 -0.16 -0.02 -0.12 0.27 -0.07 -0.09</td>
</tr>
<tr>
<td>(10) Number of employees</td>
<td>-0.03 -0.19 0.03 -0.11 0.32 -0.04 -0.08</td>
</tr>
</tbody>
</table>

Diagonal elements are the square roots of AVE. Off-diagonal elements are the correlations between the latent variables calculated in Smart-PLS. AVE is only suitable when constructs are measured with reflective indicators (Barclay et al., 1995).

Table 3
Path coefficients: test and control variables

<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>Path coefficient</th>
<th>t-statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benchmarking -&gt; Cost allocation (H4c)</td>
<td>0.26</td>
<td>1.96**</td>
</tr>
<tr>
<td>Benchmarking -&gt; Innovation and learning (H4a)</td>
<td>0.45</td>
<td>3.50****</td>
</tr>
<tr>
<td>Benchmarking -&gt; Performance (H4b)</td>
<td>0.16</td>
<td>1.57</td>
</tr>
<tr>
<td>Cost allocation -&gt; Performance (H1)</td>
<td>0.17</td>
<td>1.79*</td>
</tr>
<tr>
<td>Employee significance -&gt; Performance (H5)</td>
<td>0.42</td>
<td>2.81***</td>
</tr>
<tr>
<td>Innovation and learning -&gt; Performance (H6)</td>
<td>0.32</td>
<td>2.01*</td>
</tr>
<tr>
<td>Adoption of accounting systems -&gt; Cost allocation (H3a)</td>
<td>-0.10</td>
<td>2.75***</td>
</tr>
<tr>
<td>Adoption of accounting systems -&gt; Performance (H3b)</td>
<td>0.05</td>
<td>1.05</td>
</tr>
<tr>
<td>Strategy execution -&gt; Performance (H2)</td>
<td>0.04</td>
<td>0.55</td>
</tr>
<tr>
<td>Control Constructs: Benchmarking -&gt; Employee satisfaction (No hyp)</td>
<td>0.51</td>
<td>4.22****</td>
</tr>
<tr>
<td>Adoption of accounting systems -&gt; Benchmarking (No hyp)</td>
<td>-0.18</td>
<td>3.95****</td>
</tr>
<tr>
<td>Adoption of accounting models -&gt; Strategy execution (No hyp)</td>
<td>-0.20</td>
<td>4.77****</td>
</tr>
<tr>
<td>Strategy execution -&gt; Cost allocation (No hyp)</td>
<td>0.28</td>
<td>2.17**</td>
</tr>
<tr>
<td>Control Variable: Number of budget control</td>
<td>0.19</td>
<td>1.84*</td>
</tr>
<tr>
<td>Amount of annual grant</td>
<td>-0.08</td>
<td>0.40</td>
</tr>
<tr>
<td>Number of employees</td>
<td>0.88</td>
<td>3.60****</td>
</tr>
</tbody>
</table>

*t-statistic > 1.64 is significant at p < 0.10 level; **t-statistic > 1.96 is significant at p < 0.05 level; ***t-statistic > 2.58 is significant at p < 0.01 level; ****t-statistic > 3.29 is significant at p < 0.001 level.