

Performance Management and Evaluation

Hansen, Morten Balle

Published in:
Handbook of Social Policy Evaluation

Publication date:
2017

Document Version
Accepted author manuscript, peer reviewed version

[Link to publication from Aalborg University](#)

Citation for published version (APA):
Hansen, M. B. (2017). Performance Management and Evaluation. In *Handbook of Social Policy Evaluation* Edward Elgar Publishing.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Performance Management and Evaluation (Morten Balle Hansen)

Preprint of chapter to appear in Handbook of Social Policy Evaluation:

Hansen, M. B. 2017. Performance Management and Evaluation. In B. Greve (Ed.), *Handbook of Social Policy Evaluation*. London: Edward Elgar Publishing.

"Performance Management" (PM) will be defined more extensively below, but are essentially about management with a focus on results, the results meaning more than just the effects of an intervention. There is some criticism of the PM (e.g. Petersen, 2008; Pollitt 2013), much of which is justified, but this chapter's overall approach to public PM is that it (1) is necessary, (2) is here to stay and (3) have both advantages and disadvantages. Politicians, managers and employees in the public sector therefore do well to familiarize themselves thoroughly with the PM in order to increase its advantages and minimize its disadvantages.

Many accounts of the development of Performance Management (PM) in the public sector starts with New Public Management (NPM) in the 1980s (Hood, 1991) and finds that the public sector at the beginning of the 21st century is in "... an era of governance by performance management "(Moynihan, 2008, p.3). But public sector PM goes further back and as early as in 1943 the American city manager association (International City Management Association) published a book about measuring municipal performance (Ridley & Simon, 1943). Outside the public sector the Scientific Management Movement (Kanigel, 1997; Taylor, 1992) from the early 1900s is a well-known and controversial early version of PM.

Although PM has a long history, its development in the public sector has been more in focus in the past 30 years of management reforms around the world - especially among the relatively wealthy OECD members. Within social policy the increasing use of PM has for instance been observed in labour market participation programs (van Berkel and Knies 2016) and in eldercare provision (Hansen and Vedung 2005). The trend to make increasing use of PM is not unique to the public sector and has also been subject to debate in the private sector (Mintzberg, 2009), but here we will concentrate on the public sector.

Performance Management has had different names in the international public policy and administration literature such as Management by Objectives and Results (MBOR) (Lægreid et al 2006), Management By Objectives (MBO) (Swiss 1991) and performance measurement (Kuhlmann 2010). These names indicate differences in perspective and emphasis but the approach here is to highlight their similarities and in this chapter the most common international term - performance management (PM) – is used.

The purpose of the chapter is, first, to give an introduction to PMs idea and practice in the public sector. Second, to clarify the similarities and differences between PM and other forms of evaluation. And third, to analyse the challenges that characterizes the PM in the public sector. Although the focus is on the public sector, many of these challenges are of a more general nature and relevant to many private sector organizations as well.

In what follows first the basic characteristics of PM are analysed. Second, PMs characteristics are compared to other forms of evaluation. Third the three basic elements of PM are analysed in turn based on three basic questions: (1) What is performance in the public sector? (2) How is performance measured in the public sector? (3) How are performance measures used in public sector processes of decision-making and organizing? Fourth, the key challenges and criticisms related to PM-systems are analysed and discussed and finally a brief conclusion is given.

What is Performance Management?

Similar to the concept of evaluation, there is no one official definition of public sector PM as everyone agrees. A sort of lowest common denominator is that PM is characterized by "the development and integration in policy and management of performance measurement systems" (Bouckaert & Peters, 2002: 359). A more comprehensive definition, which most PM researchers could also agree on, is that PM is a system in which:

" ... Organizations and individuals are given objectives, measurable targets are derived from the objectives, and then a wide variety of instruments of authority or incentive are deployed to encourage staff to hit or exceed their targets." (Pollitt, 2013: s.346-347).

Although PM is defined in different ways in the literature, virtually all definitions include the integration of three fundamental processes in the routines of the organization: (1) a process in which the most important goals, targets and performance criteria are identified and agreed on. In many PM models it is called a process of goal formulation. (2) a process of gathering relevant information on these performance goals. A key effort is to develop relevant key performance indicators (Key Performance Indicators, KPI). These KPI's are usually but not necessarily quantitative. A basic challenge in this process is to ensure a meaningful (hopefully) translation of fundamental objectives to measurable indicators. (3) a process where the generated performance data is used to motivate, make decisions and to document and improve performance.

All public sector PM systems therefore raise three basic descriptive questions:

1. What is performance in the public sector?
2. How is performance measured in the public sector?

3. How are performance measures used in public sector processes of decision-making and organizing?

Different PM systems have developed different answers to these three questions and a widespread criticism of the theory and practice of PM systems has been that Question two (performance measurement) has received too much attention at the expense of Question one (formulation and clarification of performance) and three (use of performance measures) (Marr, 2009).

Performance Management and Evaluation

Comparing the conceptualization of PM above with some of the most common definitions within evaluation research, it becomes clear that PM can be seen as a special form of evaluation with some additional features or amendments related to governance and management.

The US evaluation researcher Michael Scriven provide a kind of minimal definition of evaluation as "... the process of Determining the merit, worth, and value of things" (Scriven, 1991: 1), while the Swedish evaluation researcher Evert Vedungs definition of evaluation is more compact " ... careful retrospective assessment of the merit, worth, and value of administration, output, and outcome of government interventions, which is intended to play a role in future, practical action situations" (Vedung 1997, p.3).

The first PM-process, which establishes performance targets, thus corresponds to define the standards from which the value of an activity can be evaluated. The second PM-process, in which specific indicators are defined and routinely measured, corresponds to the generation of data in an evaluation process. The third PM process in which performance data are used to make decisions, motivate employees and document performance to the public corresponds to the efforts of evaluation researchers' to make decision-makers use the findings of their evaluations. And just as in the case of evaluations, there are many examples from the PM literature that the data generated in the PM systems are not used or are used in ways creating unintended side effects.

The additional features of PM, as compared with evaluations, are, firstly, that PM-systems are embedded in the routines of the organization. This enables in principle and often in practice the PM-systems to have great impact on organizational priorities and activities based on the assumption: "what you measure is what you get" (Kaplan & Norton 1992, p.71). Secondly, a managerial perspective characterizes PM-systems. It is the challenges of management that are in focus and PM-systems are developed as tools for management, which is not necessarily the case for evaluations. Thirdly - and closely related to the previous two characteristics - PM-systems are characterized by a focus on how motivation and incentive structures related to the PM system affects the organization's employees and their activities. Due to the integration of PM in the routines of the organization they can provide powerful signals to guide employee priorities. For instance within the area of social policy Riccucci et al found that "[S]taff who

are monitored on employment tasks and outcomes, for example, believe that promoting employment is a high priority, while those who are monitored on claims processing tasks not only believe that eligibility determination is most important, they also give lower priority to employment goals.”(Ricucci et al. 2004:446).

Some PM-models, particularly models originally developed to the private sector, is also characterized by a focus on the individual- and organisational level (for instance the performance of a teacher or a school), while performance at other more aggregated levels are neglected. For instance performance at the meso-level of a policy sector (e.g. the performance of an elementary school system) or the macro-level of a country (for instance the performance of the entire public sector) is neglected in some PM-systems.

The dynamic relationship between the micro, meso and macro performance is, however, a central issue in the public sector, where performance in many areas depends on interdisciplinary collaboration and the impact of "multilevel governance" (Bouckaert & Halligan, 2008; Pollitt & Bouckaert, 2011).

The discussion above indicates that PM, in relation to the evaluation literature, can be conceptualized as an institutionalized evaluation system (Hansen 2013) designed as a tool for management.

Since the PM is thus a form of institutionalized evaluation, PM has some of the same purposes as other evaluations: (1) Learning in order to improve; (2) Learning in order to control (internal) and document (external); and (3) Learning in order to enhance general societal knowledge development (Vedung, 2009). In addition PM, like evaluations, can be used in political games on resources, can act as a sort of Potemkin scenery that belies performance (Vedung, 2009, pp. 175-176, who call it the strategic purposes). PM can also be an expression of a ritual reflection (we need to) that in the short term does not really have any real impact, but in the longer term can get quite tangible consequences (Dahler-Larsen, 2011).

What is performance in the public sector?

Performance can in principle take any goals that a management sets out to achieve as its points of reference, but in the public sector, there are some institutionalized performance criteria that recurs in most public organizations. There are two key choices concerning performance that public decision-makers, elected politicians as well as bureaucrats, rarely completely decide by themselves. These choices are often the subject of disagreement between actors, but must be taken when it comes to delineate and define what counts as performance (M. B. Hansen, 2008; Vedung 1997). One key choice can be termed a "performance criterion" and the other a "point of reference". Performance criteria are about from which criteria a judgement about performance are made. A "point of reference" is about who or what you compare with (see Table 1).

Table 1. Institutionalized performance criteria and points of reference in the public sector

Performance criteria	Points of reference
1. Economy (ressources utilized)	1. Goals (Did we accomplish our goal?)
2. Efficiency (costs per delivered unit)	2. The past (Are we better or worse than last year? (e.g. higher or lower efficiency))
3. Effectiveness (outcome/impact – often related to costs (cost-effectiveness))	3. Intra-national (Are we better or worse than other comparable public organisations?)
4. Equity and fairness (e.g. equal possibility, due process, equal distribution)	4. International (Are we better or worse than other comparable countries?)
5. Responsiveness (e.g. democratic decision process, involvement, user surveys)	5. Benchmark (How do we score compared to best empirical practice?)
6. Quality (e.g. of the output, the outcome, the process)	6. Minimum/Optimum (How do we score compared to some predefined minimum/maximum values?)
Note: The two lists are adapted from Hansen 2008 and Vedung 1997. To assess performance at least one criterion (left column) and one point of reference (right column) must be used. Many different combinations of criteria and points of reference can be combined in an overall assessment of performance.	

The two lists in Table 1 are not exhaustive, but they indicate a general version of some of the most commonly used performance criteria and points of reference in the public sector. The list of performance criteria gives rise to refute the widespread misconception that performance is just about effects (see eg Ejler et al, 2008: 12). There is no doubt that a significant effort in public administration since the 1980s NPM-wave has been to develop more and better measures of outcome (effects). But there is also not the slightest doubt that a key performance objective in the public sector in many countries has been to keep costs down or shrink the public sector (Feigenbaum, Henig and Hamnett 1998). Either understood as attempts to rollback the welfare state as the Thatcher reforms in the UK (Hood & Dixon 2015) or as attempts to limit the growth of public expenditure as in other public management reforms (Ejersbo & Greve, 2005; Pollitt & Bouckaert, 2011).

Compared to classic economic performance criteria this is an expanded list, with economic performance typically only including the top three performance criteria – often called the three E's of performance in the international performance literature (Johnsen 2005; Liu et al, 2010). A review of the literature in public administration and political science, will, however, make it clear that the list is not exhaustive (see, eg, Boyne, 2003b; Christensen, 2003; Rothstein, 2011; Rothstein & Teorell, 2008). Public performance is a very complex multidimensional concept that is not reducible to a simple input-output-outcome model (Pollitt & Bouckaert, 2011; Rosenbloom, 1983), although such a model can be quite useful (Vedung, 1997). The institutionalized performance goals above are goals which is difficult to ignore to public decision-makers in the long run (although they may be forgotten in the short

run) and that are more or less routinely generated in most policy areas, both at the organizational, the national and the international level. Many resources are devoted to develop reliable and comparable indicators for these institutionalized performance criteria. In principle, and sometimes in practice, each of the performance criteria at the left in table 1 can be combined with any one of the reference points at the right in table 1. For instance economy (costs) may be compared to (1) goals, (2) the past, (3) other public organisations, (4) other countries, (5) a benchmark of best empirical practice, or (6) some optimum or minimum value.

Economy is the term that is most easily measured (although it is certainly not unproblematic) and which has been a priority for decades from both the EU and changing governments. Paradoxically, the NPM movement with the slogan of "from input to output management" articulated cost management as old-fashioned at the same time that it has been given high priority. But such paradoxes are well known in management research (Mintzberg 2009). The public sector spending must be controlled so that they are not going to be too large a share of the economy (how large "too large" is, is of course a matter of debate). Targets are set and adopted in budgets that specifies the acceptable cost level for different activities. Often changes in cost over time are monitored and compared intra- and internationally. Internationally such comparisons are usually at the country and sector levels (e.g. size of public sector and costs of eldercare).

Efficiency is probably the classic economic performance goals used most frequently. It is often called "productivity" in the international literature, (see. Johnsen, 2005; Pollitt & Bouckaert, 2011). It is sometimes called internal efficiency. Activities in school can for example be calculated as the cost per student or cost per delivered lecture. Reference points are often the productivity of other municipalities or international comparisons. To make sense of such comparisons an "all else being equal" (ceteris paribus) assumption are often used, presuming that the quality of the service is the same. Alternatively indicators of quality are included - in elementary schools often grade level. Efficiency has recently been put on the agenda as a key performance target (Productivity Commission, 2013) in the Danish society, but it has been the a key performance criteria for many years in most countries. The focus on efficiency has been a core feature of the NPM movement in recent decades as is also preeminent in social policy areas such as elder- and healthcare (Bertelsen and Rostgaard 2013, Hansen and Vedung 2005, McIntyre and Gilson 2002)

Effectiveness is another frequently used performance criteria, where the outcome of an intervention are assessed and possibly related to the cost of the effort (cost-effectiveness) (Pollitt & Bouckaert, 2011). It is often regarded as the primary performance criteria, but it represents some difficulties. One reason is that reliable and relevant performance indicators are often far more difficult to get and quantify than cost (economy) and productivity (efficiency) indicators. Another difficulty is that the outcome of an intervention often depends

heavily on the context in which the intervention is carried out. A third difficulty is that the outcome of an intervention can change dramatically over time. From a management perspective, this performance criteria also has the disadvantage that a management has dominion over only a small part of the many factors that determine the outcome of an intervention (Scott & Davis, 2007: 326-328). Nevertheless, the performance criteria is in principle very important and many resources are spend to develop better "cost-benefit" and "cost-effectiveness" measures to calculate effectiveness. The entire evidence movement (H. F. Hansen & Rieper, 2009) is an ambitious attempt to improve public policy by enhancing knowledge of the reletive outcome of differen types of intervention.

Equity and fairness (Boyne, 2003a) is another widespread performance criteria in the public sector, which is related to the concept of justice. It has many different more or less operational variants. Part of equitys many dimensions are different criteria of equality, such as a more equitable social and geographical distribution of wealth in society, equality before the law, equal access to, equal utilization of and equal outcomes of services. The Gini coefficient is an example of a standardized indicator designed to measure the degree of economoc equality in a society and is used in historical and international comparative analyzes. Issues related to equity dimensions of performance in social policy are numerous and tend to be rediscovered once in a while – partly as a reaction to a too dominant efficiency agenda (McIntyre and Gilson 2002). For instance within public health “... horizontal equity refers to the equal treatment of equals, vertical equity refers to the unequal (but equitable) treatment of unequals.” and most “equity strategies within health systems have focused on establishing mechanisms for achieving horizontal equity in health care delivery (such as ‘equal access for all’ or ‘universal access’) whilst recognising the importance of vertical equity in relation to health care payment mechanisms (i.e. payment on the basis of ability).” (McIntyre and Gilson 2002:1639)

Responsiveness is part of the normative starting point for a democracy and is also embedded in PM systems through user committees and satisfaction surveys, etc., while *quality* in a PM system among other things is to deliver the promised quality of public services on time (Hansen & Olsen, 2004; Ministry of Social Affairs, 2002). How responsiveness and quality are included in PM-systems vary considerable between organizations and policy areas. For instance in Danish eldercare the elected municipal politicians responsible for delivering eldercare are by law required to formulate clear and transparent quality standards concerning different types of eldercare (e.g. cleaning and practical help) and these standards are used as performance criteria both by the elderly recipients of care and their interest organizations and by the municipality when controlling the delivery of services (Hansen and Vedung 2005).

Other candidates for the institutionalized performance criteria in table 1 are *economic growth* (e.g. measured in GDP), *openness* and *innovation*, and of course a wide range of more specific performance goals that vary from organization to organization and from policy sector to policy sector.

The above discussion of what counts as performance in the public sector shows how the establishment of performance standards in individual public organizations must be reconciled with the larger public context in which it is embedded. The research in evaluation and public performance in recent decades has helped to develop and refine the conceptual and empirical understanding of the performance concept and its application within public organization and management on both a micro (Boyne, 2003b) and macro level (Pollitt & Bouckaert, 2011; Rothstein, 2011).

How is performance measured in the public sector?

The above review of the key institutionalized performance criteria in the public sector also gives a hint of how performance is measured. Data are generated both internationally (for example the PISA studies in elementary school), nationally (e.g. the national statistics organizations) and in the specific organization (eg local satisfaction among citizens and employees).

International performance measures are generated and developed routinely by organizations like the UN, the World Bank, OECD, EU and numerous other organizations. These measures are used for various purposes including ranking of countries and organizations in multiple performance dimensions. "The Quality of Governance" Institute at Gothenburg University has done a great useful work by gathering and preparing many of these data for analytical use and make them available to researchers and other interested parties (<http://qog.pol.gu.se/>).

National performance measures are often based on raw data generated in the national statistics organizations (e.g. Statistics Denmark). Raw data (many based on public organizations' reporting of budgets and accounts) to many of the public performance indicators in many countries are thus generated in the National Statistics Organizations and are often publicly available statistics (e.g. in Denmark www.statistikbanken.dk). These data are often processed and compiled to more accessible actual Key Performance Indicators (KPI).

Organizational performance measures are developed and generated partly in the public organizations themselves and partly in the sector (e.g. labor market, health, education), in which the organization operates. Many employees work full-time to generate and develop these data and present them in an understandable way to decision makers and other stakeholders.

There are numerous challenges associated with measuring performance, as it would lead too far to go into here, but some major challenges can briefly be summarized. Key Performance Indicators (KPI's) can in principle be based on qualitative data (observations, interviews), but

in practice they generally consists of quantitative indicators to count the magnitude of a phenomenon. Thus, there is a tendency for a number of well-known imbalances or distortions in the way performance is measured: (1) It is easier to measure quantity than quality, which imply the danger that the importance of quality is given too little attention in PM systems; (2) It is easier to measure output (activities) than the outcome (effects) which often imply that output in PM practice is given priority due to all things equal assumptions concerning outcome and (3) It is easier to measure input (costs) than output (Mintzberg, 1989) which may sometimes imply that cost arguments are given priority over efficiency. In addition, (4) it is easier to measure short term than long term outcome, implying a tendency to myopia and for many reasons (5) there are also quite often errors in the data generated - both in public and private organizations (Mintzberg, 2009).

Since the decision context in the public sector is bureaucratic and often characterized by overload, there is a danger that decisions are made mechanically with too little sober judgement and excessively based on numbers, which poorly illustrate the fundamental performance of the activities of an individual, an organisation or a program. The danger of being too much obsessed with formal performance measures is also wellknown from private corporations. Under the heading "Where Has All the Judgment Gone?" the management thinker Henry Mintzberg suggest the rule-of-thumb to "... measure what you can, but then be sure to judge the rest: don't be mesmerized by measurement." (Mintzberg 2009:225).

How are performance measures used in public sector processes of decision-making and organizing?

Generally, it is not well documented how the measurement of performance is used in decision-making in the public sector. A review of 20 years of research (Pollitt, 2006) called it "a bit amazing" that there is so little analysis of what elected politicians line-up with performance information. Since then a few important studies has been made (Askim, 2007, 2009, Askim, Johnsen, & Christophersen, 2008), but the overall picture is still true. We have many studies of how performance management affect frontline bureaucrats and service production (see below), but very little about how the PM is used for top-level political and administrative decision-making.

Learning, documentation and perverse impact – the strength and weaknesses of PM

Above we discussed PMs general characteristics, history and relation to other forms of evaluation. But what exactly is the rationale for using PM? To use a term from evaluation research one may ask: what is the intervention theory behind PM's? An intervention theory consists in its most developed form of a situation theory, an impact theory and a normative theory (Hansen & Vedung, 2005, 2010), but we will here concentrate on the normative theory.

The PM literature has proposed various rationales for the use of PM systems (Behn, 2003; De Bruijn, 2007; Kristiansen, 2014), where particularly the following can be highlighted:

1. *Transparency*: PM systems can create more openness and transparency of the activities and results of public organizations', which is an important accomplishment in itself in a public sector in which all citizens have a legitimate interest in the activities and the results generated.
2. *Learning*: Thanks to the transparency PM systems can help public organizations learn from what they do well and less well in relation to the performance criteria they have developed. Learning can also be to revise, downplay or prioritize certain performance targets differently in light of the findings from the performance measures. In principle the learning purpose of PM is extremely important and can hardly be overestimated. Although perhaps overstated there is an important insight in the premise that "Organizations which do not not receive one sort of signal when they do something right, and another sort of signal when they do something wrong, will in the long run do more wrong things than right." (Rothstein 1998:201). In practice the matter is of course more complicated as we shall see below.
3. *Documentation*: PM-systems makes it possible in principle to document relative performance to the relevant stakeholders of the organization, which in the public sector in principle is all citizens. Such documentation of relative performance would combine selected performance criteria and selected reference points such as lower costs compared to last year or other comparable organizations.
4. *Motivation*: PM-systems make it possible to link performance with incentive structures so that good performance is rewarded and poor performance punished. Such rewards can take the form of both "soft" (more exciting assignments) as "hard" (pecuniary bonus) incentives. This may create motivation to make an extra effort in relation to the prioritized performance targets.
5. *Performance Culture*: PM systems can contribute to a more performance-oriented culture among public employees, where employees develop ownership of the performance objectives developed in the organization.

There are many examples of the above advantages of the introduction and development of public PM systems. In many case PM-sysems in practice by and large and up to a point works as intended. But there are also a number of well-known challenges associated with these systems. Among other things, there seems often to be a trade-off between the need for internal learning and external documentation, where the two purposes of the system have difficulties in being served by the same PM system. Sometimes these challenges leads to a kind of decoupling in which the PM systems are not really used properly and the investments does not really pay off. In other cases they have some unfortunate side-effects.

Many of the difficulties of PM systems are linked to the importance of the paradox associated with two theses: (1) what you measure is what you get and (2) you can not measure, what you really want to acomplish. There will almost always be more or less difference between the

principal objectives an organization sets out to (or is connected to) achieve and the operational Key Performance Indicators (KPI), it is possible to get on activities, output and outcome (Dahler-Larsen, 1998; Mintzberg, 1989) and very much of the PM literature can be seen as a discussion of the importance of the two theses. To what extent are they right on various policy areas? What implications do they have? How do you handle these implications?

There are many examples from case studies in PM research that indicates that there can be quite serious negative implications of PM systems, although the extent and significance are inherently difficult to identify (De Bruijn, 2007; Pollitt, 2013). Some of the most important are briefly explained below:

1. Goal displacement: If there is a difference between the goals of the organisation and the indicators measuring goal achievement, so the possibility arises for goal displacement away from the original objectives, towards the measurable goals. The more recent history is rich in anecdotes about the unfortunate impacts of goal displacements. From Soviet Russian planned economy (the performance of furniture production where measured in weight resulting in unusually heavy furniture) to the American war in Vietnam (accounts of the number of deaths as an indicator of success on the battlefield proved illusory). More generally, it is usually easier to measure quantity than quality, easier to measure cost than activities easier to measure activities than outcome, and easier to measure short-term effects than long-term effects. Bluntly put, the PM systems thus if taken at face value have a tendency to promote a short-term, quantitative focus on costs and activities, while long-term, qualitative outcome will be downgraded.

2. Down-grading of innovation: Innovation is about creation and deployment of something new in order to create some kind of improvement. But although innovation is increasingly recognized as essential in public administration, the nature of PM systems tend to have a conservative bias in favor of the tried and tested well-known technologies. The short-term experience with most innovation processes is that the performance goes down for a period (Repenning & Sterman, 2001). All beginnings are difficult and it requires time and patience before you can begin to benefit from innovation. Furthermore, the long-term gain will be uncertain, and implies a certain risk. The short-term profit orientation that characterizes many PM systems thus, other things being equal, tends to imply an under-investment in particularly more radical forms of innovation.

3. Creaming and other types of gaming: Various types of gaming are well known phenomena in both public and private organizations (Lipsky 1980) and PM systems are likely to encourage this kind of behavior unless they are designed to counteract it (De Bruijn 2007). Creaming refers to one "foaming" the cream by taking care of the "easy", self-propelled clients / customers, which makes it easy to create fast measurable successes, while rejecting the difficult "heavy" resource-demanding clients. For instance the probability of success for a course in job search measured as a percentage share of having a job after a few weeks, will be much greater if the participants are short-term unemployed with a higher education than if

they are long-term unemployed with no or little education. The same can be said about the productivity of any surgical operation, if you are able to send patients with low chance of successful operation on to other units. PM-systems are at risk at lowering the professional ambitions, by enhancing this type of creaming behavior in order to score cheap points in the PM system. It is difficult to measure the extent of creaming and other types of gaming in quantitative research since such controversial issues are hard to get valid answers to in questionnaires and because the impact of PM systems is difficult to distinguish from other factors such as caseloads (See however van Berkel and Knies 2016), but examples of the importance of gaming from qualitative research in social work are numerous (Brodkin 2015, Pollitt 2013, Thorén 2008).

The examples above and many other well-known challenges illustrate that PM systems can have quite significant adverse consequences. Based on the literature (e.g. Bouckaert & Peters, 2002; De Bruijn, 2007; Hondeghem & Van Dorpe, 2013; Hood, 2007, 2012; Marr, 2009; Pollitt, 2013) a number of attention points regarding the development and adaptation of PM systems (see table 2) can be suggested.

On the one hand these attention points can hopefully be of some use in attempts to increase the positive and reduce the negative impacts of the increasing importance of PM systems in the public sector. On the other hand they at least provide a summary of some of the most important difficulties facing the design of public sector PM-systems

At a very general level, the literature on the design of PM systems provide some reasonable rule-of thumbs and suggestions concerning how to answer some of the above questions in the public sector (e.g. De Bruijn, 2007; Marr, 2009) as well as the private sector (e.g. Kaplan & Norton, 1992, 2007). The answers can provide some guidance, but they can rarely if ever stand alone and they require creative intelligent translation to the specific context public policy makers are facing. Within the public sector there are furthermore huge differences between how a sensible PM-system might look like in different policy sectors.

Table 2: Attention points when designing public Performance Management Systems

Elements in PM systems	Attention points related to elements in PM systems
Defining performance	<p>1: How to define and forge a consensus on performance criteria in a political system where the key decision makers - elected officials – are partly elected on the basis of marked disagreement?</p> <p>2: How to select and prioritise among many desirable but often competing objectives that involve trade-offs and dilemmas?</p> <p>3: Who should be involved in the definition and formulation of performance criteria?</p>
Measuring performance	<p>4 How to delimit and construct reasonably accurate and valid indicators of the decided goals ?</p> <p>5: How to develop appropriate performance measures for complex tasks completed in advanced professional bureaucracies and tasks completed in the network of mutually dependent but formally different networks of actors?</p> <p>6: Which one (s) target group (s) should the performance measures primarily be related and adapted to? The public? The political decision-makers? The professionals performing the tasks? Others?</p>
Utilizing performance measures	<p>7: How should the performance measures be linked to incentives and sanctions for employees and others?</p> <p>8: How to avoid or minimize known adverse impact linked to PM systems, such as "creaming", "gaming" and "Teaching to the test"?</p>
Cross-cutting attention points	<p>9: How are tensions and discrepancies between goals and performance indicators at different systemic levels (individual, organizational, industry and society) handled, so that problems with suboptimal behavior are avoided or minimized?</p> <p>10: How are trade offs between short-term and long-term goals in PM systems handled in order to ensure a proper balance between for example long-term innovation and capital expenditures and short-term bottom line and daily operations?</p> <p>11: How to prioritise and handle possible tensions between the three overarching objectives of public PM systems: To provide external documentation, internal control and learning to improve performance?</p>

Conclusion

Performance Management is one of the most influential forms of governance, developed in the public sector in the wake of the New Public Management reforms of the 1980s. Different versions of PM permeates today the public sector everywhere and represent a cornerstone in most conceptions of public management. PM is a form of institutionalized evaluation system that attempts to integrate certain modes of management and evaluation in the routines of the

public sector. PM consists of the integration of three processes: (1) Strategic political processes where the understanding of performance are elaborated and development and overall objectives are formulated. (2) Processes in which data on key performance indicators related to overall objectives routinely are generated. (3) Processes in which the generated data are used in decision-making control, documentation, learning and motivation. The chapter has explained the main characteristics of PM and the challenges associated with the development and adaptation of PM systems in the public sector. The basic approach has been sought to be critical but constructive. There is much justified criticism of PM systems and their application in the public sector. But nevertheless, the chapter's thesis is that PM-system are necessary and are here to stay. Politicians, managers, employees and other stakeholders in the public sector therefore do well to familiarize themselves thoroughly with the PM in order to increase its benefits and minimize its costs.

References

- Askim, J. (2007). How do politicians use performance information? An analysis of the Norwegian local government experience. *International Review of Administrative Sciences*, 73(3), 453-472.
- Askim, J. (2009). The Demand Side of Performance Measurement: Explaining Councillors' Utilization of Performance Information in Policymaking. *International Public Management Journal*, 12(1), 24-47.
- Askim, J., Johnsen, A., & Christophersen, K. A. (2008). Factors behind organizational learning from benchmarking: Experiences from norwegian municipal benchmarking networks. *Journal of Public Administration Research and Theory*, 18(2), 297-320.
- Behn, R. D. (2003). Why measure performance? Different purposes require different measures. *Public Administration Review*, 63(5), 586-606.
- Bertelsen, TM and T Rostgaard. 2013. "Marketisation in Eldercare in Denmark: Free Choice and the Quest for Quality and Efficiency. I: Meagher, G. & Szebehely, M.(Red.) Marketisation in Nordic Eldercare: A Research Report on Legislation, Oversight, Extent and Consequences." *Stockholm Studies in Social Work* 30.
- Bouckaert, G., & Halligan, J. (2008). *Managing Performance. International Comparisons*. London: Routledge.
- Bouckaert, G., & Peters, B. G. (2002). Performance measurement and management: The Achilles' heel in administrative modernization. *Public Performance & Management Review*, 25(4), 359-362.
- Boyne, G. A. (2003a). Sources of Public Service Improvement: A Critical Review and Research Agenda. *Journal of Public Administration Research and Theory*, 13(3), 367-394.
- Boyne, G. A. (2003b). What is public service improvement? *Public Administration*, 81(2), 211-227.
- Brodkin, E. Z. 2015. "The inside Story: Street-Level Research in the Us and Beyond." Pp. 25-42 in *Understanding Street-Level Bureaucracy*, edited by P. Hupe, M. Hill and A. Buffat. Bristol: Policy Press.
- Christensen, J. G. (2003). *Velfærdsstatens Institutioner*. Århus: Aarhus Universitetsforlag/Magtudredningen.

- Dahler-Larsen, P. 2011. *The Evaluation Society*. Stanford: Stanford University Press.
- Dalsgaard, L., & Lassen, M. (1996). Policydannelse og præstationsmåling: Aalborg Universitetsforlag.
- De Bruijn, H. (2007). *Managing Performance in the Public Sector* (2 ed.). London: Routledge.
- Ejersbo, N., & Greve, C. (2005). *Moderniseringen af den offentlige sektor*. Copenhagen: Børsens Forlag.
- Ejler, N., Seiding, H. R., Bojsen, D. S., Nielsen, S. B., & Ludvigsen, F. (Eds.). (2008). *Når måling giver mening – Resultatbaseret styring og dansk velfærds politik i forvandling*. København: Jurist- og Økonomforbundets forlag.
- Feigenbaum, Harvey, Jeffrey R Henig and Chris Hamnett. 1998. *Shrinking the State: The Political Underpinnings of Privatization*: Cambridge University Press.
- Hansen, H. F. (2013). Systemic evaluation governance. New logics in the development of organisational fields. *Scandinavian Journal of Public Administration*, 16(3), 47-64.
- Hansen, H. F., & Rieper, O. (2009). The evidence movement the development and consequences of methodologies in review practices. *Evaluation*, 15(2), 141-163.
- Hansen, M. B. (2008). *Offentlig Organisation og Forvaltning. Strukturer, processer og resultater*. Retrieved from Odense:
- Hansen, M. B., & Olsen, G. V. (2004). *Projekt Evaluering af Fælles sprog: Fælles sprogs anvendelse i kvalitetsstandards i den kommunale ældrepleje*. Odense: Institut for Statskundskab, Syddansk Universitet.
- Hansen, M. B., & Vedung, E. (2005). *Fælles sprog i ældreplejens organisering*. Odense: Syddansk Universitetsforlag.
- Hansen, M. B., & Vedung, E. (2010). Theory-Based Stakeholder Evaluation. *American Journal of Evaluation*, 31(3), 295-313.
- Hondeghem, A., & Van Dorpe, K. (2013). Performance management systems for senior civil servants: how strong is the managerial public service bargain? *International Review of Administrative Sciences*, 79(1), 9-27.
- Hood, C. (1991). A Public Management for All Seasons. *Public Administration*, 69(1), 3-19. Retrieved from
- Hood, C. (2007). Public service management by numbers: Why does it vary? Where has it come from? What are the gaps and the puzzles? *Public Money & Management*, 27(2), 95-102. Retrieved from
- Hood, C. (2012). Public Management by Numbers as a Performance-Enhancing Drug: Two Hypotheses. *Public Administration Review*, 72(s1), S85-S92.
- IKAS. (2013). *Den Danske Kvalitetsmodel, DDKM – kvalitet i bevægelse*. Retrieved from Århus: Indenrigsministeriet, Ø.-o. (2014). *Kommunal udligning og generelle tilskud 2015*. København: Økonomi- og Indenrigsministeriet.
- Johnsen, A. (2005). What does 25 years of experience tell us about the state of performance measurement in public policy and management? *Public Money & Management*, 25(1), 9-17.
- Kanigel, R. (1997). *The one best way. Frederick Winslow Taylor and the Enigma of Efficiency*. New York: Viking.
- Kaplan, R. S., & Norton, D. P. (1992). The Balanced Scorecard - Measures That Drive Performance. *Harvard Business Review*, 70(1), 71-79.
- Kaplan, R. S., & Norton, D. P. (2007). Using the balanced scorecard as a strategic management system. *Harvard Business Review*, 85(7-8), 150-+.

- Kristiansen, M. B. (2014). *Resultatstyring i den offentlige sektor*. København: Jurist- og Økonomforbundets forlag.
- Kuhlmann, S. 2010. Performance Measurement in European local governments: a comparative analysis of reform experiences in Great Britain, France, Sweden and Germany. *International Review of Administrative Sciences*, 76(2): 331-345.
- Lipsky, M. 1980. *Street-level Bureaucrats - Dilemmas of the Individual in Public Service*. New York: Russell-Sage Foundation.
- Liu, W., Cheng, Z., Mingers, J., Qi, L., & Meng, W. (2010). The 3E methodology for developing performance indicators for public sector organizations. *Public Money & Management*, 30(5), 305-312.
- Læg Reid, P., Roness, P. G., & Rubecksen, K. 2006. Performance Management in Practice: the Norwegian Way. *Financial Accountability & Management*, 22(3): 251-270.
- Marr, B. (2009). *Managing and Delivering Performance. How government, public sector and not-for-profit organizations can measure and manage what really matters*. Oxford: Elsevier.
- McIntyre, Di and Lucy Gilson. 2002. "Putting Equity in Health Back onto the Social Policy Agenda: Experience from South Africa." *Social Science & Medicine* 54(11):1637-56.
- Mintzberg, H. (1989). A Note on that Dirty Word "Efficiency" *Mintzberg on Management* (pp. 330-334). London: The Free Press.
- Mintzberg, H. 2009. "Managing Effectively." Pp. 195-235 in *Managing*. London: Financial Times, Pentice Hall.
- Moynihan, D. P. (2008). *The Dynamics of Performance Management. Constructing Information and Reform*. Washington D.C.: Georgetown University Press.
- Opstrup, N. (2014). *Causes and Consequences of Performance Management at Danish University Departments*. (PhD), University of Southern Denmark, Odense.
- Petersen, V. C. (2008). *Vildveje i velfærdsstaten: fællesskab i opløsning, styringsillusioner, udveje*. København: Informations forlag.
- Pollitt, C. (2006). Performance Information for Democracy: The Missing Link? *Evaluation*, 12(1), 38-55.
- Pollitt, C. (2013). The logics of performance management. *Evaluation*, 19(4), 346-363.
- Pollitt, C., & Bouckaert, G. (2011). Results: through a glass darkly. In C. Pollitt & G. Bouckaert (Eds.), *Public Management Reform. A Comparative Analysis: New Public Management, Governance and the Neo-Weberian State* (3 ed., pp. 126-160). Oxford: Oxford University Press.
- Produktivitetskommissionen. (2013). *Styring, ledelse og motivation i den offentlige sektor*. Retrieved from København:
- Repenning, N. P., & Sterman, J. D. (2001). Nobody ever gets credit for fixing problems that never happened: Creating and sustaining process improvement. *California Management Review*, 43(4), 64-88.
- Riccucci, N. M., M. K. Meyers, I. Lurie and J. S. Han. 2004. "The Implementation of Welfare Reform Policy: The Role of Public Managers in Front-Line Practices." *Public Administration Review* 64(4):438-48.
- Ridley, C. E., & Simon, H. A. (1943). *Measuring municipal activities: A survey of suggested criteria for appraising administration*: The International City Managers' Association.
- Rieper, O., Ejersbo, N., Hansen, P., Antonsen, M., & Jørgensen, T. B. (2000). Målstyring i kommunerne *Forandringer i Teori og Praksis - skiftende billeder fra den offentlige sektor* (pp. 179-214): Jurist-og Økonomforbundets forlag.

- Rosenbloom, D. H. (1983). PUBLIC ADMINISTRATIVE THEORY AND THE SEPARATION OF POWERS. *Public Administration Review*, 43(3), 219-227.
- Rothstein, Bo. 1998. *Just Institutions Matter. The Moral and Political Logic of the Universal Welfare State*. Cambridge: Cambridge University Press.
- Rothstein, B. (2011). *The quality of government: Corruption, social trust, and inequality in international perspective*. Chicago: University of Chicago Press.
- Rothstein, B., & Teorell, J. (2008). What is quality of government? A theory of impartial government institutions. *Governance-an International Journal of Policy and Administration*, 21(2), 165-190.
- Scott, W. R., & Davis, G. F. (2007). *Organizations and Organizing. Rational, Natural and Open System Perspectives* (Pearson International Edition ed.). Upper Saddle River, New Jersey: Pearson Prentice Hall.
- Scriven, M. (1991). *Evaluation Thesaurus*. Newbury Park: SAGE Publications.
- Swiss, J. E. 1991. Management by Objectives, ***Public Management Systems: Monitoring and Managing Government Performance***: 61-104. New Jersey: Prentice-Hall, Inc.
- Taylor, F. W. (1992). Scientific Management. In J. M. Shafritz & A. C. Hyde (Eds.), *Classics of Public Administration* (Vol. 3, pp. 29-32). Belmont, California: Wadsworth Publishing Company.
- Thorén, Katarina H. 2008. "Activation Policy in Action": A Street-Level Study of Social Assistance in the Swedish Welfare State: ProQuest.
- van Berkel, R. and E. Knies. 2016. "Performance Management, Caseloads and the Frontline Provision of Social Services." *Social Policy & Administration* 50(1):59-78. doi: 10.1111/spol.12150.
- Vedung, E. (1997). *Public Policy and Program Evaluation*. New Brunswick: Transaction Publishers.
- Vedung, E. (2009). *Utvärdering i Politik och Förvaltning*. Lund: Studentlitteratur.