

**Leading the Future of the Public Sector: The Third Transatlantic Dialogue**  
**University of Delaware, Newark, Delaware, USA**

May 31–June 2, 2007

**Workshop 5: Leadership and the New Public Management**

**Getting to performance through leadership,  
reform and measurement:  
A theoretical exploration of some social  
mechanisms.**

Wouter Van Dooren

**Contact for correspondence:**

Dr. Wouter Van Dooren  
University of Leuven  
Parkstraat 45 PB3609  
3000 Leuven  
Belgium

Wouter.vandooren@soc.kuleuven.be  
0032/16/323603

**ABSTRACT**

This text studies the theoretical relation between performance on the one hand and leadership, reform and measurement on the other. Three kinds of social mechanisms that explain the relationship between the aforementioned phenomena (leadership/reform/ measurement) and leadership. First, we studied mechanisms representing conventional wisdom (x determines performance). Secondly, we looked at what we called counterintuitive mechanisms (x is a consequence of performance). Thirdly, we searched for social constructivist approaches (x as performance). All three approaches yield sensible hypotheses.

## Introduction

Much of the performance debate in the last decades was centered around the issues of reform, measurement and leadership. This has led to vigorous debates between believers and non-believers on the virtues of reform, measurement and (entrepreneurial) leadership (Williams, 2000). These three dimensions will also be part of our analysis. All three can be studied at the organizational level as well as on the societal level.

*Reform.* Government reformed in the name of performance. In particular, in the UK and the USA, this led to cutback management and a reduction of the size of government (Dunleavy, 1986). Other countries followed other trajectories. Pollitt and Bouckaert (2004; Bouckaert & Peters, 2002) identify four strategies: to minimize (privatize), to marketize (bringing private sector techniques and values into government), to modernize (changing public sector techniques and values) and to maintain (using the old techniques more intensely). The societal demand for a high-performing public sector filters through to the organizational level.

*Measurement.* Bouckaert & Peters (2002) argue that performance measurement is the ‘Achilles Heel’ of many public sector reforms. The availability of performance information is a necessary - not sufficient - condition for the success of many reform initiatives. Yet, often the availability of performance information is assumed. The presence of performance information is one of the most decisive and susceptible aspects of the recent tide of public management reforms.

*Leadership.* Reforms have also challenged traditional leadership (Kaboolian, 1998) (Cook, 1998). Management reforms have advocated an entrepreneurial role for public managers. Entrepreneurial leaders, according to Osborne and Gaebler, cut red tape, empowered employees, make customer satisfaction a priority, and transformed their organizations into lean, mission-driven, results-oriented enterprises (Terry, 1998). Entrepreneurialism is linked to improved performance. Obviously, this new leadership role is not without critique. Critics charge that the public entrepreneur's anti-traditionalist orientation, rule-breaking, power politics, risk-taking, and radical change conflicts with democratic theory (Terry, 1998).

*Linkages between reform, measurement and leadership discourse.* Arguments about reform, leadership and performance are often part of the same discourse. A central tenet of reforms is that enabling entrepreneurial leadership (“to let the managers manage”) requires more performance measures in order “to make the managers manage” (Kaboolian, 1998). (Lynn, 2006), in a review of Radin's book ‘challenging the performance movement’ (Radin, 2006), notes that the performance movement is about governmental reform, has multiple, complex, and hard-to-measure consequences, some unfortunate, some useful, some ambiguous for government performance broadly defined (p.636).

This being said, by studying these dimensions, other dimensions are left out of the analysis. For instance, implementation (including issues of street level bureaucracy) will be profoundly related to performance. Structural dimensions, such as decentralization, will also affect performance. Resources (financial, personnel, infrastructure, ...) and performance are another duo. Other factors are the ability to cooperate in networks, the cultural aspects, organizational capacity, ... Just to mention a few.

The text is organized as follows. First, we discuss the way we deal with theory. We use the metaphor of social mechanisms, which is based on Merton's Middle Range theorizing. The following three chapters explore the social mechanisms in the triangle

# Social mechanisms

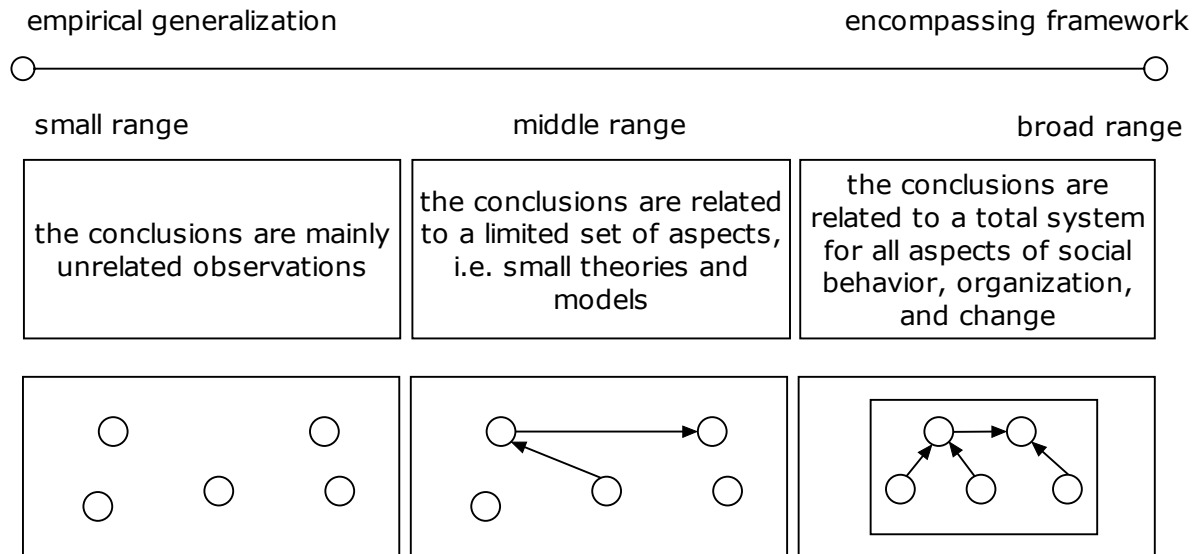
## *Middle range theorizing*

The metaphor of a social mechanism points to the ambition to go beyond ‘black box’ thinking; put more A and less B in and performance will come out (for an elaborate discussion of social mechanisms, see Hedstrom & Swedberg (1998). Such recipe style conclusions are typical results of statistical correlations. Every introductory methodological handbook warns that correlation is not causation. A correlation does not tell anything about why, how, when and where organizations get to performance. Social mechanisms are precisely these processes that explain why correlations are observed.

Secondly, the metaphor of a social mechanism refers to the middle range character of the this kind of theorizing (Merton, 1968). According to Merton, theories can be distinguished by their intrinsic characteristics. The extreme positions of the range are the total theoretical systems on the one hand and the necessary working hypotheses that evolve in abundance during day-to-day research on the other hand (Figure 1).

Middle range theories differ from daily empirical generalizations because the middle range theories consist of assumptions that make it possible to deduct testable hypotheses. A good theory acquires theoretical value through its use. The usefulness of the theory may be tested by its capacity to posit new questions. These questions can be both theoretical and empirical.

Middle range theories differ from the total theories because the former only explain one specific aspect of social behaviour, social organization and change<sup>1</sup>. Theories of the middle range are not necessarily deducted from an encompassing framework, but they may often fit in afterwards. Comprehensive sociological theories are sufficiently loose-knit, internally diversified and mutually overlapping that a given theory of the middle range, which has a measure of empirical confirmation, can often be subsumed under comprehensive theories, which are themselves discrepant in certain respects (Merton 1968: p. 43).



**Figure 1: Small, middle and broad range theories (Van Dooren, 2006).**

A total system in which observations about every aspect of social behaviour, organization, and change find their place is according to Merton not feasible yet . Unlike the specialization in the physical sciences, differentiation in

social sciences yielded theoretical systems that are held to be mutually exclusive. Rather than a division in tasks, specialization in the social sciences leads to a competition between the specialisms.

This paper is concerned with middle range theories that explain the relation between performance and respectively leadership, reform and measurement. We will explore three kinds of mechanisms. Two have a realist slant. Some represent conventional wisdom while others are more counterintuitive. Thirdly, we will discuss some mechanisms with a constructivist flavour.

### ***Conventional wisdom***

A first kind of social mechanisms can be categorized as conventional wisdom. Performance is in these theories the dependent factor that can be explained by studying leadership, reform, measurement or other dimensions of organisations or societies. The proposition is generally structured as follows;

X (Leadership, Reform, Measurement, ...) → Y (Performance)

One of the most fertile research traditions studying the conventional wisdom is the “management matters for performance – hypothesis”<sup>ii</sup>. These studies want to see whether quality and the nature of management does affect performance of organisations vis-à-vis other factors such as fiscal, legal and political constraints (Wilson, 1989). This is a fundamental issue for public management research and practice. Moynihan & Pandey (2005) for instance show how external environmental influences and internal management factors combine to create performance. Cogburn & Schneider (2003) place managerial capacity alongside other more commonly studied state characteristics such as public interest group activity, government ideology, and citizen ideology as an important influence on government activities. Like many other authors, data from the Syracuse Government Performance Project are used. Nicholson-Crotty & O’Toole, Jr. (2004) found amongst others that active management improves arrest rates.

Policy evaluation is another field that mainly follows the conventional line of reasoning. In particular the discourse on evidence based policy making (what works?) is about what matters for performance (for an overview of the British scene, see Davies et al. (2000)).

### ***Counterintuitive mechanisms***

It is always useful to challenge the conventional wisdom; not to do away with it, but to avoid academic common mode failure. The counterintuitive mechanisms social mechanisms we want to describe in this paper reverse the conventional direction of the causal linkage between independent and dependent factors. Performance is seen as the independent variable that explains leadership, reform, measurement. The propositions thus are structured as follows;

X (Performance) → Y (Leadership, Reform, Measurement, ...)

Sociologists have some useful concepts that make use of this counterintuitive line of reasoning. Merton formulated two middle range theories of this sort. A first example is the theory of the self-fulfilling prophecy stating that an initially false definition of the situation evokes new behaviour that makes the originally false conception come true (Merton, 1949) p. 423). Another counterintuitive middle range theory, also mainly acclaimed to Merton, is the so called Matthew effect, stating that those who have will get more<sup>iii</sup>. Deleeck et al., (1983) applied this theory on social policy, discussing the phenomenon that in welfare states the middle classes tend to be the main beneficiaries of social benefits and services, even if these are primarily targeted at the poor.

For what our interest, performance in the field of public administration, is concerned, similar hypotheses can be formulated. Are those organisations that perform well, more likely to perform better in the future? Will organisations that are said to perform well, change their behaviour in order to comply with the initial praise?

## ***Social constructivist mechanisms***

A third type of mechanisms is social constructivist. The main characteristic of this kind of theory making is its subjective notion of performance. Unlike the other two types, concepts such as performance are seen as constructed by human beings rather than an objective reality. Performance does not necessarily relate to real behaviour. If we take this perspective, the relation between performance, leadership, reform and measurement alters fundamentally. Rather than looking for causal relations between the realities, the social mechanisms should explain how performance, as a construct, comes into being.

Constructivist arguments are well established in organization studies, a theoretical field where Public Administration and public management often relies on. Neo-institutionalism for instance use a great deal of constructivist arguments. Scott (2001) identifies three kinds of institutions; cognitive, normative and regulative. He argues that regulative institutions imply a social realist approach while cognitive institutions are shared definitions of social reality, and thus are socially constructed. Normative institutions fall in between, but incline towards the constructivist side. DiMaggio & Powell's renowned (1983) study of what makes organisations so similar is a good example this tension between different kind of social mechanisms – or in their words isomorphic processes. The first process, coercion is a realist one. Organisations are confronted with rules and power and act in accordance with these rules. The other processes, mimetic and normative, are social constructivist. Mimetic processes are about imitation of those organisations that are perceived to be successful. Normative processes concern the shared beliefs of professionalism that are acquired through education.

The volume 'translating organizational change', edited by Czarniawska-Joerges & Sevon (1996) is a good outline of the constructivist argument. Organisational life is seen as an ongoing process of materialization of ideas, of turning ideas into objects and actions and again into other ideas. According to Czarniawska-Joerges and Sevon, ideas are translated rather than diffused. This is not a trivial semantic discussion. It touches on the fundamental difference between social constructivist and realist approaches. Diffusion, a realist notion, implies that there is a objective and real idea that disperses in the same way that, in physics, oxygen diffuses in the air. The constructivist notion of translation does away with the 'objective and real' idea. Ideas are inter-subjective and only exist as shared understandings of realities.

## **What is Public Performance?**

### ***Performance; a conceptual lacuna in the performance literature***

Before we look at the relationship between performance and any other variable, we need to define what performance of public sector organisations means. The performance literature is mainly about performance measurement and performance management. Less consideration has been given to the definition of performance in itself. Of course, performance is conceptualized as outputs and outcomes and its corollaries efficiency and effectiveness. Yet, the question remains, what are outputs and outcomes? Little conceptual work has been done to enrich the notion of performance. This lacuna may be filled by the literature on public sector values.

Both public values and public performance are said to be public because there is a collective benefit (and not a mere individual benefit) to it. For both public performance and public values, the critical point is the exact definition of collective benefit. The public values research has sought to define what generally is regarded as being of collective benefit (De Bruijn & Dicke, 2006). The performance literature may draw on the public values literature to also better define the collective benefit that performance managers are expected to pursue.

Jorgensen & Bozeman (2002) list 13 public values. Further research on the "public values universe" led to a list of about 80 public values (Jorgensen & Bozeman, 2007). A particularly useful distinction is developed by Hood (1991). He distinguishes between sigma type values, theta type values and lambda type values.

Lambda-type values are set to keep the public sector robust and resilient. Government has to keep operating even in adverse 'worst case' conditions and to adapt rapidly in a crisis (p14). Reliability is often an argument for choosing public production instead of private production.

Theta-type values intend to keep government fair and honest. Government has to pursue honesty, fairness and mutuality through the prevention of distortion, inequity, bias, and abuse of office (p.13). These values are institutionalized in appeal mechanisms, public reporting requirements and ethical codes.

Sigma-type values allege to keep it lean and purposeful - to match resources to defined tasks. Thus, frugality of resource use in relation to given goals is the criterion of success, while failure is counted in terms of instances of avoidable waste and incompetence (p12).

### ***Public performance and public values***

The next question then is how public values and public performance conceptually relate to each other. Performance can be seen as one public sector value amongst other values. Alternatively, performance and values are seen as a distinct concepts, that however share a need to define the public character.

The most common approach is the former. Performance is seen as one value amongst others. In Hood's framework, performance would roughly be equivalent to sigma-type values. This approach is taken in Beryl Radin's challenge of the performance movement. It is in essence a critique of a set of predominant public values that is promoted by people in academia and government and society. The objects of the study is a movement. A movement is a informal form of organization (Jary and Jary 1999). Movements are bound together by shared thoughts and practices. The performance movement is promoting values such as efficiency, effectiveness and accountability and is proposing practices such as performance measurement, performance management and the like. Radin is part of another movement, with different thoughts and practices and promoting different values. She argues amongst others that the performance movement forgets about the context, interferes with professionalism, is not concerned with equity, is a-political.

There is a lot of truth in her observations. However, one substantial critique on the critiques of the performance movement is its conceptual vagueness. As we argued above, this is also a result of the some conceptual lacunas in the performance literature itself. Yet, the question remains. What exactly has been criticized? Performance measurement (counting)? Performance pay? Performance contracting? The publication of performance information? Benchmarking? NPM reforms? Agencification? The policy management divide? What exactly are we talking about?

An alternative approach is to see performance and values as distinct concepts that share a need to define the public character vis-à-vis the private. Performance is in this respect an analytical concept rather than a label for a set of values. This approach has some advantages.

First, it allows to use the theoretical power of the concepts. Dubnick (2005) points to the many instances of performance; *"outside of any specific context, performance can be associated with a range of actions from the simple and mundane act of opening a car door, to the staging of an elaborate reenactment of the Broadway musical "Chicago". In all these forms, performance stands in distinction from mere "behavior" in implying some degree of intent (p391).*" Performance is about intentional behaviour, what is done or what is believed to be done. Values is about what people attach importance to, not about what they do.

Values are the frame of reference for the assessment of performance. Values and performance ask different questions about the operational dimensions of the general interest (integrity, robustness, economy, etc.). A performance assessment will analyse to what extent the general interest is safeguarded. A value assessment will ask questions about the which operational dimensions prevail, whether they are conflict, etc. In order to make this assessment of the dominant values in public sector, researchers may want to have a look at behaviour. They can study intentional behaviour that is aimed at the fulfilment of the general interest (performance). They could also look at non-intentional behaviour.

Another advantage is that the study of performance management and performance measurement becomes richer. Are these tools useful to achieve objectives that are embedded in different value systems? For instance, the development of indicators for integrity might be seen in this light.

## **Leadership and performance**

### ***Conventional mechanisms: Leadership → Performance***

In the theoretical development of leadership studies, this is the main proposition. The first studies attempted to identify the traits of successful leaders; originality, adaptability, self-confidence, integrity, conviction, etcetera. The next step was about the behaviours of leaders. The main thing to look at in behaviour is to what extent leaders show a concern for results on the one hand and for people on the other. The successful leader, obviously, looks at both. Blake & McCanse (1991) speak of team management; work accomplishment is from committed people with an interdependence through a common stake in the organizational purposes. There is trust and respect. Contingency theories were a third development. According to these theories, there needs to be a fit between leadership styles and organisational characteristics.

Many of these theories, however, seem to have an implicit assumption about good leadership. Other studies do not have such a predisposition. The question then is; does leadership matter for performance? Brewer and Selden (2000) confirmed in a study in 23 US federal agencies the importance of leadership on organizational performance. The leadership and supervision index they constructed was positively and significantly related to employee perceptions of organizational performance. Yet, they also found the weak predictive power of the index disappointing. Brewer & Selden (2000) provide three possible explanations for this finding: “(1) first, leadership and supervision may not be as important as the literature and anecdotal evidence suggests; (2) second, our measure of the concept may be incomplete and thus obscure its true importance; (3) third, leadership and supervision may contribute to organizational performance indirectly. For instance, skilled leaders may help create and sustain organizational cultures that promote high performance (p704)”.

### ***Counterintuitive mechanisms: Performance → Leadership***

The relation between leadership and performance may well be the other way around. High performing organisations may attract strong leadership. This may be the case for leadership at the strategic level at the top ranks in the organisation as well as for leadership on by the bottom ranks.

Good performance may attract good leadership at the top. Patrick Dunleavy's bureau shaping model may explain how this might work. Dunleavy (1991) argues that high-level bureaucrats seek to improve their social welfare rather than budget maximization. Bureaucrats more positively value staff functions over line functions, a collegial atmosphere over a corporate atmosphere, and a central location over a peripheral location. In all probability, top civil servants will also value organisations that are seen as good performers over low performers, since the success may reflect on the leader. Dunleavy (1991) notes that if (top) officials have these values, they can most effectively pursue these objectives at an individual level, searching for career or promotion paths which leads them to an appropriate rank in a suitable sort of agency (p.202).

Good performance of teams may also attract good leadership. Yet, the mechanism at play at the bottom of the organisation may be different. Good performance lead to self confidence and trust (a kind of a winning mood). This may in turn to an ownership of the purposes of the organization. Some people will also display behaviour typical for leaders; caring about co-workers and caring about results. The first person to look at would be the front line supervisors. However, we can also think about informal leadership of a 'primus inter pares'.

## ***Social constructivist mechanisms leadership as performance***

The social constructivist approach sees leadership and performance in a different way. Smircich & Morgan (1982) conceive leadership situations as those in which there exists an obligation or a perceived right on the part of certain individuals to define the reality of others. This has implications for the relationship between the concepts. Rather than one objective reality explaining the other, subjective phenomena are transformed through interpretative processes. There seem to be two mechanisms that relate leadership to performance, depending on who is doing the transformation; the leaders or the led. In social constructivist language; who is the agent?

In the first case, the led are the agents. The led see leaders as symbols. Pfeffer (1977) explains how this works using Kelley's (2006) attribution theory. Kelley concludes that persons are not only interested in understanding their world correctly, but also in controlling it. Controllable factors, it is argued, will have high salience as candidates for causal explanation. Controllable explanations (such as leadership) are not necessarily the most important explanations, but it is more comfortable to attribute success and failure to leadership than to intangible contextual causes. The identification of certain organizational roles as leadership position guides the construction of meaning.

Secondly, the leaders are the agents. Leaders are seen as translators that interpret the complex reality for the led (Czarniawska-Joerges & Sevón, 1996). Given the likelihood that success and failure will be attributed to leadership, it is of particular importance for leaders to confer the 'right' notion of performance. As we demonstrated above, performance is a multi-layered concept. This will leave substantial room for interpretation. Smircich & Morgan (1982) call this activity the management of meaning.

## **Reform and performance**

### ***Conventional mechanisms: Reform → Performance***

Better performance is the ultimate argument for reforms. In recent decades, performance, in accordance with the sigma values, has increasingly been defined as better efficiency and effectiveness. In these arguments, reform is the independent variable that should lead to performance. Most studies follow this line of reasoning.

An important theoretical formulation is contingency theory. The central proposition is summarised by Donaldson (2001) with the acronym SARFIT; structural adaptation to regain fit. Organisational structures have to be able to satisfy the requirements of the environment. This is the only way it can subtract resources from its environment (resource dependency). Another kind of fit is between the tasks (e.g. complexity) and the structure. It is fit that affects performance.

Recent efforts to study the determinants of performance are justified based on the reform agenda. It is argued that since better performance is the stated objective of reformers, we have to pursue a better insight in the determinants of performance in order to infuse evidence in the reform agenda (Boyne, 2003). There have been for instance vigorous debates on whether privatization, a popular reform strategy, has led to better performance. In general, the conclusion seems to be that it is not (public or private) ownership that determines performance. The level of competition is the main determinant (Savas, 2000).

### ***Counterintuitive mechanisms: Performance → Reform***

Reform may lead to performance. Yet, performance may also lead to reform. The line of reasoning introduces capital as an intermediate variable. The main thesis is that reform has its costs, and high performing organisations have more capital to bear the costs of reforms. A more negative formulation is to talk of organizational slack (March, 1981; Nohria & Gulati, 1996). In contrast to capital, the notion of slack suggests that the resources that are used for fostering change were previously unproductive. Capital can be interpreted in a strict financial sense, or in a broader sociological sense.

Although reformers most of the times promise a public sector ‘new style’ that is able to do more for less money, the reforms themselves usually cost money. Consultants are brought in, staff needs to be trained, infrastructure needs to be (re-)build, ICT systems are introduced, etcetera. In the private sector, restructuring costs are motivated by the promise for increase profits in the future. Profitable firms will have less problems, but also fewer incentives, to bear restructuring costs. Similar arguments could be made for the public sector. Wildavsky (1987) argues that before undertaking an evaluation, it is necessary for managers to ascertain who will pay the costs of change (see also (Boyne et al., 2004). If sufficient finances are not available, evaluators will either resist change, or they will do only what they consider to be “administratively feasible” and “organizationally attainable”.

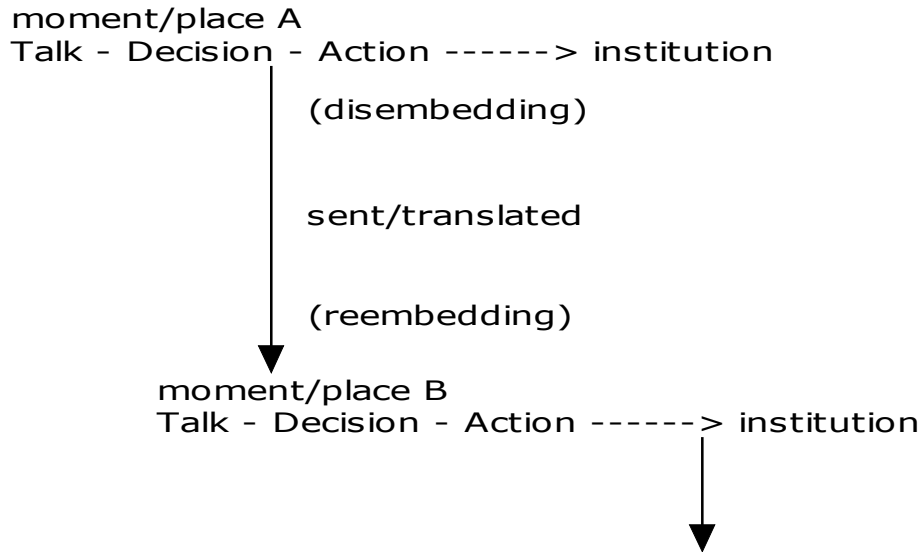
Capital can also be broadly interpreted in a sociological sense. Bourdieu (1986) theorized that people have three forms of capital; economic capital (command over resources), cultural capital (knowledge, skills, education), and social capital (group membership, networks, influence, support). Economic capital relates to capital in the strict sense (see above). Social and cultural capital too can be translated to an organisational level. High performing organisations are likely to possess more knowledge, skills, networks, relationships. These social and cultural assets in turn may be needed to bear the cost of reforms. Organisations that lack the necessary social and cultural capital may not be able to bear the costs of reform, and thus are less likely to successfully reform.

### ***Social constructivist mechanisms: Reform as Performance***

Some of the main constructivist texts on reform and performance have been written by the so-called Scandinavian institutionalists. A key feature of this stream of research is the disconnection words and deeds. (Brunsson, 2002). It is not uncommon for organisations to talk in one way, to decide in another and to act in a third. In this way, organisations can meet the often conflicting demands that are imposed on them. Inconsistency between ideas and action is easier and, therefore, more likely to happen than consistency.

An important point are the inconsistent of environments in which organizations operate. Different interests in the environment demand different things from organizations. This is the connection with performance. As we argued above, the definition of performance depends on the public values of the stakeholders that are demanding for performance. Organisations have to provide in different performances in order to respond to these often conflicting demands. One strategy is to talk about reform, in order to deliver the promised performance. The French saying ‘*plus ça change, plus c’est la même chose*’ (the more it changes, the more it remains the same), summarizes this strategy quite well. Another strategy is to actually reform, without talking about it. Another French saying goes ‘*pour vivre heureux, vivons cachés*’ (for living happy, we have to live concealed’).

But how do these external demands get to organisations? Czarniawska & Joerges (1996) write about ideas that travel. Figure 2 represents the process. At a given moment and place, there is talk, decision and action. Organisational talk and definitions may become disembedded from their initial context, travel and then become reembedded in a new context. Performance of organisations is defined through interactive processes of disembedding, translation and reembedding.



**Figure 2: travels of ideas (Czarniawska & Joerges, 1996)**

This framework is applied on New Public Management. The main argument is that to a certain extent, organisations talk about reform as a substitute for ‘real’ performance. That is, instead of delivering outputs and outcomes they speak about reforms as a promise of future performance. Christensen & Laegreid (2003) conclude from an analysis of the symbolic nature of administrative reforms that symbols matter. However, they also assert that reforms are not only talk and symbols. “Symbolic behaviour and processes of interpretation are important for the development of identities, loyalties, trust and compliance, but they can also be used to manipulate and exercise power. Turning symbols into practice is, however, not an easy task to do. There is a coupling between the world of reform policy symbols and the world of practice, even if it is not as tight as might be expected from an instrumental perspective (p.24)”.

Another study by Pollitt (2001) is about the process of disembedding and reembedding. The question is whether there is global convergence in the public management as a result of the internationally promoted NPM agenda. NPM is by many seen as a global paradigm. Institutions such as the OECD’s Public Management division have done their part in the disembedding and dissemination of national reform practices (mainly in New Zealand). But are these reforms retranslated to other national context? Pollitt argues that there are three levels of convergence “(1) Discursive convergence – more and more people are talking and writing about the same concepts (performance budgeting, TQM, etc.). The conceptual agenda is converging. (2) Decisional convergence – the authorities (governments, legislatures, boards, CEOs) publicly decide to adopt a particular organizational form or technique. (3) Practice convergence – public sector organizations begin to work in more similar ways. For example, many jurisdictions use competitive tendering for a larger and larger proportion of their budgeted expenditure. (4) Results convergence – this is when reforms produce their intended (and unintended effects) so that the outputs and outcomes of public sector activity begin to converge. For example the unit costs of issuing passports may fall in every country that re-engineers the issuing process in a certain way. (p.477-478)”. Convergence, he argues has mainly been discursive, somewhat decisional, rarely practical and almost never concerning the results. Reform and performance are disconnected. Reform is seen as (a substitute) for performance.

## **Performance measurement and performance**

Finally, we turn to the relation between measurement and performance. Often, there is no clear distinction between performance measurement and performance. This is because authors refer to the performance movement rather than to the practice of performance measurement and the concept of performance. Talbot’s (2005) discussion of

performance management for instance is mainly about performance measurement. In this text, measurement and performance are seen as two distinct phenomena.

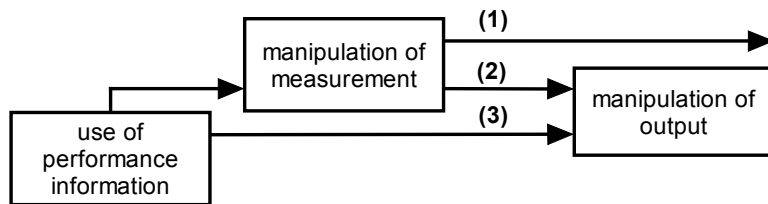
***Conventional mechanisms: Measurement → Performance***

The conventional relationship is about the performance of performance measurement. Research has usually been operationalised in terms of effects of measurement. In essence, two types of questions can be formulated; does performance measurement have positive effects (intended behavioural changes), and does performance measurement have negative effects (unintended behavioural changes). Surprisingly, most of the research on the effects of measurement has addressed the latter question: what are the negative consequences of performance measurement? The positive effects are until now more a matter of belief than of research evidence.

The social mechanism behind the negative effects of measurement is usually called gaming. Gaming is a well-documented reaction to measurement. In particular for what performance measurement in organizations concerns, a multitude of gaming strategies has been described. Generally, gaming strategies consist of either a manipulation of the measurement process or a manipulation of the organizational output (Figure 3) (van Dooren et al., 2006).

The first set of effects mainly affects the measurement process. Measurement is skewed due to these dysfunctions. Yet, the output of the organization is not necessarily affected. This is represented by (1) on Figure 3. For instance, in a case where measurement is pure window dressing, bad measurement may not impact the day-to-day operations of the organizations.

Secondly, some dysfunctions do alter the daily operations the organization. The practices of organizations are changed. Different quantities and/or qualities of output are pursued. These dysfunctions may materialize even with a perfect measurement system - (3) on Figure 3. However, with an imperfect measurement system, the chances of operational dysfunctions are higher - (2) on Figure 3. Flawed measurement can be a catalyst for distorted practices. Flawed measurement results can become a self-fulfilling prophecy.



**Figure 3: effects of the use of performance information.**

There is substantive evidence in the public administration literature of gaming actually taking place. An illustrative document in this respect is the report of the Public Administration Select Committee of the UK House of Commons (2003). The report provides substantial evidence of manipulation of performance indicators across sectors. For instance, hospitals had cancelled appointments and reduced follow up visits in order to meet waiting time targets. Another similar report is titled “performance indicators; good, bad and ugly” (Bird et al., 2005). Bevan & Hood (2006) provide evidence from the British health care sector. Heinrich (2002) points to some dysfunctional effects of measurement in a major US job training programme. Other texts are Bouckaert & Balk’s (1991) description of measurement pathologies and Smith’s description of the non-intended consequences of measurement in the national health service (Smith, 1995).

A lot of this research is (mostly unintended) deceptive because of conceptual imprecision. Performance measurement is seen as a container for all kinds of performance arrangements. In our view, a clear distinction between amongst others performance measurement, performance information and performance management is needed. Performance measurement is an activity that, if successful, produces performance information. The use of performance information is another activity. There are several uses of performance information. Performance

information can be used in order to support a style of management that is results oriented (performance management). Performance information can also be used for better accountability towards stakeholders. Another use is for learning and improvement (Van Dooren, 2006).

There is only an indirect relation between measurement and negative effects. The use of performance information is the intermediate variable that explains negative effects. Different uses exert different incentives for gaming. Carter et al.(1992) make a distinction between indicators as dials or indicators as tin-openers depending on how much room for interpretation the uses of performance information leave. Indicators that are used as tin openers only open the discussion. The main question is what is inside the box. Indicators as dials do not leave room interpretation. The measured number is tightly coupled to the final judgement.

The key question in assessing the performance of performance measurement thus is the utilization of performance information – both manifest and latent. A lot of evidence of the perverse effects of measurement stems from the England (not Wales or Scotland). A typical feature of English system is the high level of centralization<sup>iv</sup>. Under the banner of accountability, performance indicators systems seem to support this centralisation agenda, and thus the performance information is used in a more intimidating way. This leads to substantial frictions between higher levels in the hierarchy (“London”) and professionals in the field. In other countries, there seems to be less of a conflict between professions and measurement systems.

Other mechanisms where measurement is predicted (not) to affect performance, suffer from the same omitted variable bias. Meyer and Gupta’s (1994) performance paradox is an example. This theory states that indicator systems tend to lose their capacity to differentiate because of four processes. The first process is positive learning. As performance improves, indicators lose their capability of detecting bad performance. The second process is perverse learning. When organizations or individuals have learned which aspects of performance are measured (and which are not), they can use that information to manipulate their assessments. Earlier, we called this process gaming. The third process, selection, refers to the replacement of poor performers with better performers. Only good performers remain and the indicator loses its discriminating value. And fourth, suppression occurs when differences in performance are ignored. All these reactions to measurement will mostly occur when the indicators are used as dials, not as tin openers.

### ***Counterintuitive mechanisms: Performance → Measurement***

High performing organizations may measure more. The mechanism is somewhat similar the one explaining how performance may lead to reform. Measurement requires resources. High performing organisations may have more resources to sustain a measurement system. There seems to be little research on the measurement costs, although the substantial costs are regularly acknowledged (Kouzman et al., 1999). Some indications from adjoining fields may give an clue. The rule of the thumb for costing evaluation projects is 10% of the total budget of a programme.

### ***Social constructivist mechanisms: Measurement as performance***

A lot of evidence on how measurement constructs rather than depicts reality has emerged from sociological studies on social statistics. In an era where rationality is highly valued, measurement adds to the perceived importance of a phenomenon. The number defines the problem (De Neufville, 1975).

The social construction of reality by statistics becomes very clear in the process of census taking. Anderson & Fienberg (1999) studied the undercount in the American census. Blacks, urban poor and other minorities in particular were hard to count. It was argued that this undercount lead to underrepresentation in political institutions. More broadly, it also leads to underrepresentation in public life. The extent of particular problems will be underestimated and as a result, policy responses will also be inadequate. The social construction of reality is even more apparent in the decision on what to count. A revelatory issue in this respect is the racial registration (Arel & Kertzer, 2002). Registration is constituting a racial discourse. This racial discourse may shape and explain policy outcomes, and therefore is important. Others argue however that such a discourse leads to stigmatization of groups

in the population. This is the reason why in countries such as Belgium and France, ethnic registration is highly contested.

The question is at which point statistics are taken for granted as the source of evidence for policy making? De Neufville (1975) studied the history of some major social statistics, unemployment rate, Consumer Price Index (CPI), the standard budget and the crime rate, and found 10 factors that point to high institutionalisation. Porter (1995) further studies how we come to trust numbers, helped by technocrats, bureaucrats, professions and academic communities who infuse a sense of rationality in it. The edited volume 'demystifying social statistics' (Irvine et al., 1979) is another text in this field.

Why do we seem to be so susceptible to the truth expressed in numbers. One reason is the statistical illiteracy of citizens, politicians, bureaucrats and journalists (Murray et al., 2001) (Best, 2001). One of the first efforts was Darrell Huff's book on how to lie with statistics, published in (1954). These authors argue that we fail to ask some basic questions when we are confronted with statistics and they usually have ample examples. Since numbers explain and shape policies, it is argued, we have to use the best data available.

There is less evidence on the social constructive power of numbers in public administration and public management. Yet, two mechanisms can be distinguished; *measurement* as performance and *measures* as performance.

Measurement as performance. Modell (2004) finds a cycle of control myths in organizations. Myths come and go. He argues that goal-directed, multidimensional performance measurement models, such as the Balanced Scorecard, will gradually replace the old myth that public service provision may be improved by heavy reliance on financial control. Noordegraaf & Abma (2003) This research resonates the argument made by Feldman & March (1981) that the use of information symbols a commitment to rational choice. Displaying the symbol reaffirms the importance of this social value. The activity of measurement thus is a substitute for managerial performance ("we are performing, because we are measuring").

Measures as performance. This is a slightly different argument. Here we turn from the activity of measuring to the measures themselves. Managers (and others) have a bounded rationality and thus are seeking to make sense of their world and their job in a satisficing way (Vakkuri & Meklin, 2006). Measures are means to reduce the complex nature of performance. Performance measures are used to make sense of the many meanings of performance. Measurement then serves as a proxy for performance. Etzioni and Lehman (1967) call this phenomenon 'concept reduction' resulting in a lack of coincidence between a social concept and its operational definition. Obviously, this mechanism has its risks. Meyer (2002) refers to this state of mind as operationalism.

## conclusions

This texts studied the theoretical relation between performance, leadership, reform and measurement. Rather than attempting to look for an encompassing theory, we looked for social mechanisms explaining concrete relations (Hedstrom & Swedberg, 1998). This is in line with Merton's middle range approach to theory in the social sciences.

We used performance in this text as concept rather than a label for a movement. In our view, the literature on public values can inspire the performance literature. A broader notion of public performance may be developed.

Three kinds of social mechanisms that explain the relationship between the aforementioned phenomena (leadership/reform/measurement) and leadership. First, we studied mechanisms representing conventional wisdom (x determines performance). Secondly, we looked at what we called counterintuitive mechanisms (x is a consequence of performance). Thirdly, we searched for social constructivist approaches (x as performance). A summary is provided on the next page.

	X -> performance: conventional	Performance -> X; counterintuitive	X as performance; constructivist
Leadership	Leadership traits, behaviour, contingency determine performance "Management matters for performance" literature gives some evidence	Top of the organization; Bureau shaping, performers attract good leadership Bottom of the organisation; Performance gives confidence and trust to informal leadership	The led attribute success and failure preferably to controllable factors such as leadership. Leadership is management of meaning. Leaders give meaning to performance.
Reform	Contingency theory; structural adaptation of organizations to regain fit. Fit affects performance.	Reform requires resources, high performers are more likely to have these resources Capital can also be broadly interpreted as social and cultural capital	Some performance requirements are met with organisational talk, others with decisions or action. Performance requirements are defined through processes of disembedding, translation and reembedding.
Measurement	The use of information (not its mere measurement) has behavioural impacts	Measurement requires resources. High performers are more likely to have these resources	Measurement as performance. Measurement gives managers, with bounded rationality, a sense of control. Measures as performance. The measures define what is seen as performance – concept reduction.

## Reference List

- Anderson, M. J. & Fienberg, S. E. (1999). *Who counts? The politics of census taking in contemporary America*. New York: Russell Sage Foundation.
- Arel, D. & Kertzer, D. I. (2002). *Census and identity: The Politics of Race, Ethnicity, and Language in National Censuses*. Cambridge: Cambridge University Press.
- Best, J. (2001). *Damned Lies and Statistics: untangling numbers from the media, politicians, and activists*. University of California Press.
- Bevan, G. & Hood, C. (2006). What's measured is what matters: targets and gaming in the English public health care system. *Public Administration*, 84, 517-538.
- Bird, S. M., Sir David, C., Farewell, V. T., Harvey, G., Tim, H., & Peter, C. (2005). Performance indicators: good, bad, and ugly. *Journal of the Royal Statistical Society*, 168, 1-27.
- Blake, R. R. & McCause, A. A. (1991). *Leadership Dilemmas - Grid Solutions*. Houston: Gulf Pub.Co.
- Bouckaert, G. & Balk, W. (1991). Public Productivity Measurement: Diseases and Cures. *Public Productivity & Management Review*, 15, 229-235.
- Bouckaert, G. & Peters, B. G. (2002). Performance Measurement and Management. *Public Performance and Management Review*, 25.
- Bourdieu, P. (1986). The forms of capital. In J.G.Richardson (Ed.), (pp. 241-258). New York: Greenwood Press.
- Boyne, G. A., Gould-Williams, J. S., Law, J., & Walker, R. M. (2004). Toward the Self-Evaluating Organization? An Empirical Test of the Wildavsky Model. *Public Administration Review*, 64, 463-473.
- Boyne, G. A. (2003). Sources of Public Service Improvement: A Critical Review and Research Agenda. *Journal of Public Administration Research and Theory*, 13, 367-394.
- Brewer, G. A. & Selden, S. C. (2000). Why Elephants Gallop: Assessing and Predicting Organizational Performance in Federal Agencies. *Journal of Public Administration Research and Theory*, 10, 685-712.
- Brunsson, N. (2002). *The Organization of Hypocrisy: Talk, Decisions and Actions in Organizations*. Copenhagen: Abstrakt Liber.
- Carter, N., Klein, R., & Day, P. (1992). *How Organisations Measure Success: The Use of Performance Indicators in Government*. London: Routledge.
- Christensen, T. & Lægreid, P. (2003). Administrative Reform Policy: The Challenges of Turning Symbols into Practice. *Public Organization Review*, 3, 3-27.
- Cogburn, J. D. & Schneider, S. K. (2003). The Quality of Management and Government Performance: An Empirical Analysis of the American States. *Public Administration Review*, 63, 206-213.
- Cook, B. J. (1998). Politics, Political Leadership, and Public Management. *Public Administration Review*, 58, 225-230.
- Czarniawska, B. & Joerges, B. (1996). Travels of ideas. In B.Czarniawska & G. Sevón (Eds.), *Translating Organizational Change* (56 ed., pp. 13-48). Berlin: de Gruyter.

- Czarniawska-Joerges, B. & Sevón, G. (1996). *Translating Organizational Change*. Berlin: Walter de Gruyter.
- Davies, H. T. O., Nutley, S. M., & Smith, P. C. (2000). *What Works?: Evidence-based Policy and Practice in Public Services*. London: The Policy Press.
- De Bruijn, H. & Dicke, W. (2006). Strategies for safeguarding public values in liberalized utility sectors. *Public Administration*, 84, 717-735.
- De Neufville, J. I. (1975). *Social Indicators and Public Policy: Interactive Processes of Design and Application*. Elsevier.
- Deleeck, H., Huybrechts, J., & Cantillon, B. (1983). *Het Matteuseffect. De ongelijke verdeling van de sociale overheidsuitgaven in België*. Antwerp: Kluwer.
- DiMaggio, P. J. & Powell, W. W. (1983). The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. *American Sociological Review*, 48, 147-160.
- Donaldson, D. L. (2001). *The Contingency Theory of Organizations*. Thousand Oaks: Sage Publications.
- Dunleavy, P. (1986). Explaining the privatization boom: public choice versus radical approaches. *Public Administration*, 64, 13-34.
- Dunleavy, P. (1991). *Democracy, bureaucracy and public choice: economic explanations in political science*. Harvester Wheatsheaf.
- Etzioni, A. & Lehman, E. W. (1967). Some Dangers in "Valid" Social Measurement. *The ANNALS of the American Academy of Political and Social Science*, 373, 1.
- Feldman, M. S. & March, J. G. (1981). Information in Organizations as Signal and Symbol. *Administrative Science Quarterly*, 26, 171.
- Hedstrom, P. & Swedberg, R. (1998). *Social Mechanisms: An Analytical Approach to Social Theory*. Cambridge University Press.
- Heinrich, C. J. (2002). Outcomes-Based Performance Management in the Public Sector: Implications for Government Accountability and Effectiveness. *Public Administration Review*, 62, 712-725.
- Hood, C. (1991). A public management for all seasons. *Public Administration*, 69, 3-19.
- Huff, D. (1954). *How to lie with statistics*. New York: Norton.
- Irvine, J., Miles, I., & Evans, J. (1979). *Demystifying Social Statistics*. London: Pluto Press.
- Jorgensen, T. B. & Bozeman, B. (2002). Public Values Lost? Comparing cases on contracting out from Denmark and the United States. *Public Management Review*, 4, 63-81.
- Jorgensen, T. B. & Bozeman, B. (2007). Public Values: An Inventory. *Administration & Society*, 39, 354-381.
- Kaboolian, L. (1998). The New Public Management: Challenging the Boundaries of the Management vs. Administration Debate. *Public Administration Review*, 58, 189-193.
- Kelley, E. (3-10-2006). Measuring health care quality at the international level: Progress in the OECD Health Care Quality Indicators Project: paper presented at the Joint OECD/ONS/Government of Norway Workshop on the Measurement of non-market output in education and health.  
Ref Type: Slide
- Kouzmin, A., L'Affler, E., Klages, H., & Korac-Kakabadse, N. (1999). Benchmarking and performance measurement in public sectors. *International Journal of Public Sector Management*, 12, 121.

- Lynn, L. E. J. (2006). The Performance of Performance. *Journal of Public Administration Research and Theory*, 16, 633-636.
- March, J. G. (1981). Footnotes to Organizational Change. *Administrative Science Quarterly*, 26, 563-577.
- Merton, R. K. (1949). *Social Theory and Social Structure: Toward the Codification of Theory and Research*. New York: Free Press.
- Merton, R. K. (1968). On Sociological Theories of the Middle Range. *Social Theory and Social Structure*, 39-72.
- Meyer, M. W. (2002). *Rethinking Performance Measurement: beyond the balanced scorecard*. Cambridge: Cambridge University Press.
- Meyer, M. W. & Gupta, V. (1994). The performance paradox. *Research in Organizational Behavior*, 16, 309-369.
- Modell, S. (2004). Performance Measurement Myths in the Public Sector: A Research Note. *Financial Accountability & Management*, 20, 39-55.
- Moynihan, D. P. & Pandey, S. K. (2005). Testing How Management Matters in an Era of Government by Performance Management. *Journal of Public Administration Research and Theory*, 15, 421-439.
- Murray, D., Schwartz, J., Lichter, S. R., & Tokarski, C. (2001). *It Ain't Necessarily So: How Media Make and Unmake the Scientific Picture of Reality*. Hammondsworth: Penguin.
- Nicholson-Crotty, S. & O'Toole, L. J., Jr. (2004). Public Management and Organizational Performance: The Case of Law Enforcement Agencies. *Journal of Public Administration Research and Theory*, 14, 1-18.
- Nohria, N. & Gulati, R. (1996). Is slack good or bad for innovation? *Academy of Management Journal*, 39, 1245.
- Noordegraaf, M. & Abma, T. (2003). Management by Measurement? Public Management Practices Amidst Ambiguity. *Public Administration*, 81, 853-871.
- Pfeffer, J. (1977). The Ambiguity of Leadership. *The Academy of Management Review*, 2, 104-112.
- Pollitt, C. & Bouckaert, G. (2004). *Public Management Reform: A Comparative Analysis*. Oxford University Press.
- Pollitt, C. (2001). Clarifying convergence. Striking similarities and durable differences in public management reform. *Public Management Review*, 3, 471-492.
- Porter, T. M. (1995). *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton Univ Pr.
- Public administration select committee (2003). *On target? Government by measurement* London: House of Commons.
- Radin, B. A. (2006). *Challenging the Performance Movement: Accountability, Complexity, and Democratic Values*. Georgetown University Press.
- Savas, E. S. (2000). *Privatization and public-private partnerships*. New York: Chatham House.
- Scott, W. R. (2001). *Institutions and Organizations*. Thousand Oaks: Sage Publications.
- Smircich, L. & Morgan, G. (1982). Leadership: the management of meaning. *Journal of Applied Behavioral Science*, 18, 257-273.
- Smith, P. (1995). On the unintended consequences of publishing performance data in the public sector. *International Journal of Public Administration*, 18, 277-310.

- Talbot, C. (2005). Performance Management. In E.Ferlie, L. E. Lynn, & C. Pollitt (Eds.), *The Oxford handbook of public management* ( Oxford: Oxford University Press.
- Terry, L. D. (1998). Administrative Leadership, Neo-Managerialism, and the Public Management Movement. *Public Administration Review*, 58, 194-200.
- Vakkuri, J. & Meklin, P. (2006). Ambiguity in Performance Measurement: A Theoretical Approach to Organisational Uses of Performance Measurement. *Financial Accountability & Management*, 22, 235-250.
- Van Dooren, W. (2006). *Performance measurement in the Flemish public sector: a supply and demand approach*. Faculty of social Sciences.Leuven, Belgium.
- van Dooren, W., Manning, N., Malinska, J., Kraan, D. J., Sterck, M., & Bouckaert, G. (2006). *Issues in Output Measurement for Government at a Glance* Paris: OECD.
- Wildavsky, A. B. (1987). *Speaking Truth to Power: The Art and Craft of Policy Analysis*. New Brunswick: Transaction Publishers.
- Williams, D. W. (2000). Reinventing the Proverbs of Government. *Public Administration Review*, 60, 522-534.
- Wilson, J. Q. (1989). *Bureacracy: What Government Agencies Do and Why They Do It*. Washington: US, BasicBooks.

---

<sup>i</sup> The most prominent example of a grand theory is Parsons' functionalism. Nowadays, Luhman's social system theory is an important example of a grand theory (Brans and Rossbach 1997). Another example of a contemporary broad range theory is Gidden's structuration theory (Giddens 1986).

<sup>ii</sup> see Boyne (2003) for an overview of this literature up to 2003

<sup>iii</sup> The phrase from the Bible goes as follows; *For unto every one that hath shall be given, and he shall have abundance: but from him that hath not shall be taken away even that which he hath."*

<sup>iv</sup> Based on the discussions in the Netherlands Institute of Government - Public Services Programme Joint Workshop, Analysing Performance Indicators: Data, Behaviour, Impacts 15/16 March 2007.Erasmus Universiteit Rotterdam. <http://www.publicservices.ac.uk/NIG.asp>