

Performance measurement: the mismatch of manifest and latent functions

Paper presented at the conference of the Netherlands Institute of Government
29 October 2004, Rotterdam

Willem Trommel (University of Twente)
Taco Brandsen (Tilburg University)
Mirjan van Heffen-Oude Vrielink (University of Twente)
Maaïke Moulijn (University of Twente)

1 Introduction

Performance measurement is still on the rise. According to Power (1994, 1997), modern welfare states are developing into “audit societies”. In the same vein, Strathern (2000) argued that audit regimes have become the moral fabric of modern society. Both scholars feel that new public management (NPM) has furthered a culture of accountability by promoting “styles of governance which commonly presuppose the necessity and benefits of auditing in its various forms” (Power 1994: 143). Indeed, empirical evidence for the rise of audit systems, including all sorts of performance measurement tools, is overwhelming. Simultaneously, many scholars have shown how performance measurement can give rise to serious complications, in the shape of technical problems and all kinds of “diseases” and unintended consequences (cf. Bouckaert, 1995; Van Thiel & Leeuw, 2002). In this type of analysis, problems and solutions generally focus on improving our understanding of the information asymmetries involved in performance measurement. Such reasoning draws on institutional economics, claiming that the risks involved in principal-agent relationships (such as opportunistic behaviour and moral hazard) can be kept in check through rational strategies on the part of the principal.

We would argue, though, that a singular focus on information asymmetries may disregard the fundamental causes of failures in performance measurement. This is because it is based on the theoretical assumption of a singular, stable principal-agent relationship. However, many governance systems lack such a clear and stable relationship, and it is that which is at the heart of many problems in performance measurement. We will demonstrate this through an analysis of developments in four areas in the public domain (provisions for disability, labour reintegration, social housing and home care) in which performance measurement has been applied. These cases show that an analysis of performance measurement (and its problems) must take account of its historical development and position within the governance system: where it originates, what functions it has and how different functions are related. If this is omitted, one may be trying to solve the wrong type of problem. It is important to stress that such an analysis does not imply a rejection of research on information asymmetries; rather, it is an attempt to position this more clearly within the broader governance debate.

In paragraph two we will explain our approach to the analysis of performance measurement problems in more detail. Next, the four cases will be presented in paragraphs three through six, respectively provisions for the disabled, labour reintegration, social housing and home care. This will be followed by a comparative analysis and conclusion in paragraph seven.

2 Performance measurement and governance

In the public sector, the ideas from the NPM school put the issue of performance measurement on the agenda of institutional reform. Osborne and Gaebler, for instance, have argued that “measuring results” is a neglected field in the Weberian, input-based model of public administration. “If you don’t measure results, you can’t tell success from failure, and if you can’t reward success, you’re probably rewarding failure” (Osborne and Gaebler, 1992). There are various settings in which performance measurement has been applied. First, there is the typical situation in which (decentralised or “autonomised”) agencies perform public tasks, accounted for by means of “realised output”. Second, public tasks may be performed by a so-called “quasi-market”, in which private providers compete for clients or the custom of collective purchasers (Brandsen, 2004). Often, these quasi-markets lack adequate information on quality and prices, which means that performance measurement may be a necessary precondition for the well-functioning of the market (cf. LeGrand and Bartlett, 1993). Finally, public governance increasingly involves hybrid (public-private) organisational networks, in which public services are co-produced. Since such networks are not embedded in the traditional systems of political control, performance measurement tools become part of efforts to encourage different types of accountability.

Like anything in life, performance measurement does not always work as one would like, and many authors have addressed the problems. In such literature on the “diseases” and unintended consequences, the focus is largely on the behaviour of organisational actors and their reactions to incentives connected to the various measurement instruments. For instance, public agencies may stop investing in aspects of their work which are no longer rewarding in terms of new output indicators, or try to withhold information potentially damaging to their reputation. Information is the central concept here, implying that the key to controlling performance is to know the facts of it. When framed in terms of a principal-agent relationship, attention is often focused on information asymmetries: the agent often has more information about actual practice (the facts) than his principal, which allows the former to manipulate assessments of performance.

Although this type of explanation certainly covers some of the failures of performance measurement, its explanatory potential is intrinsically hampered by the way in which it conceives of the principal-agent relationship. To clarify this, it is important to distinguish between the conceptual and the normative use of the principal-agent notion. Conceptually, it is a perspective that presupposes a certain type of relationship between actors. Although limited by its bilateral focus, it is essentially neutral: theoretically, the roles could be attributed to any empirical actor, and any actor can be both principal and agent in certain respects. When used normatively, though, the notion implies that the relationship between two empirically identified actors should be one of principal and agent, in other words, of hierarchy. In many studies on performance measurement, the conceptual and the normative have become tangled up: the principal-agent

notion is used as a perspective to analyse failures, but with the implicit assumption that government must be the principal, who controls the other actors involved in public service provision. The nature of this relationship is not questioned, which, reflecting back onto the conceptual notion, means that roles are fixed and the relationship is stable. The consequence is that explanations for failure are only examined from within such a specifically defined relationship, and are not extended to the relationship itself.

Yet many governance systems lack the clarity that such an approach supposes. Whether desirable or not, many of the quasi-markets and networks mentioned above are characterised by multiple and ambiguous relationships, where it is not clear who is agent or principal, and where the nature of the relationship is contested. For example, service providers may feel that government's attempts at control overstep the bounds of legitimacy; in other words, that it is less of a principal than it tries to be. The analysis of performance measurement tools cannot be isolated from such a context. When actors have no consensus on the nature of their mutual relationship, they may have different views of what performance measurement means. To clarify this, we will identify four potential functions.

Given that the principal-agent notion implies hierarchy, ambiguity at least partially concerns uncertainty about the level of coercion that public authorities can legitimately exercise. On a scale of rising coercion, De Bruin (2002) distinguishes four functions of performance measurement. Respectively, performance measurement can contribute to:

- (a) Transparency: other actors are informed about the nature of performance;
- (b) Organisational learning: transparency creates opportunities for a debate on quality, in terms of procedures, products and services;
- (c) Assessment: measurement allows evidence-based judgement on the quality of performance;
- (d) Sanctioning: assessment is tied to binding decisions with respect to budgets, continuation of contracts and so on.

Some functions can only be fulfilled under specific institutional (governmental) conditions. For instance, the function of sanctioning presupposes a multi-layered governance system rooted in an undisputed, legitimate distribution of authority. However, the nature of governance and the (use of) performance measurement tools are not necessarily well-aligned. To demonstrate this, we have selected four empirical cases along the following two dimensions:

- Is the governance system in question based on clear principles, or characterised by ambiguity?
- Have performance measurement tools been deliberately designed as part of the governance system, or have they developed in a spontaneous process?

This has resulted in four cases, which can be characterised as follows:

1. *Well-designed governance + public control over performance measurement*

Case: Decentralisation of service provision for disabled people.

The analysis will show what happens when the actors involved challenge the legitimacy of the underlying governance system.

2. Ambiguous governance +/- public control over performance measurement

Case: Private provision of labour reintegration tasks.

This case demonstrates how the need for performance measurement tools increases in a context of opaque and hybrid governance, whereas this context itself impedes effective implementation of such tools.

3. Ambiguous governance + absence of performance measurement regulation

Case: Quasi-market in social housing.

The case shows that this has caused a situation in which several actors have developed performance measurement tools of their own, which stand in an unclear, overlapping and even conflicting relationship.

4. Well-designed governance + absence of performance measurement regulation

Case: Introduction of a quasi-market for home-care delivery.

The case shows how a lack of transparency regarding public funding provokes a need for performance measurement, which turn from an instrument of accountability into a tool for inciting organisational learning.

In the following paragraphs, the four cases will be analysed at greater length. In each case, the analysis will focus on the governance structure, the nature of performance measurement and its problems, the (intended and actual) functions of performance measurement, and finally, the relationship between governance and performance measurement. It will become clear that the actual functions of performance measurement often differ from its manifest ones, which means that an analysis of its problems must also necessarily shift.

3 Provisions for disabled persons: the renegotiation of governance

Introduction

In the Netherlands, the provision of facilities, such as wheelchairs, for disabled persons is considered to be a public service. In order to keep costs down and enable adaptation to local circumstances, central government decentralized the responsibility for provision in 1994. Since then the provision is arranged in a Framework Act, in which municipalities are held responsible for policy making concerning disabled persons in their community and the distribution of facilities (Provision for Disabled Persons Act¹, article 2).

Along with decentralization and deregulation, municipalities also assumed financial responsibility, meaning that the Dutch Ministry of Health, Welfare and Sports allocates an annual budget to municipalities. The Ministry is not concerned as to how a municipality spends the budget as long as it administers the facilities on a minimal level, as defined by the Act. Municipalities are permitted to use any budget surplus for other purposes; however, other municipal resources bear the costs in the case of a deficit. In this way, municipalities have an incentive to work efficiently.

The Ministry monitors the administration by performance measurement tools. It will be shown that these tools suit the governance system. Therefore at first glance performance measurement seems to work well. However, some problems regarding the performance measurements arise. Although these problems seem to be classic NPM problems, evidence indicates that discontent over (the evolution of) the governance system could be the major cause. In this way performance measurement fulfils another, unintended function, namely as a means to renegotiate the governance system.

Well-considered governance system

The Framework Act came into effect after an intensive debate about the governance system. The parties involved agreed upon a well-considered system, in which the responsibility for policy formulation and implementation lies within a municipality. It is assumed that the democratically chosen city council and an advisory board of clients monitor these tasks, while the mayor, aldermen and bureaucrats are held responsible for the execution. In this way the Ministry aims at horizontal accounting within a municipality instead of hierarchical or vertical accounting between the Ministry and a municipality (RMO, 2002: 34). The only external control is through Courts, which are supposed to ensure a minimal level of service provision in individual cases, based on the Framework Act (House of Parliaments, 1993: memorandum of understanding).

The responsibility for the Ministry is limited to the maintenance of the system. Its main activities are to ensure a sufficient macro-budget² and to formulate fair distributive criteria. In order to fulfil these tasks, the Ministry requires information about the services delivered and costs involved. For example, the Ministry regularly investigates if the macro-budget is adequate, for this purpose it needs information about municipal costs. In the case of an inadequate macro-budget, it examines the causes of the deficit. If causes can be found within municipal policies and implementations, municipalities are held responsible. However, in case external causes can be identified, like the increase of prices for facilities or a growth in the number of recipients, the ministry is expected to increase the macro-budget, as it did in 2000 (Ipsa Facto and SGB0, 2001: 22). Thus, the main contribution of performance measurement, in this case, is transparency.

Public control over performance measurement

Although municipalities have a large degree of autonomy, they are legally obliged to hand over information required by the Ministry (Provision for Disabled Person Act, article 10). The Ministry has hired consultants to collect the required information and interpret data of municipal performances. The main method of data-collection is an annual scan, which is a questionnaire regarding financial and other quantitative indicators³ that municipalities have to complete. These data and interpretations are presented in evaluations of the Act (in 1995, 1997 and 2001) and in annual overviews of quantitative data. In this manner the Ministry has control over the system of performance measurement.

Performance measurement in practice

Superficially seen, performance measurement in this governance system is a success as it results in transparency. After all, the gathered information allows the Ministry to perform its tasks. However, if one looks beneath the surface it seems to be a can of worms.

The first problem is that organizations representing the interests of disabled persons do not share the positive official conclusions about performance. According to the evaluations, the performance of municipalities is quite good: they deliver facilities for the disabled at a reasonable standard and for a reasonable price. On the other hand, organizations of disabled claim many municipalities mess up because they do not provide the services that disabled persons need and instead provide cheaper ones in order to keep costs down (Moerkamp, 2000). To strengthen their argument, organizations of disabled conduct their own investigations and present their own interpretations of the performance measurements results, which are not as positive as the official ones (e.g. CG-raad, 2001). The differences in results led to question marks about the impartiality of the official conclusions, especially considering one of the hired consultants is in fact an associate of the Dutch Society of Municipalities⁴.

These frictions could be the result of typical NPM performance measurement problems, for instance the difficulty of measuring performances because it is hard to present indicators that give a realistic idea of municipal service provision. However, some evidence shows that disagreement about the legitimacy of the governance system could be the real underlying cause of this problem. It appears that municipalities prefer the designed system, in which they have a lot of autonomy. Based on the official results, they argue that the Ministry does not have any reason to interfere with municipal autonomy or change the governance system. Therefore, the official results are of benefit to municipalities. Organizations of disabled plead for a governance system in which municipalities are bound to national regulation and the Ministry has an extended role. According to them, horizontal accounting is undesirable or at least not enough accounting, because it leads to differences between municipalities in legal equality and security. So their not so positive results are in favour of their argument. Disputes about the performances seem to veil the real problem of disagreement about the governance system.

The second problem is the rising non-cooperative attitude of municipalities toward data-collection, which makes performance measurement impossible. Although the principal-agent relation can also explain non-cooperative attitude of agents, in this case it seems to arise out of discontent with the way the governance system develops itself.

It appears that many representatives of the Lower House support pleas of disabled in favour of ministerial interference (e.g. Smits, 2000). Therefore they changed the governance system de facto, since they established a comprehensive national protocol in 2002, which reduced municipal autonomy considerably. According to many municipalities this change is unacceptable because the official results showed they did well. They feel that the data they provided to the Ministry are not being used in decisions of the Lower House; consequently some of them are unwilling to contribute to the data-collection any longer⁵. Moreover most municipalities are very unsatisfied with this protocol, since they are unsure about how they can be financial responsible when they are only seen as executors of policy. Disputes about cooperating or not with performance measurements seem to veil the real problem of disagreement about the evolution of the governance system.

Politicising of performance measurement

Failures in the performance measurement can be understood as classical principal-agent problems. However, in this case discontent over (the evolution of) the governance system seems to be the major reason for problems with performance measurement. Actors in the field seem to

use the language of performance measurement and performance measurement problems to express their (implicit) objections against the governance system, and to renegotiate the institutional responsibilities in that system. Thus, performance measurement is more or less politicised and fulfils in this way another, unintended function.

4 Labour reintegration: the legitimising aspects of performance measurement

Introduction

The analysis in this section is based on research into the structure and conduct of a market-based system of labour (re)integration policy in the Netherlands(cf. Svensson & Trommel 2003). Until recently, a large share of governmental activities in this area was carried out at the municipal level by social service organisations and labour exchange agencies. However, in the second half of the nineties the Dutch government initiated a major institutional revision in the fields of both social security and labour integration, aiming at a more efficient and effective implementation of policy. An important element of this operation was the decision to outsource most labour reintegration activities.

The analysis focuses on one of the larger cities in the east of the Netherlands (which will here be referred to as “Easttown”). Easttown was among the first municipalities introduce the new approach to labour reintegration. This happened at the beginning of 2001. We studied the design of the new system and its functioning during the first one and a half years.

At the municipal level labour reintegration mainly concerns the group of citizens receiving a social assistance benefit. Within this population there are a wide variety of social problems, ranging from lack of educational training to social isolation, physical handicaps or severe drugs problems. This implies that reintegration plans must be tailored to individual needs. Usually, such a plan is based on a first diagnosis, followed by different activities in the area of social empowerment, educational training, job searching and the like. Under the new system, the chain of necessary activities, -from diagnosis to actual re-entry on the labour market- is referred to as a “reintegration trajectory”. At the beginning of 2001 Easttown hired five private reintegration agencies to handle an agreed-upon number of individual “trajectories”.

Our study shows it was extremely difficult to measure the performance of these agencies. Beneath we discuss the nature of these problems. It will be shown that specific problems in the area of performance measurement are strongly related to aspects of the design of the new system.

Performance measurement: major problems

At the beginning of 2001 about 1.200 trajectories were outsourced. To be more precise: the social policy department of the municipality selected five commercial agencies, which were then licensed to handle a fixed number of trajectories (a so-called “package”). However, the agencies did not receive their clients all at once. During the following year, officials from the social services department sent clients to the agencies on an individual basis.

So how did the private reintegration agencies perform? As indicated above, it is very hard to pronounce sound judgement on this. In our study we found large variance between the agencies

with regard to the most important output variable, the share of reintegrated clients. However, these findings are highly unreliable, given the following considerations:

- I. *Selection of clients.* We found that municipal consultants (case managers) developed specific preferences with respect to their use of the agencies. This resulted in an (implicit) division of labour between the agencies: some dealt especially with the complex cases –i.e. clients with the largest distance from the labour market- while others were allowed to concentrate on the easiest cases. Therefore, it would be neither fair nor valid to judge the agencies merely by output standards.
- II. *Interdependencies.* For a reintegration agency to be successful, it is very important to have good access to municipal resources (client information, use of legal sanctions, reintegration instruments etc.). As far as this was concerned, it appeared that some agencies had a much better relationship with local officials (c.q. case managers) than their competitors, and therefore much better access to resources.
- III. *Uncertainty about the workload.* While one of the agencies received only 14% of the “package” that was agreed upon, another agency received 37%, and yet another one as much as 89%. Evidently, the resulting uncertainty prevents agencies from good organisational and personnel planning. The more the actual workload deviates from the expected one, the more one can expect providers to suffer serious financial problems.
- IV. *Lack of reliable data.* We found perplexing discrepancies between the data (on clients, products, results) delivered by the agencies and those of the municipal administration. Although contracts with the agencies included clear arrangements regarding data delivery, in practise the municipal department did not enforce the implementation of these arrangements or set up an effective administrative system.

In short, the municipal authorities were unable to assess the performance of the agencies under contract. This was partly a data-related problem; however, even if adequate data on progress and results had been available, it would still have been impossible to make a fair comparison between the efforts of the reintegration firms, for the reasons mentioned above. One would need very precise knowledge on the selection of clients and the organisation of the work processes within the reintegration agencies in order to cast on actual achievements. The difficulty in measuring performance very much relates to the design of the new system, as we shall discuss now.

Managed competition

Clearly, the new approach to labour integration in Easttown is inspired by a market-oriented concept of governance. This new approach aims to introduce “managed competition” or, as Le Grand and Bartlett (1993) put it, a quasi-market for the provision of public goods. The basic concept is that public authorities withdraw from the delivery of goods and services and leave those tasks, under clearly specified conditions and regulations, to commercial organisations. It is important to note that managed competition does not necessarily call for a performance measurement system. Theoretically, it is assumed that markets will produce adequate information on quality and prices, and that clients will punish poor service delivery by choosing an alternative provider.

Still, this is not the way it works in Easttown. There seems to be a strong need for performance measurement which can be explained from two characteristics of the new institutional design: a) the use of so-called “block contracts” and b) the importance of the “package-principle”.

Block contracts

Labour (re)integration is a highly complex product, consisting of a large variety of services, such as educational training, psychological counselling and drug rehabilitation programs. “Deconstructing” the reintegration trajectory into separately priced and contracted services may be a very complicated and costly. Given these high transaction costs, it is understandable that Easttown prefers the use of block contracts, which enable the purchase of complete reintegration trajectories, leaving a large measure of discretion (with respect to the construction of the trajectory) to private providers.

However, such an approach negatively affects both competition and transparency. In fact, reintegration agencies become the *preferred* providers for a whole range of (interlinked) activities, which makes it hard to tell if prices reflect actual market conditions. In turn, this will increase the need for an ex-post performance assessment.

The “package-principle”

In a typical market context, clients must be able to select the provider of their choice. The less this basic principle is realised, the less providers will be inclined to increase quality and reduce costs. Seen from this perspective, Easttown agreements on the number of reintegration cases per agency (the “package”) seem to make no sense, as they contradict the principle of free choice.

Still, there are some arguments to justify this specific design. The precondition for individual choice would be that clients are awarded a personal budget, enabling them to act as consumers on the reintegration market. Here some serious risks are involved. First, clients might be opportunistic in their purchasing behaviour, for instance, if their (silent) ambition is not re-entry on the job market but continuation of social assistance benefits. In that case, clients will tend to select providers with a poor reputation. Second, providers might try to seduce clients by offering services with little relevance to the public aim of labour reintegration. For these reasons, it is no surprise that a system based on personal reintegration funds is still controversial within the political arena.

However, the result is that local authorities must accept a method of implementation that does not produce the disciplining effect one would expect in a typical market-context. Given the time-consuming character of reintegration services, there is a real danger that poor performance will only manifest itself after a considerable period of time. Therefore, the need for interim performance assessment increases dramatically.

Managed co-operation

While the need for a monitoring system increases, local authorities also become increasingly aware of the difficulties involved in realising such a system. As we know, performance measurement is complex enough for agencies operating at arms’ length from public authorities. But reintegration concerns private firms with even more autonomy both in financial and legal terms. Why would they make their organisations fully transparent? How reliable are their reports

on progress and results? Here, well-known problems with respect to principal-agent relationships are likely to be ubiquitous and highly persistent.

In our study, we found that local officials started searching for means of reducing the risk of opportunistic behaviour on the part of private reintegration agencies. They did this by initiating forms of co-operation. This practice first developed at the stage of client-selection. It appeared that close relationships developed between municipal case managers on the one hand and (some) representatives of reintegration firms on the other hand. Of course, this co-operation may encourage the effective assignment of clients, given that case managers acquire knowledge about the strengths and weaknesses of the contracted agencies. However, it may also result in an uneven (and/or unfair) distribution of cases. We found evidence for the latter, which meant that we could not trust our data on realised output.

Furthermore, several forms of co-operation developed during the implementation of reintegration trajectories. During this process, some firms appeared to have much better access to municipal resources than others, based on higher levels of trust in dealing with local case managers. Again, it is possible that the development of networked co-operation may well contribute to effective reintegration, but nonetheless it does not match the logic of fair competition and valid performance measurement.

Conclusion

The design of the market process has major shortcomings. It fails to copy the selective mechanisms of the market with respect to quality and price. Potential effects cannot even be observed, since poor performance will not show in the conduct of clients. Given the risks associated with long-term block contracts and the use of packages, it is no surprise that local case managers start investing in co-operative and trust-based relations with private reintegration agencies. However, without being too cynical, one could wonder about the surplus value of this new practice over the previous, bureaucratic system of policy implementation. In some respects, reintegration agencies are increasingly treated as (over-paid?) divisions of the municipal organisation...

All in all, one can observe a hybrid situation that gives rise to a strong need for sound performance measurement, while simultaneously obstructing the realisation of such a system. Above all, the function of performance measurement in this case is that it covers-up failures in the governance system while feeding the impression that nonetheless an effective and efficient mode of governance has been installed.

5 Social housing: hidden confusion over governance

Introduction

In the Netherlands, the social housing stock has traditionally been in the care of private nonprofit providers (for an overview, see Brandsen, 2004; Brandsen and Helderma, 2004). In the postwar period, these former charities became closely linked to state housing policy through a complex system of subsidies and regulation. They built extensively and became the dominant players in the rental market, at present still controlling well over a third of the total housing stock. Many are

still fairly small organisations, with a few dozen employees and an average of about three thousand homes in their possession. Recently, though, mergers have led to the rise of a number of large, nationally operating providers.

In the early 1990s, the financial autonomy of the social housing providers was restored, with the state buying off future subsidies in exchange for scrapping all outstanding loans. At the same time, regulatory constraints on their room for manoeuvre were relaxed. Nonetheless, the providers are still expected to fulfil their traditional function of affording housing for low-income families and people with special needs. In exchange for this, they retain a number of privileges such as tax exemptions (though these have been partially lifted). Since part of their capital has been accumulated through past subsidies, and as they have the responsibility to carry out public duties, the providers are subject to government supervision.

The shape of performance measurement

Within the governance system of Dutch social housing, formal supervision rests with the Ministry of Housing, but the system of performance measurement operates on two tiers. At the national level, the government has laid down a number of broad standards within the *Besluit Beheer Sociale Huursector*, the act outlining the basic rules of the game to which social housing providers have to submit. These standards amount to the following:

- I. Maintaining the quality of the social housing stock;
- II. Preferential treatment for low-income households;
- III. Encouraging combinations of housing and health care for the elderly;
- IV. Maintaining the social and physical quality of the housing environment;
- V. Giving voice to tenants in organisational decision-making.

It is evident that these are rather broad standards. In practice, there are many ways in which they can be interpreted. In fact, it is even desirable that there should be a wide variety of interpretations. The housing market is divided into various local and regional submarkets, each different in terms of the demographic and socio-economic composition of the population, urban density, the size and composition of existing housing stock, the scope for expansion, and the portfolio and financial soundness of the individual housing providers operating within the submarket. To define detailed performance criteria for all providers concerned would be meaningless; to define them separately for each provider or group of providers would require a huge amount of information about local conditions. It is for this reason that the specification of targets has been decentralised to the local level. Local authorities, or groups of local authorities within a region, agree on contracts with social housing providers (usually collectively, with specifications for each organisation). Sometimes, other parties such as tenant organisations are included in the negotiations. The Ministry of Housing strongly encourages detailed contracts, though neither the content nor the level of detail are officially specified. It is hoped that local knowledge and priorities will lead to the development of generally accepted and tailor-made performance criteria.

In addition to this formal system of performance measurement, there are a number of self-regulatory methods through which social housing providers measure their own performance, notably benchmarking and peer review procedures. While these partly focus on other criteria, e.g.

relating to internal organisational matters, there is an overlap with the official system of performance measurement. The development of these instruments is still at an early stage, though, and participation in their use has not (yet) been adopted as membership policy by the national association of social housing providers Aedes.

Performance measurement in practice

There are two basic problems associated with this elaborate system of performance measurement. The first concerns the specification of performance, the second the relations between the various instruments involved.

To begin with, there is no general consensus over how performance should be measured, not just in terms of indicators, but also in terms of values and scope. What with the expansion of non-landlord activities, housing providers are increasingly defining their task more widely than simply providing a roof. For example, they also invest in activities relating to community building, lifestyles and empowerment. Since they have a great deal of money at their disposal, their willingness to invest in other types of activities has been greeted heartily by other local organisations. However, the diversity and fuzziness of the scope of their tasks complicate performance measurement greatly. At present, many areas of activity covered by performance contracts are left vague and unspecified. In such general terms, of course, it is easy enough to agree on the criteria that social housing providers should meet. They should invest in good housing, help disadvantaged groups, invest in neighbourhoods, develop housing for the elderly, and so forth. The difficulty is in the combination of the criteria, in determining how much effort in one type of activity can be offset against another, and the development of specified indicators.

A further problem is that it is not clear how exactly the different instruments of performance measurement are related. There are two basic areas of tension. The first concerns the role of formal supervision in relation to local performance contracts. When a contract is set up between municipalities and housing providers, it is formally a matter of private law. In case of transgressions, there is the formal possibility of recourse to the legal system (though in practice this never occurs). Alternatively, transgressions can be reported to the Inspection, which can in turn ask the Minister of Housing to impose sanctions. This sometimes does happen, but only rarely. There is a theoretical clash of sanctioning instruments here, which is never realised. There is a real problem here, though. The performance contracts are not subject to prior approval by the Inspection, as they are of a private nature and involve municipal authorities over which the Inspection has no jurisdiction. This can put the latter in a difficult position. When contracts are left vague and unspecified, without detailed indicators, the Inspection is left with little to supervise, nor can it formally press for elaboration of the contracts. In practice, it mostly lobbies behind the scenes.

Another area of tension involves legal (vertical) rules and self-imposed (horizontal) regulation. Although there have been some moves by the Ministry in a more hierarchical direction, there remains an institutional gap concerning specifications of performance. Hierarchical regulation is mostly “negative”, in that it only prevents certain activities from being undertaken. Positive indicators that allow organisational learning and competition between providers are mostly developed by the organisations themselves and by their branch association, but the status of such instruments is hanging in the air. At present, they are not tightly coupled to any system of

rewards or sanctions, which means their function is largely internal. The matter is complicated further by the existence of competing self-regulation instruments, since providers do not yet agree among themselves which are the proper instruments.

The root of the problems is disagreement over prevalent co-ordination mechanisms, and ultimately, what is the nature of the housing providers. In this quasi-market type of governance system, the market is obviously a key method for bringing supply and demand together. In that sense, providers are businesses, a point they themselves stress to underpin their claim to self-regulation. Their historical status as private non-profits provides further fuel for this claim and protects them from the direct interventions that public bodies could be subject to. However, given that the housing market functions imperfectly, that part of the capital held by providers is of public origin, and that certain public criteria have to be met, there is also a strong case for strong and continued public regulation and performance measurement. It is unclear to what extent benchmarks, peer reviews and contracts should be used for the purpose of learning, of assessment, or sanctioning.

Conclusion

The problems in defining specific indicators appear to have less to do with the nature of the products and services than with the diverse claims and balanced powers of the parties involved. Discussions over performance measurement instruments are vehicles for the more fundamental discussion over responsibilities and levels of autonomy within the present governance system. Another way in which this translates itself is in assessments of the levels of investment providers are financially capable of while safeguarding their long-term continuity. This is a situation in which failure to develop suitable indicators regarding public tasks cannot be ascribed (only) to information problems, but must also be attributed to a lack of agreement on the functions of performance measurement, and fear on how other actors might use the information it produces.

6 Benchmarking home care: between transparency and secrecy

Introduction

In the last decade the Dutch health care system has been reformed drastically. In the early nineties a quasi-market replaced the traditional system in which suppliers received public or insurance money for a previously negotiated supply of (medical) care. The introduction of market activity in the health care sector was based on the assumption that regulated competition among health care providers and insurers would create a system of financial incentives. The need to compete with other providers and insurers would cause them to have an interest in better efficiency. In the home care sector the introduction of market activity has been an incremental process.

Home care is regarded to be a basic health provision in the Dutch health care system. It consists of home help, community nursing and maternity care. About 70 million hours of community nursing and home help is provided to about two million clients yearly.⁶ The Dutch government considers the provision of home care to be insurable by private insurers or Medicaid. For that reason community nursing and home help are included in the General Act on Exceptional Medical Expenses (AWBZ). The home care entitlement traditionally was executed by an

authorised (recognised) home care organisation. In 1994 these organisations for the first time had to compete with private home care providers. To some extent the “regular” providers responded to this incentive by increasing their efficiency. However, it also caused a wave of amalgamations and alliances that largely offset competition. The introduction of new competitors therefore did not create the expected financial incentives (cf. Breedveld, 2003).

The entrance of private providers on the home care market also had some other unintended side effects. To begin with, the private providers only offered their services to lucrative clients, burdening the regular providers with the less profitable clients. Because of this “unfair” competition several regular home care organisations encountered financial problems. Secondly, the policy reform encouraged home care providers to pursue a market-oriented strategy. But a number of them seemed to reduce costs at the expense of the quality of care. Thirdly, some of the regular home care organisations started a private service branch, which among other things diminished the transparency of the use of public funding. The Dutch government responded to these unintended side effects by bringing the reform to a temporary halt. However, in recent years competition is brought back to the scene. In 1998 the ministry of Health, Welfare and Sports (VWS) decided to transform the supply-side funded health care system into a demand-side funded system. It is expected this will empower health care consumers, which in turn will induce suppliers of (medical) care to better efficiency and higher quality of care. In the *Beleidsagenda 2004* (Ministerie van Volksgezondheid, Welzijn en Sport, 2003: 17) this line of policy is refined and broken down into various actions that should improve the performances of insurers and care suppliers, e.g. more competition among insurers, and performance measurement by means of benchmarking.

The shape of performance measurement: benchmarking

Benchmarking is considered a suitable instrument to enable citizens, governmental agencies or other stakeholders to form an opinion about whether they receive “value for money”. The introduction of competition in the home care sector raised a need for transparency regarding the efficient and proper use of public money. Almost all home care organisations receiving public funding were affiliated to the National Association for Community Nursing and Home Help (LVT). This branch organisation arranged home care benchmarking, which is financed by the ministry of VWS.¹ The branch organisation and the ministry of VWS together called in a large firm of consultants to perform the benchmarking. The results regarding the efficiency and quality of home care provision have been submitted to them on an aggregated level. In addition to that the consultants informed the home care organisations that were subject of the benchmark on their individual level of performance. Although the Dutch government was financially involved, it did not enforce performance measurement regulation. This allowed the LVT to act at its own discretion. As a consequence the focus of benchmarking gradually shifted from an accounting instrument into a tool to orchestrate organisational learning.

After a pilot in 1997 the first integral benchmark arranged by the LVT took place in 1999. Considering the doubts regarding the spending of public funding and the financial problems of some of the regular home care organisations, it will hardly surprise that the pilot entirely focused on the expenditures and production of home care providers. The benchmark functioned as an

¹ *Landelijke Vereniging Thuiszorg: baanbreker in prestatie-indicatoren*, <http://zorgaanzet.nl/opinie>, 28-01-2003.

instrument to account for the use of public money. However, the renewed competition in the home care sector soon induced the LVT to broaden the scope of the next benchmark. Instead of a single focus on efficiency - in terms of putting public money to good use – the benchmark was used to compare individual performances. It was hoped for that this would stimulate home care organisations to improve their efficiency and quality of care. Accordingly, the first integral benchmark not only concentrated on expenditure and production but also looked into client and professional satisfaction.

In 2002 the second integral benchmark took place. Compared to the first integral benchmark its successor gave a more refined representation of the performances of home care organisations. Moreover, the LVT hoped to gain insight into best practices by linking the performances to the way in which the organisations provide care, but it turned out that there was no significant correlation. The LVT now tries to find out what goals members that perform best set and how they achieve these goals, which is reflected in the main function of the benchmark. On its website the LVT defines the benchmark as: “a continuous systematic process for measuring costs and quality of products and services, and selecting best practices for the purpose of organizational improvement”.⁷ By relating outcomes to processes the LVT wants to pinpoint “performance indicators”. In this way it hopes to offer its members not only insight in their individual level of performance but also an understanding of the way in which they can improve themselves. Those LVT-members that achieve the best performances based on a combined evaluation covering aspects of efficiency and quality will be identified as “best practices”. These best practices serve as a tool for identifying points in respect of which LVT-members can improve their performances. In the end the LVT strives at an integration of performance indicators. This shows that the main function of the benchmark has changed over time: it turned from a tool to improve transparency into an instrument to orchestrate organisational learning.

Consequences of the change toward organisational learning

How does the change in the design of the benchmark affect the transparency that is required on a successful quasi-market? The main concern of branch organisations is to lobby and to protect their member’s interests. Accordingly the branch organisation and its members use the results of the benchmarks to consolidate their position in a quasi-market. We therefore may expect that the changed function of the benchmark - from a tool to improve transparency into an instrument to orchestrate organisational learning - will affect the chances of stakeholders to form an opinion about whether they receive “value for money”. The benchmark will be used as an accounting instrument only when the results prove that their members stand out in some specific part of service provision, for example the high rating of LVT-members on client satisfaction. Moreover, when future benchmarks reveal how best practices are linked to the process of care provision a further closure of information is likely. As the ministry of VWS will not be financially involved in the next benchmark, which is scheduled for 2005, the results of the benchmark will come exclusively into the possession of the branch organisation and its members. Their emphasis on the learning from best practises will reduce their willingness to share the results with “outsiders”. After all, the more transparency the easier best practices can be copied by competitors and the more vulnerable home care organisations are to governmental cut backs.

Main conclusion

When a benchmark model is designed to orchestrate organisational learning, the willingness to share the results with stakeholders diminishes. They will be informed only when this is profitable from the point of view of those who commissioned the benchmark. Complete openness about the results may conflict with a market strategy that aims at learning from best practises in order to distinguish oneself from competitors. We conclude from this that benchmarking is less appropriate to increase transparency when it is used primarily as an instrument to facilitate improvement of performances and less as an accounting instrument.

7 Analysis and conclusions

The four cases show that the official functions of performance measurement tools are frequently not realised. Though in two cases (disability + home care) a certain level of transparency was achieved, none of the cases showed clear examples of assessment and/or sanctioning, although these were the formally intended functions of the performance measurement tools. Yet to refer to these developments as “unintended consequences” would be confusing, because they are not unintended, but different than originally and/or formally intended.

A more suitable terminology comes from a classic article by Robert Merton, in which he argued that functional analysis must keep its distance from the subjective intentions of actors in social systems: functional analysis is about “objective consequences” (Merton, 1949). However, he also recognised that many scholars in the field were unable to explain the functioning of social systems without reference to the motivations of the participants in the system. Therefore, Merton suggested distinguishing between manifest and latent functions, the first referring to the obvious and intended functions of elements in a system, the latter to functions that are not overtly intended or recognised. A governance system, including its performance measurement instruments, has intended (manifest) functions, recognised by its designers; in its actual functioning it may provoke unintended consequences (for instance due to information asymmetries) *and* it may develop new (latent) functions that were not thought of beforehand.

As the cases show, some of these latent functions are especially functional to the goals of some particular participants in the system, rather than to the over-all system’s purposes, and must in some cases be considered dysfunctional. We distinguish four such functions.

Politicisation

Once performance measurement has been established, as in the disability case, actors may start challenging various aspects of the system, such as indicators. Although this might look like a technical debate on the feasibility of the methods involved, it may actually be the result of fundamental objections against the governance system and its distribution of institutional responsibilities. In an audit culture, performance measurement discourse may have become the natural way to express particular political and administrative interests. Or, in the vocabulary of Michel Foucault, questions of “governmentality” are increasingly raised in terms such as “visible results” and “opportunities for inspection”.

Regaining public control

Reforms inspired by NPM usually involve a withdrawal of direct bureaucratic control over the implementation of public policy, as was the case in Dutch social housing. Yet, if the results of reform are disappointing in this respect, the point of no return may have been passed, in the sense that the formal institutional structure cannot be restored without great difficulty. Performance management tools may then be used in an effort to re-introduce public aims on the agenda of relevant actors in the field, for example, through “contractual relationships” between public authorities and autonomous agencies, as an. This might be labelled a case of “principals in search for agents”.

Masking failures

Institutional reforms do not necessarily result in transparent, unambiguous governance systems. As the case of municipal labour reintegration showed, system designers may combine elements of quasi-markets, outsourcing, bureaucratic control and interorganisational co-operation within a single structure, without much reflection on its coherence. The ambiguity with respect to institutional responsibilities and achievements may lead to urgent calls for (better) performance measurement tools. At the same time, however, the ambiguity of the governance system will not allow for an effective use of these tools. In cases like this, such tools are adopted from the institutional environment, as a “rationalised myth” that contributes to the suggestion of effective governance (cf. Meyer & Rowan 1977).

Disruption of institutional logic

The case of home care provision showed that private providers in a quasi-market may have incentives to hide relevant information, to prefer secrecy over transparency, and to defy or bypass institutional pressures towards “objective information services” (cf. Oliver 1991). This is typical of “regular” markets, but it means that performance measurement, when left in the hands of the providers themselves, becomes a learning tool rather an instrument of accountability. The kind of transparency required by the public aspect of the system clashes with the need for intransparency on the part of providers.

In these cases the failure to realise the manifest functions of performance measurement is not the result of technical problems, but of the ambiguity, imperfection and hybridity of the newly developed governance systems. Performance measurement tools and discourse seem to have developed into the vehicles by which these systems are changed and challenged, irrespective of evidence regarding the actual success or failure of the tools. No doubt, the birth of the “audit society” is closely related to the emergence of NPM, but its diffusion and continuation is also propelled by the tendency of performance measurement tools to take on latent functions.

Of course, our effort has only signalled this aspect of performance measurement. Further comparative research in other sectors and countries would provide a far richer picture of the latent functions of performance measurement allowing a more systematic typology of these functions in relation to particular features of governance systems. In this context, it would be useful to shift more attention to the organisational level and to concentrate on the different dynamics behind the emergence of functions and dysfunctions. It also seems desirable to improve our understanding of the ways in which principal-agent and institutional arguments relate to each other. The former tends to focus on manifest functions and, as a result, emphasizes the risks caused by information asymmetries. The latter concentrates on latent functions following from

institutional dynamics. One may have to reflect not only on current applications of the principal-agent concept, but also on its applicability in governance systems characterised by increasing multiplicity and ambiguity.

Bibliography

- Bouckaert, G. (1995) "Improving performance measurement". In G. B. A. Halachmi, (ed.), *The enduring challenges in public management: surviving and excelling in a changing world*, San Francisco: Jossey-Bass, pp. 379-412.
- Brandsen, T. (2004), *Quasi-market governance: an anatomy of innovation*, Utrecht: Lemma.
- Brandsen, T. and J.-K. Helderma (2004) *Sociaal Kapitaal en Sturend Vermogen*, The Hague: Wetenschappelijke Raad voor het Regeringsbeleid (National Scientific Council).
- Bruijn, H. de (2002), *Managing Performance in the Public Sector*, London: Routledge.
- CG-raad (2001), *Wet zonder recht; bespreking vanuit gebruikersperspectief van de derde evaluatie Wvg*, Utrecht: CG-raad.
- House of Parliament (1993), *Vergaderjaar 92-93: 22815, nr 2*, Staatsblad, p. 545.
- Ipsos Facto and SGBO (2001), *Een verstrekkende wet, derde evaluatie Wet voorzieningen gehandicapten*, Den Haag: Sdu uitgevers.
- Landelijke Vereniging voor thuiszorg (1997), *Home care in the Netherlands*, http://www.thuiszorg.nl/docs/home_care_in_the_netherlands.pdf
- Landelijke Vereniging voor thuiszorg (2003), *Performance Monitoring, Improvement and Benchmarking in Home Care*, outline of a paper presentation in Orlando.
- Le Grand, J. and W. Bartlett (eds) (1993), *Quasi-markets and social policy*, London: Macmillan Press Ltd.
- Merton, R.K. (1949), "Manifest and Latent Functions". In: Merton, R., *Social Theory and Social Structure*, Glencoe, Ill, The Free Press.
- Meyer, J.W., B. Rowan (1977), "Institutionalised Organisations: formal structure as myth and ceremony". In: *American Journal of Sociology*, no. 83, pp. 340-363.
- Ministerie van Volksgezondheid, Welzijn en Sport (2003), *Beleidsagenda 2004. Keuzes voor een houdbare en beter presterende zorg*, Den Haag.
- Moerkamp, J. (2000)., "Uitvoering Wvg: gemeenten hebben alle krediet verspeeld". In: *Binnenlands Bestuur*, 25-08-2000, p. 28.
- Moulijn, M. (2003), "Decentralisatie: een geluk of een ongeluk?" In: *Bestuurskunde*, jrg. 12, nr. 7, pp. 286-295, Den Haag: Reed Business Information.
- Oliver, C. (1991), "Strategic Responses to Institutional Processes". In: *Academy of Management Review* 16, pp. 145-179.
- Osborne, D., T. Gaebler (1992), *Reinventing Government*. New York: Plume.
- Power M., (1997), *The Audit Society: Rituals of Verification*. Oxford: Oxford University Press.
- Power, M. (1994), *The Audit Explosion*. London: Demos.
- RMO (2002), *Bevrijdende kaders*, Den Haag: Sdu uitgevers.
- SGBO (2003), *Kerncijfers Wet voorzieningen gehandicapten 2002*, Den Haag: Ministerie van Volksgezondheid, Welzijn en Sport.
- Smits, J. (2000), "Wvg". In: *Binnenlands Bestuur*, 21-04-2000.
- Strathern, M. (ed.) (2000), *Audit Cultures*. London, Routledge.
- Svensson, J., W. Trommel (2003), *Arbeidsreïntegratie in de markt*. In: *Over.Werk 4*, pp. 107-111.
- Thiel, S. van, and F.L. Leeuw (2002) "The performance paradox in the public sector" In: *Public Performance & Management Review* (25), pp. 267-81.

¹ In Dutch: Wet Voorzieningen Gehandicapten (Wvg).

² The macro-budget can be defined as the national budget that is available for service provision to disabled, before it is allocated to municipalities.

³ This scan includes questions about the number of applications, number of allocated facilities, percentage applications that is permitted, costs of facilities and costs of implementation (SGBO, 2003).

⁴ In Dutch: Vereniging Nederlandse Gemeenten (VNG).

⁵ For example, in 2002 20% of the municipalities did not participate in the scan research (SGBO, 2003: 2).

⁶ Community nursing and home help is a collective term for a number of divergent functions and services that are provided in the home situation.

⁷ http://www.thuiszorg.nl/docs/final_orlando.pdf