

Pollitt, C. & Dan, S. (forthcoming). Searching for impacts in performance-oriented management reforms: A review of the European literature. *Public Performance and Management Review*.

SEARCHING FOR IMPACTS IN PERFORMANCE-ORIENTED  
MANAGEMENT REFORM  
A REVIEW OF THE EUROPEAN LITERATURE

CHRISTOPHER POLLITT

Emeritus Professor, Public Management Institute, KU Leuven

Parkstraat 45, Box 3609, Leuven, 3000, Belgium

E-mail: [christopher.pollitt@soc.kuleuven.be](mailto:christopher.pollitt@soc.kuleuven.be)

SORIN DAN

Researcher, Public Management Institute, KU Leuven

Parkstraat 45, Box 3609, Leuven, 3000, Belgium

E-mail: [sorin.dan@soc.kuleuven.be](mailto:sorin.dan@soc.kuleuven.be)

Tel: 0032 16 32 36 33

**ABSTRACT:** This paper reports an analysis of 519 studies describing the impacts of NPM-type reforms across Europe. The aim was to establish whether performance-oriented reforms had led to changes in outputs and/or outcomes. This was thus a test of a central proposition in NPM theory. A number of significant conceptual and methodological problems were immediately apparent, and we describe our route through them. In conclusion, one principal finding is that although the population of studies is large, the number of high quality studies that focus on outputs is quite modest and the set addressing outcomes is very small. Another finding is that the identified impacts are distinctly ‘mixed’, with substantial proportions of studies indicating that specified outputs or outcomes are unchanged or ‘down’. Significantly, the database examines contextual influences which facilitate (or hinder) NPM reforms. Despite enormous attention, our understanding of the impacts of NPM remains both fragmentary and fragile.

**KEYWORDS:** New Public Management (NPM), evaluation, impacts, Europe

The New Public Management (NPM) has spawned a huge and international academic literature. Many commentators have attempted summaries or overviews (for example Barzelay, 2000; Christensen & Lægreid, 2011; Hood & Peters, 2004; Lane, 2000; Pollitt, 2003) and it might therefore be thought that there was little left to say on the subject. For at least three reasons, however, we argue that there is room for more. First, we focus here on

what differences have performance-oriented management reforms made to outputs and outcomes – whereas many of the overviews and other studies are (perfectly legitimately) focused on defining NPM and discussing how it has spread and evolved on largely internal organizational changes of process and structure. Yet most definitions of NPM make it clear that performance improvements are at the heart of this kind of reform (Aristigueta & Van Dooren, 2007). Therefore we ask what is known of shifts in the relevant performances. Second, we take Europe as our domain. On the one hand this is a very wide canvas, but on the other it has until recently represented something of a *terra incognita* in NPM studies. The main weight of scholarship has usually fallen on those states considered to be the originators and most vigorous proponents of the brand – the Anglophone states of Australia, New Zealand, the UK and the US (see for example Lynn, 2006; Pollitt & Bouckaert, 2011). Here we survey the whole of Europe, west, central and eastern. Third, as described later, we have assembled a large-scale, multi-lingual database and subjected each individual study to considerable analysis as to scope, methods and findings.

The layout of the paper is straightforward. We first set out what we mean by 'NPM' and by 'impacts'. Second, we take the reader through the processes by which we have selected and categorized the studies in our database. Third, we go on to review the substance of our findings.

### **Theoretical background**

Our main purpose is to seek to answer an apparently straightforward empirical question: 'what have been the impacts of NPM reforms across Europe?' To address the question we have assembled a large database of studies and we look across this database to try to see the 'big picture'. The studies in the database adopt a wide range of theoretical perspectives (and some of the official reports try as hard as they can to avoid 'theory' altogether). We have

refrained from trying to impose some master theory or meta-theory on top of all this variety – we simply seek to report it. Yet we cannot, of course, avoid theory that easily.

The ‘practitioner theory’ behind NPM may not be either entirely homogenous or entirely coherent, but it does posit performance improvement as a central result of NPM-type reforms. There are several strands to this. First, efficiency will result from the additional authority given to professional managers. Second, it will also be enhanced by the use of market-type mechanisms such as contracting out, performance league tables, and performance-related pay. Third, in so far as service users can be given choice, they will exercise this in favor of those providers who offer a better service. Thus, with NPM, services will become both more efficient and more responsive to service users.

To assess the degree to which these claims have been born out in practice we must look at changes not merely in processes and structures (important though these may be) but also in *outputs* and *outcomes*. Efficiency is most commonly defined as the ratio between inputs and outputs, and therefore its assessment must involve some measurement of outputs (mere input reductions score as increased economy, but not necessarily increased efficiency). Services which are more responsive to users and give them greater satisfaction are achieving higher *outcomes*. These definitional issues can be both complex and (at times) fraught, but we believe we are here applying widely accepted versions of the relevant concepts. Therefore we are using the database to test an important part of a major body of practitioner theory.

In assembling the database we sought to identify any significant contextual factors that have influenced outputs and/or outcomes. This reflects our broad sympathy for a ‘critical realist’ approach:

‘Context, we have gradually learnt, is not unwelcome noise, not a confounding variable to be controlled for – but an integral part of a programme’ (Pawson, 2013, p. 36)

To put it simply, we assume it is the *combination* of the *intervention* (NPM policy or technique) and *context* which give rise to specific impacts. So a given technique may produce quite different outputs and outcomes in different circumstances. Contexts consist of complex assemblies of different elements, including organizational cultures, structures and procedures, legal frameworks and the political climate (Pawson, 2013).

### **Defining NPM**

Defining NPM is the first step but it is not easy. In its origins NPM is strongly associated with UK Prime Minister Thatcher and US President Ronald Reagan, and with the New Zealand Labor governments of 1984-1990. Neither Mrs. Thatcher nor Ronald Reagan were any friends of the ‘planning’ approach which had been the orthodoxy in the US and UK public sectors of the 1960s and early 1970s. During their periods in power in the 1980s they, and many of their advisers, favored what they considered to be a more ‘business-like’ approach. Gradually, partly through doctrine and partly through trial and error, this general attitude crystallized into a more specific set of recipes for public sector reform. By the early 1990s a number of influential commentators appeared to believe that there was one clear direction – at least in the Anglophone world. This general direction was soon labeled as the New Public Management (NPM) or (in the US) Re-inventing Government (a seminal article here was Hood, 1991). A pair of American management consultants, who wrote a best-seller entitled *Reinventing government* and then became advisers to the US Vice President on a major reform program, was convinced that the changes they saw were part of a global trend. They claimed that ‘entrepreneurial government’ (as they called it) was both worldwide and ‘inevitable’ (Osborne & Gaebler, 1992).

There have been many definitional disputes and ambiguities about exactly what the key elements of this widespread trend were supposed to be (Dunleavy & colleagues, 2006; Gow & Dufour, 2000; Hood, 1990; Hood & Peters, 2004). For our purpose we will assume

(like Dunleavy and colleagues) that the NPM is a two-level phenomenon. At the higher level it is a general theory or doctrine that the public sector can be improved by the importation of business concepts, techniques and values (Gore, 1997, Lord President of the Council, 1981). Such a large-scale importation, it should be noted, is not merely a technical operation, because it ultimately has implications for the role and size of the state itself. Then, second, at the more mundane level, NPM is a bundle of specific concepts and practices, particularly including:

- Greater emphasis on ‘performance’, especially through the measurement of outputs. Here we should immediately note that, while performance-oriented reforms are central to NPM, they are not exclusive to NPM (Pollitt & Bouckaert, 2011). In this paper, however, we focus only on reforms that are seen by commentators in the countries concerned as being distinctly NPM-ish in character
- A preference for lean, flat, small, specialized (disaggregated) organizational forms over large, multi-functional forms (for instance semi-autonomous executive agencies – see for example Pollitt & colleagues, 2004)
- A widespread substitution of contracts for hierarchical relations as the principal coordinating device
- A widespread injection of market-type mechanisms (MTMs) including competitive tendering, public sector league tables, performance-related pay and various user-choice mechanisms
- An emphasis on treating service users as ‘customers’ and on the application of generic quality improvement techniques such as TQM.

Dunleavy & colleagues (2006, p. 470) have usefully summarized this as ‘disaggregation + competition + incentivization’.

It would be mistaken to assume that this formula was internally consistent (Hood, 1995). There is some tension between different intellectual streams that feed into the NPM, particularly between the economic, principal-and-agent way of thinking, which is essentially low trust, and the more generic managerial way of thinking which is more concerned with leadership and innovation – and more trusting of the inherent creativity of staff, if only they are properly led and motivated (Pollitt, 2003). The former stream emphasizes the construction of rational systems of incentives and penalties to ‘make the managers manage’. The latter emphasizes the need to ‘let the managers manage’ by facilitating creative leadership, entrepreneurship and cultural change. Other writers have drawn a parallel distinction between ‘hard’ and ‘soft’ versions of NPM (Ferlie & Geraghty, 2005).

### **Defining and assessing impacts**

We have distinguished between impacts on *activities/processes*, *outputs* and *outcomes*. Within these broad categories we have paid particular attention to impacts on efficiency, effectiveness, and the attitudes of those who use public services. ‘Impacts’ are things that can reasonably be supposed to have been *caused* by the reforms. In thinking about impacts we adopted a fairly standard model of policymaking (*Figure 1*).

### ***Place Figure 1 here***

An organization or program is conceived as a set of *activities or processes*. These would include organizational arrangements like the division of responsibility, the allocation of authority, the standard operating procedures, and so on. These procedures enable the organization to deploy and redeploy its resources which are collectively termed *inputs*. From

these activities and processes the organization or program then produces a set of *outputs*. These outputs are, in a sense, what the organization ‘gives’ to the outside world – to citizens, to civil society associations and to business firms. They are like messages, passing across the membrane that separates (on the one hand) the state from (on the other) the market sector and civil society. Outputs are invariably intended to produce desirable *outcomes*. An outcome is something that happens in the world outside the organization and the program: it is an effect ‘out there in the real world’. Some analysts make distinctions between ‘intermediate outcomes’ and ‘final outcomes’, but we decided that such a distinction, though useful in some contexts, would be of limited value in our field of study, where outcomes measures of any kind turn out to be rather rare.

There is also the vexed issue of *attribution* which needs to be discussed upfront. To be an impact the particular outcome or output must appear, *prima facie*, to be the *result* of that reform, not of other developments which may be happening at the same time. There must be a plausible causality. This condition is, however, often hard to satisfy (in the studies we have examined frequently it is not satisfied at all). A classic example would be shifts in public satisfaction, or trust, in government or in the civil service. If there is a shift (either upwards or downwards) it *could* be the result of a preceding public service reform. However, it *could* also be the result of a general well-being factor, linked to an economic boom, or to ephemeral shifts in party politics (a new leader, a scandal) or to pre-existing personal expectations or hopes, or any combination of these.

All in all, it can be readily appreciated that assessing the impacts of NPM reforms is far from straightforward. It involves several layers of categorization and interpretation, each of which has some effect on what is eventually ‘found’. More generally, it is important to understand that the body of work contained within the database derives from a mixed, increasingly multinational community, consisting of academics, public servants, management



consultants and politicians. Their reasons for becoming engaged with NPM reforms have differed. This mixture of motives means that the dividing line between descriptive and analytical ('is') statements and normative ('ought') statements is frequently hard to find. The desire to understand and explain is often tangled up with the desire to promote and support a particular kind of reform. Our raw materials (studies) are neither pure nor homogenous. Despite the large N this points towards an interpretive literature review with some simple counting. More advanced statistical techniques would not be appropriate for such heterogeneity.

### **The domain covered by the database<sup>i</sup>**

For practical reasons of research capacity we decided not to try to pursue every NPM reform in every European country at every level. We limited our search in various ways. We excluded acts of outright privatization (privatization is here defined as the sale of publically-owned assets to the private sector).

We concentrated on central governments. We have not attempted to cover reforms of sub-national tiers of government *per se*, although we *have* included quite a few centrally-inspired initiatives which have impacted significantly on sub-national and local tiers. Some of these have been on a large scale (for example Audit Commission, 2009). We covered the literature from all EU member states (except Bulgaria, Cyprus, Malta and Slovenia) plus Croatia, Norway and Switzerland. In all we took in studies from 26 states and the European Commission.

### **The selection of studies and construction of the database**

The process of selecting studies for inclusion was a complex one, and provided us with a forcible introduction to the problems of defining and identifying reform impacts. In order to get the pan-European coverage and linguistic diversity we were seeking we worked with

partners in ten other universities in nine other countries apart from our own. Together with these partners we went through six main steps:

- We discussed, in meetings and by correspondence, what we meant by NPM (arriving at the position outlined above)
- We discussed, in meetings and by correspondence, what we meant by impacts (arriving at the position outlined above). Neither of these first two steps was entirely straightforward, as they revealed quite a wide diversity of initial definition and usage.
- We then asked all partners to identify key texts (whether academic or practitioner) which assessed NPM impacts, both from their own country and from neighboring countries
- Once the texts were identified we asked partners to submit them to an electronic database. At this point they were also required to fill in a form (for each document) in which they were asked to classify the document and make various assessments of it. For example, did the document say outcomes or outputs or processes had improved, stayed the same or deteriorated? Did it mention significant contextual influences and, if so, what were they? We termed the answers to these questions ‘metadata’ and they were lodged in the database, linked to the actual document or study
- In addition to the documents identified and submitted by our partner universities we conducted a literature review of our own, and subsequently added further documents which we found, together with their associated metadata. This literature review is described in the Appendix
- As a further check for consistency the two authors of this paper read through a large sample (more than 150) of the database entries to check whether we

agreed with the classifications used by our partners. We also checked a smaller number (circa 30) of studies coded by each other, to assure ourselves that our own assessments were identical. In these two ways we attempted to ensure that inter-coder reliability was adequate

- We did *not* insist on one particular method or set of methods in the studies. To have done so would have radically reduced the total number of studies, and therefore undermined our wish to establish what the big picture of current (published) knowledge is.

In sum, we have taken a relatively relaxed approach to the process of admitting documents to the database, but a puritanical approach to analyzing what it finally contains (especially in the classification of outputs and outcomes). Our puritanism, if that is what it is, stems from a desire to achieve a measure of consistency across a large number of very diverse documents, together with our determination to focus, as far as we are able, on the outputs and outcomes that matter to the world *outside* public sector organizations. The question, to put it crudely, is ‘what do we *get* for all this reorganization?’ where ‘we’ are the citizenry and its representatives, civil society associations and companies. In all the multifarious literature on NPM this question has not been asked often enough, and has rarely been asked on the empirical scale that is performed here.

Most of the studies in the database (68 percent) are academic while the rest include official policy reports, external and internal evaluations or studies by international organizations. Large-N studies make up circa 22 percent of the total database. Small-N makes up circa 36 percent while the rest are analytical overviews and declarative studies without a precise identification of the number of cases. The definition of small and large N that we adopted did not follow a specific threshold, but rather it reflected the distinction between

qualitative case study approaches (small N) versus quantitative research (large N). While the generous treatment of evidence of impacts can be considered a limitation, it was in line with our goal of seeking the overall picture of the NPM literature across Europe. A few other scholars have focused on studies that employ particular quantitative methods, but inevitably they have ended up with much smaller populations of studies (for example Andrews, 2011) which are not representative of the overall character of the NPM literature.

It is plain that the UK (nearly 20 percent), France (13.8 percent), Germany (12.4 percent), the Netherlands (also 12.4 percent) and Italy (11.7 percent) heavily outnumber the numbers of studies from other countries (see Table 1 in the Appendix). Four of these ‘big hitters’ are also countries with the largest populations, but the Netherlands seems to punch above its demographic weight, possibly because it has a proportionately large population of public administration scholars and a tradition of academic involvement in public policymaking.

## **Findings**

One first finding is that most of the studies have not gone beyond reporting changes (or no change) in processes or activities. Studies of outputs are less common than studies of processes. As for outcomes, we found that only 45 studies in the total database (less than 9 percent) have gone that far (see Table 2 in the Appendix). Table 3 in the Appendix synthesizes our findings concerning the ‘success’ or ‘failure’ of such reforms. Outcomes and outputs, not just processes and activities, are found to have improved more often than not. One should note, however, that this broadly favorable picture is hardly uniform. Table 3 indicates that many studies report negative changes in outcomes, outputs or processes. Furthermore, a sizable number of studies and entries have not reported significant changes, or have reported no changes at all.

NPM-type reforms do not seem to be always successful. On the upside, outcomes, outputs and processes/activities are all reported as improving things more often than they make them worse. On the other hand, one could argue that the glass is only half full, because, if one adds together the categories of ‘no change’ and ‘things got worse’, one finds that in 56 percent of the studies where outcomes were reported, they did *not* get better, and in 47 percent of the studies where outputs were recorded they also failed to improve. This makes NPM reform look like a pretty hit-and-miss affair.

We should immediately note some limitations on the interpretation of these findings. First, we have bundled bigger and smaller changes together. Some studies cover single cases while others cover whole programs of reform. Second, improvements in one aspect may be balanced out by deteriorations in another. Thus, to say that the majority of those studies which focus on outputs and outcomes record improvements does not mean that there may not also be deteriorations in other areas recorded in some of those studies. Management reforms present complex trade-offs which (willingly or unwillingly) are often not acknowledged and assessed. The following examples of output and/or outcome studies from the database illustrate this. Andersen (2008) studied the impact of performance-oriented reform on performance, quality and inequality in Danish public schools. The results suggested that performance, measured by average exam scores, did not change significantly, quality improved slightly while the measure of inequality worsened. Jeannot & Guillemot (2010) conducted a quantitative analysis of performance management tools in French central government and deconcentrated services. They found no impact on outcomes, firm improvements in internal processes and lower staff morale. Clear improvements in processes came at the expense of lower staff satisfaction.

Many of our studies refer to important contextual factors which, they argue, facilitated or (more usually) inhibited or distorted performance-oriented reform attempts. Analysis of

the database suggests that these contextual factors may provide one of the main explanations for why a given reform may be more successful in one country or sector than in another.

To analyze contextual factors we developed an analytical framework. We proceeded inductively by following the following steps:

1. We read systematically through all the studies in the database which identified specific changes in outputs or outcomes
2. We noted wherever one of these studies mentioned a contextual factor as having been an important influence, and listed what those factors were
3. We inspected the list and built from it a set of categories that are appropriate to our specific topic – the impacts of performance reforms.

Three important features emerged. First, one can conceptualize a range from deep-seated, long-lasting influences at one pole towards medium term and then short term influences at the other. On the whole reformers should be able to foresee and anticipate many of the longer lasting influences (for example those introducing reforms in the CEE countries in the 1990s should have recognized that they were starting with a civil service that was weak on many skills, often corrupt and not at all ‘consumer focused’ – and they should therefore have included plans for dealing with these features). But the closer one comes to the other pole – sudden events and chance happenings – the more difficult it will usually be for reformers to plan or allow for them (for example reforms may be blown off course by the discovery of a major scandal, or the occurrence of a major accident or natural disaster just after the reform has been announced). There is a complication here which needs to be noted. Time runs in both directions. Some sudden, surprise events can have long term consequences in the future (9/11 being an outstanding example). Other contextual factors may have existed for a long time but have only a minor influence in the future. On the whole, however, features of political systems and public organizations which have existed for a long time – cultures, basic

structures, embedded standing operating procedures – are likely to be resistant to being changed overnight. In short, it is highly probable that long-standing, basic features of the machine will require sustained efforts over months or years before they can be removed or radically modified.

Second, different contextual effects are not equally broad in their effects. Thus the civil service culture may be expected to exert an influence right across government. Equally the effects of a centralized, powerful executive are likely to be widely spread. However not all contextual influences are on this scale. Some may be local, yet nevertheless durable and strong. For example, it may be that a particular town or a particularly prestigious organization has a different culture from the others and insists on ‘doing things differently’ (for example the case of the capital, Tallinn, in the primary health care reforms in Estonia – see Atun & colleagues, 2005). Or there may be a local incident or controversy, such as that concerning the shooting incident which led to the premature departure of a reforming English police chief, as cited in Pollitt & Bouckaert (2009).

The third important feature is that contextual influences can help or hinder reforms. Indeed, the politico-administrative culture, for example, may hinder one type of reform but help another. A political system with a powerful centralized executive may help that country quickly to implement a radical reform that improves, say, tax collection services, but on another occasion the same structural features may encourage governments to over-reform – to go on and on changing things just because they are easy to change, until confusion and cynicism begin to grow. Thus, although contextual factors most frequently gain attention when they hinder reforms, they can also facilitate implementation. For example, the socio-economic status of the local catchment area tends strongly to affect the exam results achieved by a school (Wilson & Piebalga, 2008). What kind of local contextual influence has, say, a

deprived socio-economic catchment on a national reform designed to improve exam performance? On the basis of our analytical framework it is:

- Local in scale (because the socio-economic composition may be quite different a kilometre or two away)
- Medium to long term in time (the socio-economic composition of neighbourhoods can and do change over time, but that is usually a matter of years or decades rather than weeks or months – and it cannot be controlled by the school)
- Inhibiting or distorting of reform (it does not stop the reform being implemented, but it probably reduces its effectiveness in comparison with a school that has a catchment in a locality with a high socio-economic character).

To take another example, if we look at a study of the marketization of healthcare services in Central and Eastern Europe (Nemec & Kolisnichenko, 2006) we find that the (largely) disappointing achievements of reform were deeply affected by wider economic conditions at the time.

‘probably the crucial objective issue in all countries was finance, which was tied to the level of economic performance of the country. Health reforms started to be realized in the period of more or less massive decline of GDP per capita in most CEE countries’ (Nemec & Kolisnichenko, 2006, p. 24).

In terms of the framework this would be:

- International in scale (it affected the whole CEE region)
- Somewhere between ‘quite sudden’ and ‘medium term’ in time (the decline in GDP per capita began quite suddenly but lasted for a few years – reformers could hardly have claimed to be unaware of it)
- Inhibiting or distorting of reform (the reforms went ahead but did not work out nearly as positively as had been proclaimed).



Table 4 in the Appendix categorizes and synthesizes contextual factors using the analytical framework that we introduced previously along with specific examples of studies in each category (for a more elaborate recent treatment of context see Pollitt & Dan, 2011).

### **Implications for practice and further research**

What implications for practice and future research can be drawn from our findings? First, reformers would be well-advised to include in their reform teams members with up-to-date local knowledge, both formal and tacit. Such individuals should be best placed to anticipate problems with the local culture, standard operating procedures, staff morale or tendencies to corruption. Reform teams composed entirely or primarily of consultants or reform enthusiasts from outside the target organizations or programs are inherently risky.

A second implication is that an *ex ante* systematic review of likely contextual factors constitutes a highly desirable feature of any reform plan. This analysis can identify rough areas as well as windows of opportunity which can facilitate successful implementation – or help avoid common mistakes – and ultimately the impacts of reform. In this sense our database can provide a check list of possible items to think about. These include:

- Considering the implications of the wider political system. The overall political system (broad scope) does not usually change much in the short term: reforms may therefore need adapting according to whether they are being inserted to, say, an aggressively competitive, majoritarian system or a more consensually-oriented multi-party system
- Assessing the likelihood of sudden turnovers of governments and/or restructurings of governmental structures and functions. Such volatilities may exercise a major influence on the continuity and cohesion of reform efforts in time and across governmental levels and units. This factor is more difficult to predict as it may be due

to unanticipated accidents, crises or scandals. However, in the case of some countries (Italy in the 1980s and 1990s, some CEE countries since the mid-1990s) rapid changes of government are a clear and expected pattern

- Assessing the local organizational culture. This is a longer term influence, hard to change radically in the short term but potentially malleable in the medium to long term
- Recognizing the degree of corruption existing in the particular sector/organization. Many NPM reforms are hard to implement in a high-corruption environment because they involve giving managers and front-line staff greater autonomy – in cases where corruption is prevalent it may be necessary to tackle corruption, at least in selected areas or sectors, before NPM reforms can be fully implemented. Corruption is one particular aspect of culture, and is similarly hard to transform in the short term, although it can be progressively reduced over time
- Assessing the existing capacity and skill set of the organization(s) which are to undergo reform. Skill sets can be altered in the medium term if a planned program of training and recruitment is put in place. In the short term there is little point introducing a new technique (such as regular appraisal interviews; accruals accounting) if enough staff are not trained to operate them
- Analyzing the socio-demographic situation of users and citizens that are targeted by reform, since different groups may experience the effects of reform differently and, conversely, the impacts of reform may be different, based on different socio-demographic characteristics
- Considering the role and influence of more 'organized' factors affecting implementation and possibly impacts per se, such as labor unions or (senior) civil servants who are opposed to reform and can possibly exert significant influence.

A third implication is clearly that those researching public management reform need to be particularly careful to specify the domain over which they consider their findings are likely to hold reasonably true, and to identify the main contextual influences they see as being in play within that domain. Too many of the documents in our database are loose or even silent about these issues of domain and context. Only by being more specific about these issues (and, indeed, by actively *theorizing* them) can it be hoped that sound, contextually grounded, middle level generalizations will emerge (Pawson, 2013).

Finally, we should acknowledge a number of weaknesses in our own study. Our strategy has been to seek the ‘big picture’ of European scholarship and official analyses on NPM impacts. Our picture is a picture of what the literature *says*, and one cannot guarantee that it has a very close relationship to some underlying ‘reality’. We have allowed many different kinds of studies into the database, so we cannot claim any kind of methodological homogeneity. The impact attribution claims in different studies vary widely in nature (some seeming quite weak). We cannot go beyond the crude figures reported in Tables 2 and 3 in the Appendix to seek any kind of statistical average, mean or distribution. We have no way of measuring the relative *sizes* of effects (increases, decreases etc.) or whether they are statistically significant. Furthermore there may be some double-counting, in the sense that we may have included in the database more than one study of the same reform (though where we have, the different studies usually adopt different approaches and methods). On the other hand, we may have left some important studies *out*, because of our prime focus on central government. In a number of European countries (for example Germany, Italy) the main weight of performance-oriented management reform has fallen on sub national administrations. We have picked up a few of these studies (where the reform has been initiated by central government) but there are others we have not caught in our net. Nor have

we been able to ask more subtle questions such as ‘for whom, exactly was such-and-such a reform a “success”?’ or ‘what *constitutes* success for each participant?’ All these limitations, of course, represent opportunities for future research.

### **Concluding observations**

What, finally, can we say about the impacts of central government-initiated performance reform in Europe? On the one hand there have been endless publications – academic and official – concerned with NPM-like programs and techniques. Yet, on the other, our solid, scientific knowledge of the general *outcomes* of all this thinking and activity is very limited. Even our somewhat larger understanding of *outputs* cannot be described as more than fragmentary. Claims and counterclaims outnumber hard, carefully collected evidence by a substantial margin. That was noted in the mid-1990s (Pollitt, 1995) and again by other scholars in the 21st century (Andrews, 2011; Hood, 2005; Hood & Dixon, 2012; see also Modell & Grönlund, 2007; Norman, 2007), and it remains the case today. The impacts of NPM are elusive, both definitionally and empirically.

It is true that our database holds a good deal of interesting information about changes in *outputs* resulting from performance-oriented reform, but what the collectivity of this knowledge suggests is that these vary enormously, from highly positive through ‘no change’ to rather negative, depending on circumstances. One broad conclusion is that it is often inappropriate to talk *in general* of a particular tool or technique ‘working’ or ‘not working’. Similarly the notion of an abstracted, internationally mobile ‘best practice’ begins to appear rather fragile (this point has already been noted by a number of scholars and organizations – see, in the rather different context of development administration, World Bank, 2012). Both species of generalization fail to take sufficient account of the sheer variety and power of contextual influences. To put it another way, all sorts of contextual ‘conditions’ have to be satisfied for a given performance-oriented tool to work well – or perhaps at all.

All this has implications for theory. In order to explain the impacts of performance reform, one needs a theory which grants contextual factors an explicit and substantial role and which, within that recognizes the importance of cultural and institutional factors. That, in turn, suggests that the simpler versions of rational choice theory (one of the pillars of early NPM thinking) are a problematic vehicle for this kind of explanation. Although more sophisticated versions of the theory *can* model varying contexts (Dunleavy, 1991), the typical applications in the public management literature have said little of contextual variation, and little about the vast range of cultures or institutional structures (Lane, 2000; Niskanen, 1973). They have focused mainly on individual motivations and relationships, or contractual relations which treat the contracting organizations as if they were homogenous actors – an inadequate basis for understanding the actual effects of NPM, according to many of the studies reviewed here.

Neither do generic management theories make many concessions to contextual variation – they are, by nature and intent, 'generic' (Barzelay, 1992; Kettl, 2000). But inspection of our database makes it seem improbable that 'six steps' or 'new transformational techniques' models of management will prove reliable guides for public managers inhabiting different cultures, political systems, sectors, task environments and histories.

More positively, the theoretical pre-requisites proposed in the previous paragraph *can* be satisfied. There are several theoretical approaches that might meet the bill, including historical institutionalism (Pollitt & Bouckaert, 2009), sociological institutionalism (Peters, 2005), critical realism (Pawson & Tilley, 1997; Pawson, 2013) and some varieties of organization theory (Tsoukas & Knudsen, 2003). Each of these, however, will require some modification or further refinement. The 'new institutionalisms', for example, sometimes fail to leave enough space for actors, leaders, accidents and 'situations', whereas our database shows that these can sometimes be of great importance (an election, a new leader, a scandal).

The present paper has been mainly concerned with assessing the strength of pro-NPM practitioner theory, so this is not the place to begin a deep theoretical comparison, but suffice it to say that we do not lack theoretical tools capable of tackling the range of findings which our work has brought to the surface.

Despite all these caveats, it would be mistaken to dismiss NPM as a failure. That would be just another sweeping generalization. While keeping in mind the limitations of our study, the database entries for outcomes show improvements in 44 percent of studies, for outputs the corresponding figure is 53 percent, and for processes/activities it is 58 percent. So NPM often does 'work'. Like a delicate plant, however, it seems to require the right soil and care – more the orchid than the potato. It is not robust across a wide range of environments. One 'message' would therefore be that would-be reformers need to look rather hard at several dimensions of the local context before they decide that NPM is the answer to their problems (and certainly before they decide how to implement it). This is the key take-away message from our research. It is not new, but what is new is that in our case it is based on a systematic review of an unusually large body of literature across Europe. While essentially European, our research has brought to the surface implications that practitioners and researchers in countries other than European may find helpful to consider. In this sense our implications for practice and theory are broad enough to speak to the world rather than to Europe alone.

### **Acknowledgements**

The research leading to these results has received funding from the European Union's Seventh Framework Programme under grant agreement No. 266887 (Project COCOPS), Socio-economic Sciences & Humanities. We would like to acknowledge the invaluable input in creating the NPM database made by the IT officer at the KU Leuven Public Management Institute, Anita van Gils, and our partners and collaborators inside and outside of the

COCOPS project. We would also like to thank three anonymous reviewers for their helpful comments.

## References

- Andersen, S.C. (2008). The impact of public management reforms on student performance in Danish schools. *Public Administration*, 86(2), 541-558.
- Andersen, L.B., & Pallesen, T. (2008). Not just for the money? How financial incentives affect the number of publications at Danish research institutions. *International Public Management Journal*, 11(1), 28-47.
- Andrews, R. (2011). NPM and the search for efficiency. In Christensen, T., & Lægreid, P. (Eds.). *The Ashgate research companion to New Public Management* (pp. 281-294) Farnham: Ashgate.
- Andrews, R., Boyne, G., & Walker, R. (2011). The impact of management on administrative and survey measures of organizational performance. *Public Management Review*, 13(2), 227-255.
- Aristigueta, M., & Van Dooren, W. (2007). Toward a performing public sector – the roles of context, utilization and networks. Introduction. Symposium. *Public Performance and Management Review*, 30(4), 463-468.
- Askim, J., Christensen, T., Fimreite, A. L., & Lægreid, P. (2010). How to assess administrative reform? Investigating the adoption and preliminary impacts of the Norwegian welfare administration reform. *Public Administration*, 88(1), 232-246.
- Arndt, C. (2008). The politics of governance ratings. *International Public Management Journal*. 11:3, 1-23.

- Atun, R., Menabde, N., Saluvere, K., Jesse, M., & Habicht, J. (2005). Introducing a complex health innovation – primary health care reforms in Estonia (multi-methods evaluation). *Health Policy* 79, 79-91.
- Audit Commission (2009). *Final score: the impact of Comprehensive Performance Assessment of local government, 2002-2008*. London: The Audit Commission.
- Baggot, R. (1997). Evaluating health care reform: the case of the NHS internal market. *Public Administration*, 75:2, 283-306.
- Barzelay, M. (1992). *Breaking through bureaucracy: a new vision for managing in government*. Berkeley: University of California Press.
- Barzelay, M. (2000). The New Public management: a bibliographical essay for Latin American (and other) scholars. *International Public Management Journal*, 3, 229-265.
- Bauer, M. (2009). Impact of administrative reform of the European Commission: results from a survey of heads of unit in policy-making directorates, *International Review of Administrative Sciences*, 75:3, 459-472.
- Belorgey N. (2010). *L'hôpital sous pression, enquête sur le nouveau management public*. Paris: La Découverte.
- Bevan, G., & Hood, C. (2006). What's measured is what matters: targets and gaming in the English public healthcare system. *Public Administration*, 84:3, 517-538.
- Brunetta, R. (2009). Reforming the public administration to make Italy grow, *Review of Economic Conditions in Italy*, 3, 337-368.
- Capano, G. (2003). Administrative traditions and policy change: when policy paradigms matter. The Case of Italian administrative reform during the 1990s. *Public Administration*, 81, 781-801.
- Christensen, T. (2001). Administrative reform: changing leadership roles. *Governance*, 14/4, 457-480



- Christensen, T., & Lægreid, P. (Eds.) (2001). *New Public Management: the transformation of ideas and practice*. Aldershot: Ashgate.
- Christensen, T., Lægreid, P., & Stiger, I.M. (2006). Performance management and public sector reform: the Norwegian hospital reform. *International Public Management Journal*, 9(2), 113-139
- Christensen, T., Lægreid, P., Roness, P., & Røvik, K. (Eds.) (2008). *Organization theory and the public sector: instrument, culture and myth*. London: Routledge.
- Christensen, T., & Lægreid, P. (Eds.) (2011). *The Ashgate research companion to the New Public Management*. Aldershot: Ashgate.
- Divay, S. (2009). Nouveaux opérateurs privés du service public de l'emploi: les pratiques des conseillers sont-elles novatrices? *Travail et Emploi*, (3)119, 37-49.
- Dunleavy, P. (1991). *Democracy, bureaucracy and public choice: economic explanations in political science*. Hemel Hempstead: Harvester/Wheatsheaf.
- Dunleavy, P., Margetts, H., Bastow, S., & Tinkler, J. (2006). New Public Management is dead – long live digital-era governance. *Journal of Public Administration Research and Theory*, 16(3), 467-494.
- Edelenbos, J., & Klijn, E-H. (2009). Project versus process management in public-private partnership: relation between management style and outcomes. *International Public Management Journal*, 2:3, 310-331.
- Emery, Y. (2004). Rewarding civil service performance through team bonuses: findings, analysis and recommendations. *International Review of Administrative Sciences*, 70:1, 157-168.
- Emery, Y., & Giauque, D. (2001). New Public Management, contracts for the provision of services and financial incentive mechanisms: practices and problems encountered in Switzerland. *International Review of Administrative Sciences*, 67, 151-171.

- Federal Ministry of the Interior (2006). *Focused on the future: innovations for administration*. Berlin: Federal Ministry of the Interior.
- Ferlie, E., & Geraghty, J. (2005). Professionals in public service organizations: implications for public sector reforming. In E. Ferlie, L. Lynn Jr. & C. Pollitt (Eds.). *The Oxford handbook of public management* (pp. 422-445). Oxford: Oxford University Press.
- Gore, A. (1997). *Businesslike government: lessons learned from America's best companies*. Washington, DC: National Performance Review.
- Gow, J., & Dufour, C. (2000). Is the New Public Management a paradigm? Does it matter? *International Review of Administrative Sciences*, 66(4), 573-593.
- Hammerschmid, G., Krischok, A., & Steigenberger, K. (2012). Agencification experiences in Austria. In Verhoest, K., Van Thiel, S., Bouckaert, G. & Lægreid, P. (Eds.) *Government agencies: Practices and lessons from 30 countries*. (pp. 156-165). Basingstoke: Palgrave MacMillan.
- Healthcare Commission (2009). *Investigation into Mid Staffordshire NHS Foundation Trust*, London, Healthcare Commission, March.
- Hondeghem A., & Depré R. (2005). *De Copernicushervorming in perspectief: Veranderingsmanagement in de federale overheid*. Brugges: Vanden Broele.
- Hood, C. (1990). De-Sir Humphreyfying the Westminster Model of bureaucracy: a new style of governance? *Governance*, 3(2), 205-214.
- Hood, C. (1991). A public management for all seasons. *Public Administration*, 69(1), 3-19.
- Hood, C. (1995). Contemporary management. *Public Policy and Administration*, 10:2, 104-117.
- Hood, C. (2005). Public management: the word, the movement, the science. In E. Ferlie, L. Lynn Jr. & C. Pollitt (Eds.) *The Oxford handbook of public management*. (pp. 7-26). Oxford: Oxford University Press.

- Hood, C., & Peters, G.B. (2004). The middle aging of New Public Management: into the age of paradox? *Journal of Public Administration Research and Theory*, 14(3), 267-282.
- Hood, C., & Dixon, R. (2012). A model of cost cutting in government? The great management revolution in UK central government. *Public Administration* (pre-publication version accessed 23 October 2012).
- Jacobs, R., & Goddard, M. (2007). How Do Performance Indicators Add Up? An Examination of Composite Indicators in Public Services. *Public Money & Management*, 27(2), 103-110.
- James, O. (2010). *End user satisfaction as a performance measure for public services*. Paper for the National Audit Office, Department of Politics, University of Exeter, February.
- Jeannot, G., & Guillemot, D. (2010). Réformer par les outils ou par les homes? Un bilan quantitatif de la modernisation de gestion de l'Etat. *Politiques et management public*, 27:4, 73-101.
- Kaufmann, D., Kraay, A., & Mastruzzi, M. (2007). Governance matters VI: governance indicators for 1996-2006. World Bank Policy Research Working Paper 4280.
- Kettl, D. (2000). *The global public management revolution: a report on the transformation of global governance*. Washington DC: The Brookings Institution.
- Kuhlmann, S., Bogumil, J., & Grohs, S. (2008). Evaluating administrative modernization in German local governments: success or failure of the "New Steering Model? *Public Administration Review*, September/October, 851-863.
- Laforgue, D. (2004). Le repli gestionnaire des inspecteurs d'académie. In *L'action publique au travail*, coll. Le travail en débats. Octares, 169-176.
- Lægreid, P., Roness, P., & Rubecksen, K. (2008). Performance information and performance steering: integrated system or loose coupling? In Van Dooren, W & Van de Walle, S.

- (Eds.). *Performance information in the public sector: how it is used*. Basingstoke: Palgrave MacMillan.
- Lember, V. (2008). *New Public management reforms, contracting out and Emergency Medical Service in Estonia*. Paper presented at the European Consortium for Political Research workshops in Rennes, September.
- Lane, J. (2000). *New Public Management: an introduction*. London: Routledge.
- Lord President of the Council (1981). *Efficiency in the civil service*, Cmnd. 8293, London: HMSO, July.
- Lynn, L. Jnr. (2006). *Public management: old and new*. London and New York: Routledge/Taylor and Francis.
- Macinati, M.S. (2006). Il ricorso all' outsourcing nel Ssn: i risultati di un'indagine empirica. *Mecosan*, 57, 121-144.
- Makinson, J. (2000). *Incentives for change: rewarding performance in national government networks*. London, Public Services Productivity Panel: H.M.Treasury.
- Marsden, D., & French, S. (1998). *What a performance: performance-related pay in the public services*, London, Centre for Economic Performance, London School of Economics.
- Matelly, J-H & Mouhanna, C. (2007). *Police: des chiffres et des doutes: regard critique sur les statistiques de la délinquance*. Paris : Michalon.
- Meyer-Sahling, J-H. (2009). *Sustainability of civil service reforms in Central and Eastern Europe five years after EU accession*. SIGMA paper No.44, Paris: OECD.
- Minvielle, E. (2006). New public management à la française: The case of regional hospital agencies. *Public Administration Review*, September/October.

- Modell, S., & Grönlund, A. (2007). Outcome-based performance management. Experiences from Swedish Central Government. *Public Performance and Management Review*, 31(2), 275-288.
- National Audit Office. (2009). *Private finance projects, a paper for the Lords economic affairs committee*, October 2009. London: National Audit Office.
- Nemec, J., & Kolisnichenko, N. (2006). Market-based health care reforms in Central and Eastern Europe: lessons after ten years of change. *International Review of Administrative Sciences*, 72:1, 11-26.
- Nemec, J. (2007). Performance management in the Slovak higher education system: preliminary evaluation. *Central European Journal of Public Policy*, 1(1), 64-78.
- Nemec, J. (2008). Market based health care reforms solutions and their outcomes in CEE. In Antkiewicz, S., Kalinowski, M. (Ed.). *Tendencja i uwarunkowania rozwoju finansow przedsiębiorstw i finansow publicznych*. - S. 338-352, Warszawa a Gdansk: CeDeWu a WSB Gdansk.
- Niskanen, W. (1973). *Bureaucracy: servant or master?* London: Institute for Economic Affairs.
- Norges Forskningsråd (2007). Resultatevaluering av sykehusreformen: *Tilgjengelighet, prioritering, effektivitet, brukervedvirkning og medbestemmelse*.
- Norman, R. (2007). Managing outcomes while accounting for outputs. Redefining 'public value' in New Zealand's performance management system. *Public Performance and Management Review*, 30(4), 536-549.
- OECD. (1995). *Governance in transition: public management reforms in OECD countries*. Paris: PUMA/OECD.
- OECD. (2010). *Modernising the public administration, a study on Italy*. Paris: OECD

- Osborne, D., & Gaebler, T. (1992). *Reinventing government: how the entrepreneurial spirit is transforming the public sector*. New York, NY: Addison Wesley Publishing Company.
- Painter, M. (2003). Policy, capacity and the effects of New Public Management. In Christensen, T. & Lægreid, P. (Eds.). *New Public Management: the transformation of ideas and practice* (pp. 209-229) Aldershot: Ashgate.
- Pawson, R. (2013). *The science of evaluation: a realist manifesto*. London and Thousand Oaks: Sage.
- Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. London: Sage.
- Peters, G.B. (2005). *Institutional theory in political science: the 'new' institutionalism* (2<sup>nd</sup> edition) London and New York: Continuum.
- Pollitt, C. (1990). *Managerialism and the public service: the Anglo-American experience*. Oxford: Blackwell.
- Pollitt, C. (2003). *The essential public manager*. Buckingham: Open University Press/McGraw Hill.
- Pollitt, C. (2010). Simply the best? The international benchmarking of reform and good governance. In Pierre, J. & Ingraham, P. (Eds.). *Comparative administrative change and reform: lessons learned*, (pp. 91-113). Montreal and Kingston: McGill-Queen's University Press.
- Pollitt, C. (2011). Moderation in all things: international comparisons of governance quality. *Financial Accountability and Management*, 27:4, 437-457.
- Pollitt, C., Keen, J., Packwood, T., & Roberts, S. (1998). *The Transforming Healthcare Delivery program at King's Healthcare: an evaluation*. Uxbridge: Centre for the Evaluation of Public Policy and Practice, Brunel University, February.
- Pollitt, C., Talbot, C., Caulfield, J., & Smullen, A. (2004). *Agencies: how governments do things through semi-autonomous organizations*. Basingstoke: Palgrave Macmillan

- Pollitt, C., & Bouckaert, G. (2009). *Continuity and change in public policy and management*. Cheltenham: Edward Elgar.
- Pollitt, C., Harrison, S., Dowswell, G., Jerak-Zuiderent, S., & Bal, R. (2010). Performance regimes in health care: institutions, critical junctures and the logic of escalation in England and the Netherlands. *Evaluation* 16:1, 13-29.
- Pollitt, C., & Bouckaert, G. (2011). *Public management reform: a comparative analysis: New Public management, Governance and the Neo-Weberian State* (3<sup>rd</sup> edition). Oxford: Oxford University Press.
- Pollitt, C., & Dan, S. (2011). The Impacts of the New Public Management in Europe: A Meta-analysis. Deliverable 1.1. COCOPS Project ([www.cocops.eu](http://www.cocops.eu)).
- Prime Minister and the Minister for the Cabinet Office (1999). *Modernising government*, Cm. 413, London: The Stationery Office.
- Promberger, K. (2008). *Evaluation der Verwaltungsmodernisierung In Österreich am Beispiel des Pilotprojekts Flexibilisierungsklausel*, Jubiläumsfondsprojekt Nr. 11998, Innsbruck, Juli.
- Randma-Liiv, T. (2005). Performance management in transitional administration: the introduction of pay for performance in the Estonian civil service. *Journal of Comparative Policy Analysis, Research and Practice*, 7:1, 95-115.
- Roberts, A. (1997). Performance-based organizations: Assessing the Gore Plan. *Public Administration Review*, 57(6), 465-478.
- Societatea Academică Română. (2010). Pentru o autonomie responsabilă, expansiunea universităților în România, SAR Policy Brief No. 46, București, Iulie.
- State Audit Office. (2002). Utilization of Payroll Funds Allocated for the Payment of Performance Wages. Tallinn: State Audit Office.

- Stoker, G. (2006). Public value management: a new narrative for networked governance? *American Review of Public Administration*, 36:1, 41-57.
- Șandor, D. S., & Tripon, C. R. (2008). Citizens' Perceptions about Public Administration Reforms in Romania. *Transylvanian Review of Administrative Sciences*, 24E, 111-122.
- Torres, L., & Pina, V. (2004). Reshaping public administration: the Spanish experience compared to the UK. *Public Administration*, 82(2), 445-464.
- Tsoukas, H., & Knudsen, C. (2003). (Eds.). *The Oxford handbook of organization theory: meta theoretical perspectives*, Oxford: Oxford University Press.
- Van de Walle, S., Van Roosbroek, S., & Bouckaert, G. (2008). Trust in the public sector: is there any evidence for a long term decline? *International Review of Administrative Sciences*, 74(1), 47-64.
- Verheijen, T. (2007). *Administrative capacity in the new EU member states the limits of innovation?* World bank Working paper No.115. Washington, DC: World Bank.
- Verhoest, K. (2005). Effects of autonomy, performance, contracting and competition on the performance of a public agency: a case study. *Policy Studies Journal*, 33:2, 235-258.
- Wilson, D., & Piebalga, A. (2008). Performance measures, ranking and parental choice: an analysis of the English school league tables. *International Public Management Journal*, 11:3, 344-366.
- World Bank (2001). *Salary supplements and onuses in revenue departments*, Final Report. Washington, DC: World Bank.
- World Bank (2012). *Better results from public sector institutions*. Washington, DC: World Bank.
- Yesilkagit, A., & De Vries, J. (2002). The unanticipated consequences of decentralization and reinvention: the case of the Province of South Holland. *International Review of Administrative Sciences*, 68:4, 579-597.



Christopher Pollitt is Emeritus Professor of Public Management at the Public Management Institute, KU Leuven in Belgium. He is author of more than 60 scientific articles and author or editor of more than a dozen scholarly books. Among the better known are *Public management reform: a comparative analysis* (with Geert Bouckaert – third edition 2011) and *Time, policy, management* (2008). He has also undertaken extensive consultancy and advice work for a wide variety of organizations, including the European Commission, the OECD, the World Bank, H.M Treasury, the Finnish Ministry of Finance and the Dutch Ministry of the Interior.

Sorin Dan is Researcher at the Public Management Institute, KU Leuven in Belgium. He worked on the first work package of the EU-funded project *Coordinating for Cohesion in the Public Sector of the Future* (COCOPS) and conducts research on public management reforms and coordination in healthcare and the public sector more generally.

## APPENDIX

### Literature search

Each of the ten partner universities reviewed the literature in their own country and any other country they had agreed to cover. These reviews were the basis for their selections of suitable documents/studies to upload to the database. In addition, we conducted a literature search in which we checked everything that had been published between 1980 and the middle of 2011 in a number of the leading mainstream Anglophone public administration journals:

Journal of Public Administration Research and Theory

Public Administration

Governance

Public Administration Review

International Review of Administrative Sciences

International Public Management Journal

Public Management Review

Public Policy and Administration

## Evaluation

In addition to these academic sources we scrutinized a number of promising organizational publication lists, including those of:

OECD

SIGMA

World Bank

National government websites on central government reforms (in those countries where they are available)

National audit office performance audits (in those countries where they are available)

Certain civil society organizations in Central and Eastern Europe

*Place Table 1 here*

*Place Table 2 here*

*Place Table 3 here*

*Place Table 4 here*

---

<sup>i</sup> The database of NPM reforms in Europe is available online at

<http://www.cocops.eu/work-packages/wp1-npm-meta-analysis/database-of-studies-of-npm-reforms-in-europe>