

New Public Management, science policy and the orchestration of university research – academic science the loser

*A ELZINGA**

Abstract

In many countries - during the past three decades - there has been increasing alignment of public administration with neoliberal modes of governance driven by processes of globalization and privatization. Key is New Public Management (NPM) as an instrument for applying private sector or market-based techniques to public services. The paper concerns the implications and impact of these developments as they relate to the university sector where we are seeing an influx of bibliometrics to assess performativity with measures that are frequently far from adequate. These changes are related to the broader context of a globalized privatization doctrine of science policy that has gained currency in many countries. The analysis presented here probes and discusses a nexus between NPM, bibliometric performance audits, and a new paradigm in science policy. As such the paper seeks to fill an important gap in science policy studies and the literature on New Public Management more generally. It summarizes various characteristics associated with NPM, and explicates the connection with methods of research evaluation now being introduced in academic institutions. It also takes up varying responses evoked within academe by the use of bibliometrics and current methods of ranking of tertiary educational institutions. Apart from gaining a better understanding of significant changes in the higher educational and research landscapes or disciplines and the interplay of these with broader economic and political trends in society at large, the aim of the paper is also to stimulate discussion and debate on current priorities, perceptions and policies governing knowledge production.

Keywords: New Public management, research policy, transdisciplinarity, postnormal science, Audit Society Agencification, Accountingization, peer review evaluation.

Disciplines: Public Management and Governance, Management Studies and Research Management. Science Policy, Science Studies, Political Science,

Introduction

In many countries - during the past three decades - there has been increasing alignment of public administration with neoliberal modes of governance driven by processes of globalization and privatization. In the UK New Labour social democrats and Conservatives urged this trend on, equally. Stephen Cope and others (1997), analyzing the connection between globalization and New Public Management (NPM), state that,

* Prof. Aant Elzinga, University of Gothenburg, Sweden. aant.elzinga@theorysc.gu.se

(I)t is no accident that the wave of NPM crashing through governments worldwide corresponds with increasing globalization ... States have restructured as a result of globalization. NPM has been a very prominent form of restructuring in most Western governments, though its spread has been uneven. NPM rests on the twin doctrines of removing differences between public and private sectors, and shifting 'methods of doing business in public organizations', away from complying with procedural rules /and towards 'getting results' (Hood, 1994. p. 129).

NPM is shorthand for applying private sector or market-based techniques to public services, an approach that rapidly spread through the Anglophone world in particular. As a politico-administrative regime it has been successively introduced into one sector after another, hastening, systematizing and legitimating these developments, making it appear that it is only a technical-administrative question mainly of cost reduction, increasing flexibility, and greater accountability and efficiency of performance on the part of public service agencies and their various practitioners (Lane 1994, 1995 & 2000). The process began to take form in the UK during Margaret Thatcher's government (1977-1990). The label NPM is usually attributed to the renowned Oxford professor of government Christopher Hood (1991; 1995).

Problem and purpose

The problem addressed by the paper concerns the implications and impact of the aforementioned developments as they relate to the university sector. Although academic institutions and working conditions are being transformed with an influx of bibliometrics to assess performance with measures that are frequently far from adequate, few analysts have explicitly related these changes to the broader context of a globalized privatization doctrine of science policy that has gained currency in many countries. This is ironic since social scientists themselves have analyzed the impact of globalization and privatization via NPM methods in many other sectors of society, from health care to transport systems, and other essential public services and infrastructures where privatization and "marketization" with subsequent erosion of "the commons" has been identified, analyzed and debated. But, when it comes to our own workplace – academe – a nexus between NPM, bibliometric performance audits, and a new paradigm in science policy is mostly discussed only in anecdotal terms.

The purpose of the present paper is therefore to fill an important gap in science policy studies and the literature on New Public Management more generally. In particular the present analysis develops a novel angle in science policy analysis. Additionally, on the basis of a review of the more general literature, a second purpose is to summarize at length various characteristics associated with NPM, one might say its "anatomy" and dynamics, in order to more precisely delineate and explicate the connection with methods of research evaluation now being introduced in academic institutions. A third aim of course is to stimulate discussion and debate.

Structure and argument

To set the stage the paper characterizes several features of NPM and the concomitant shift in focus from "administration" to "management". Since the changes in terminologies that attend this shift are significant they are also dealt with from a *metatheoretical* point of view.

At a more general level the shift identified by the paper reflects the change from "government" to "governance", terms that will be explained below. Other changes of meaning associated with further concepts are also touched upon.

Finally it is demonstrated how the tertiary educational sector is made a regular part of what has been called contemporary *Audit Society* and that this in its current form impacts academe negatively in as far as short term interests override the strategic need for basic research and knowledge production for the long term future.

Even though some readers may find it somewhat tedious, in order to substantiate the paper's claim that we are witnessing a global trend, reference has to be made to developments, experiences and responses in several countries,

In the USA some scholars have come to refer to the present situation as one of "academic capitalism" (Slaughter and Leslie 1997). The suggestion is not that most entrepreneurial individuals in research and teaching have become small capitalists. Rather the point is that a capitalist-like behaviour and (contractual) relationships have become rather prominent within the moral economy (norms or ethos) of academe.

This has on the other hand also evoked calls to a return to the traditional values and collegial norms associated with the ideal of the Humboldt-university of the 19th century that gave primacy to a wide-ranging freedom of research and a symbiotic interplay between research and teaching. The introduction of university ranking lists has, symptomatically, also led to protests from scientific communities (Butler 2007) that need to be understood in the context of broader trends.

In the concluding section there is a summary of some findings based on the review of relevant literature covering experience with NPM and new forms of research evaluation as well as discussion of the claim that, ultimately, *Audit Society* engenders a paradox that undermines its efficacy and hence renders it counter-functional.

Characteristics of New Public Management and Audit Society

It should be noted that NPM is an international trend, but that it has not been adopted wholesale as a package in all countries. One finds different versions in different countries. These differences have been explained by differences that exist between countries with regard to legal and administrative traditions, political style and cultural determinants (Christensen and Lægeid 2005). This means that it is wrong to assume that the increasing adoption of NPM methods brings with it a general global homogenization in methods of policy development and implementation. What we see rather is a common fundamental rationale or logic combined with variety in implementations in different countries.

One of the main arguments for the shift to business-mindedness in public utilities or services has been to introduce greater cost consciousness, flexibility and efficiency, to make do with less in order to produce more (Barzalay 2001). This is attended by a shift from administration to management in the public domain. Units responsible for delivering products or services are impelled to cover their own costs as far as possible, and to encourage this quasi-markets are set up in relation to which various units delivering similar kinds of products or services are induced to compete with each other. Another factor is to diminish traditional bureaucratic forms of administration with the intention of opening for greater sensitivity to public needs and by extension the possibility of more "user" participation in decision- and policy-making processes. In market economic terms the user of services and products supplied by public institutions is framed as customer or client and "the customer is always right". What this means in practice is however a contentious issue.

TD, 6(2), December 2010, pp. 307-332.

Manifestations of NPM differ across policy areas in any given country and between countries. This also goes for the domain of science policy, which as already noted is, moreover, a relative latecomer when it comes to the entry of NPM. Particularly interesting in the case of science policy is the issue of accountability. A feature in this connection that is common to most countries is the use of bibliometrics, that is publication counts and citations to measure research performance, and how results of such exercises are then related to economic cost-benefit assessments (econometrics). Since accountability and “accountingization” using numbers takes on a particular form in the research and tertiary educational sector it is dealt with at some length below. Science policy as a policy domain also has specific characteristics that must be briefly outlined to contextualize the issue further.

Suffice it at this point to underline more generally the occurrence of a shift from accountancy of resource inputs to evaluation and assessment of outcomes or outputs. This is a necessary corollary to the fundamental change in doctrine relating to accountability for public sector institutions. Former trust in the wisdom and behaviour of traditional civil servants and professions as non-opportunistic is replaced by mistrust. This follows logically from the premises of a regime in which the supplier of utilities, services and other goods has to mantle the role of a rational calculator of opportunity costs without concern for the impact cutbacks and speedups might ultimately have on externalities. By the same token it is therefore logical also to find an increasing demand for control by means of externally initiated evaluations and assessment or auditing procedures as well as design of internal control systems to harmonize and comply with the new external conditions or demands. Michael Power and Richard Laughlin (1992) have called this “accountingization” (also see Almqvist 2006:24). They also argue that a significant consequence for public organizations is the emergence of new types of cultures of compliance.

Three main ingredients of the NPM narrative are:

- *Competition;*
- *Agencification*, i.e., introducing contractual relationships; and
- *Accountingization.*

Service deliverers in public agencies in a given sector, like health care or transport, are brought into competition with each other through competitive tendering. Contractual arrangements between a principal and an agent are articulated to specify goals and expectations (agencification and contractualization). Together with outsourcing to supplier firms and consultancies this fragments government in client/contractor and purchaser/provider splits (Cope and others 1997:449), whence a Balkanization of the public sector and emergence of public-private partnerships may occur. Agencification, also called “corporatization”, converts departments into free standing units with downward delegation of responsibilities for tasks to lower managers while strategic choices are centralized. Decentralization-centralization goes hand in hand. In terms of human resources a corresponding feature is a move from a Weberian styled civil service career system to a contract system with short-term appointments in the lower ranks of employees as well as the emergence of an array of private-public partnerships. This is meant to increase flexibility but in practice it also introduces lack of continuity and lapses in institutional memory that ultimately can have a negative effect on the delivery of utilities, products and services in the public sector.

Degrees of goal fulfilment are determined by introducing performance assessment schemes (accountingization) to measure the same (more on this below).

A more complex schematism based on a review of pertinent literature will include further catchwords. Some of these are summarized in Table 1. It lists both descriptors and a few normative characterizations.

Table 1: Some basic NPM notions

- * More for less¹
- * Marketization (including creation of quasi-markets in administrative organizations)
- * Commoditization of health care services, welfare benefits, teaching packages and research results (also those generated by publicly funded institutions)
- * Inducing competition between task performers
- * Turning citizens into consumers and clients (this goes for students too)
- * Agencification (contractification)
- * From administration to management (fostering the entrepreneurial bureaucrat)
- * From input to output/outcome control
- * Performance-based management (and funding), Performativity metrics (accountingization)
- * Reputation and image management (PR & branding)
- * Entrepreneurialism
- * Partnering

Statements regarding impacts in three different policy sectors

In the literature one can find many case studies on the impact of NPM. Some of the consequences are particularly evident in the case of development policy in Third World countries. One author who sees NPM as a gain since it introduces flexibility and provides for cost recovery in waste management systems in big cities, for example, nevertheless observes how deregulation has led to a proliferation of private entrepreneurs in the waste collection business, many of whom do not uphold the standards called for by policies meant to promote sustainable development (van Dijk 2008). Another author notes how in development policy under NPM regimes:

Result-based management approaches that have entered the field have the paradoxical effect of disempowering those it meant to empower (Kilby 2004:207).

A case study on a public-private planning initiative in a city in the USA that used consensus-based collaborative meetings involving citizens in “visioning” to shape the future cityscape ended up with elites capturing the mechanics of the process. This led to a situation where dissent is difficult and the outcome tends to be a vision of the future that replicates the standard economic development models of powerful stakeholders (McCann 2001). The process becomes a kind of “therapy” similar to what we observed in the case of Gothenburg, Sweden, a city where, by the

¹ For some of the characteristics see Dreschler (2005).

way, NPM has produced a system where garbage collection has become the monopoly of a couple of large entrepreneurs that cut costs by employing fewer (overstressed) truck drivers and reducing the frequency of garbage runs so that at many recycling stations heaps of waste regularly accumulates between the various bins for separating two sorts of bottles, various cans, plastic containers, bulky package wrapping and cartons, newspapers, etc.

In Denmark Susan Wright a social anthropologist who is now at the Danish School of Education has turned her attention to the bibliometric auditing system of performance indicators in place in Danish universities and what impact that is having. In her anthropological perspective - as outlined in an earlier study together with Cris Shore regarding the influx of audit practices in British higher education - the new audit system and management technique are seen to function as “political technologies” for introducing neo-liberal systems of power (Shore and Wright 1999; also 1997). Now that globalization has incrementally reduced the power of government issuing from one centre and supplanted it by an orchestration of a multiple stakeholders with particularist interests “audit embodies a whole new rationality of governance”. Since the language of governance replaces that of government, our sensitive anthropologists discover a kind of doublespeak.

One of the main claims made by the architects and advocates of auditing is that it ‘enables’ individuals and institutions to ‘monitor’ and ‘enhance’ their own performance and quality, and to be judged by targets and standards that they set for themselves’

Shore and Wright argue, and go on:

This language of quality enhancement suggests that audit is an open, democratic, participatory and ‘enabling’ process; a process so uncontentious and self-evidently positive that there is no logical reason for objection. The new ‘habitual grouping’ of audit with words like ‘efficiency’, ‘effectiveness’, ‘best practice’ and ‘value for money’ disguises its hierarchical and paternalistic roots and plays down its coercive and punitive implications. (Shore and Wright 1999).

The number of examples of case studies of the influence of NPM in urban planning and other areas may be multiplied and of course one will find instances of gains in efficiency and flexibility of planning as well as some success stories. The major gist however seems to be one that confirms a shift from politics to economics and with it a systemic deficit in *citizen empowerment*, in as far as citizens are construed (or “reconstructed”) as consumers and public consultation exercises turn into opinion polling and market research (Ryan 2001). This is not the same as Participation in decision-making.

Two key terms: accounting and governance

“Accountingization” and auditing practices to satisfy it is the feature that is particularly interesting, both generally and in current discourses on science policy. It refers to an objective to make visible, break down and categorize costs in areas and endeavours where such costs earlier or traditionally were aggregated or more or less undefined. Under the auspices of NPM steering, follow up and evaluation or auditing are thus emphasized, predicated on a buyer-seller or principal-agent contractual and cost-cutting nexus that replaces former trust in providers and administrators of welfare tasks to serve citizens or provide public goods. Citizens are no longer, they are now clients and consumers and we get privatization and commoditization of former public goods including portions of scientific knowledge (*qua* intellectual property). Functions

and mechanisms for dividing public resources are thereby supposed to move from the realm of politics to that of the marketplace.

Paradoxically though, the quasi-markets are still created by political decisions. But now it is no longer called government but *governance*, i.e., a multi-level chain of delegations and decentralization in the process of orchestration and steering in tune with signals from the market. Government thus becomes negotiated or orchestrated “governance” and administration becomes business “management”. Government entails formal institutions of state and legal frameworks involving old style civil servants. Governance involves government plus looser processes of influencing and negotiating with a range of private corporate actors and public sector agencies as stakeholders where the stronger ones tend to set agendas. One gets what I want to refer to as “orchestration policy”, where “good governance” tends to get equated with “sound economic development” and in practice - when applied - is routinely taken to mean investment liberalization, services commercialization, generalized privatization and other “market” policies.

Even if NPM is now retreating at the theoretical level and making room for New Weberian Public Management thinking that once more begins to emphasize the necessity of a “commons” in order to counteract partitioning and fragmentation, it is safe to say that NPM still reigns in a lot of thinking and action in policy-making. At the same time it is meeting resistance and vocal opposition from the side of critical scholars and practitioners who demand reflexivity and attention to human rights and justice when it comes to basic needs like health and water.

NPM has fostered a predominant mode of accountability that requires reliable and hopefully fine-grained metrics of performance measures in terms of various quantitative indicators. Rewarding actors on the basis of measures of performance, some studies on health care have shown, however, may not lead to lower costs because the generation of some forms of health care that is not really necessary gets prioritized since profit-driven health care makes it competitively advantageous to encourage fairly healthy individuals to make quick extra visits to the doctor. This proves to be more profitable for the health care provider than does involvement with elderly patients with chronic illnesses, for example. The quest for profit gains thus appears to introduce a skew that is for the most part not taken into account in economic cost benefit measures. Mainstream econometrics when applied to the deliverer of services lacks the necessary categories that would be needed to consider broader social costs. In the case of policies relating to environmental protection and future climate change the precautionary principle has been introduced together with environmental economics to try and deal with such broader issues by putting a price on loss of further qualities that are seen to be essential for desirable life conditions.

Research as a domain for public policy

In the sector of tertiary education and research NPM has already had a lasting influence. This is manifested in the attention and energy that have been expended by universities to fine-tune mission statements and visions, making visible profiles and articulating various strategy documents. Some of these activities come under the heading of what is nowadays called “branding”. Public relations and alumni fundraising are also part of the picture. Some of the changes are far-reaching, while others are cosmetic. A common denominator here too is the new and fashionable notion called “governance”.

Before going into further detail concerning the use of bibliometric measurements of academic performance in various disciplines to satisfy accountability criteria rooted in NPM practices it is necessary to briefly characterize science policy as a specific public policy domain. Additionally it is useful to consider the paradigm shift in science policy doctrine that has occurred more or less contemporaneously with the end of the Cold War.

Science policy as a separate policy domain emerged after the Second World War and it became institutionalized under the influence of OECD (Godin 2003, 2005 & 2006:654-655; Elzinga 2005). Important distinctions like the one entailed in the acronym R&D (research and development) adopted in the early 1960s for comparative statistical purposes are still with us today, although now intermediate categories like mission oriented (basic) research, strategic research, frontier research (term used in Europe) and transformational research (term appearing in the USA) get considerable attention. Unless one subscribes to some species of idealism or essentialism, strictly speaking the term basic or fundamental research has no meaning outside its use as a statistical household word: in other words it is historically and socially contingent and sometimes regarded as a contested term (Stokes 1997; Godin 2007:28-43). Nevertheless, rhetorically the term remains very important in science policy when it comes to identifying longer-term priorities and counteracting the strong pull toward applied science associated with innovation policy (Gulbrandsen and Kyvik 2010). Innovation policy, when it collapses science and technology under a single (fashionable) heading called “technoscience”, gets in the way of cultivating a broader knowledge base that may be important in the longer term.

Consequently it is not unimportant what kind of sociopolitics drives the construction of indicators to measure “development” of science and technology (S&T). This measurement within nation states occurs in accordance with particular conventions together with rules of standardization codified within the intergovernmental frameworks of either the OECD (for the rich world) or UNESCO (for the so-called developing nations).

Another important distinction introduced in the early years was the one between policy for science and science for policy (Elzinga and Jamison 1995). The former has to do with the stimulation of research to develop an advanced knowledge base while the latter is what we now associate with sectoral research policies (Elzinga 1980). Under the head of science for policy one can also include various functions and mechanisms for science advice to government (Pielke 2007).

Disciplines and transdisciplinarity – from normal to postnormal science

Early on science advisors were often physicists. Today the accent has shifted towards personalities with backgrounds in biomedical and biotechnological fields - and more recently also the environmental sciences. In some cases systematic methods of reviewing the literature have been developed to provide decision-makers and international organizations with a science-base for important policy decisions. The most visible example today is the Intergovernmental Panel on Climate Change (IPCC) and its periodical reviews with lead authors and teams producing chapters in comprehensive reports regarding trends in global climate change (Elzinga 1996).

When the IPCC was first instituted it represented a novel mode of mediating between state of the art science, and political action. Bert Bolin, the first chairman of this important intergovernmental forum emphasized the importance of integration of findings from many different disciplines (*multidisciplinarity*) on the one hand and a bridging of science and policy-

making on the other (*transdisciplinarity*). Initially great care was taken to maintain a strict boundary between science and politics, but as time went on the urgency of the message of global climate change led to an emphasis on powerful visual and graphical imagery to gain the attention of policy makers and the public in an effort to press for political action.

When finally in some instances (in Working Group 2 – assessing intertwining scientific, technical, environmental, economic and social aspects of climate change) dramaturgy led to insufficient attention to empirical data (e.g. the case of glacier retreat in the Himalayas), we got a situation that in turn played into the hands of longstanding lobbies of so-called contrarians or climate change sceptics whose minority voices oftentimes are amplified beyond proportion in controversy-hungry media. Recently therefore efforts are afoot within IPCC to re-organize the *transdisciplinary* interplay between science and political action in order to strengthen the integrity of the evidential base and hence the credibility of the advice scientists can provide for decision-making in an area of strong and conflicting extramural stakeholder pressures, some of which are intimately linked to the pressures of economic globalization.

Climatology is only one area of science that is in strong interaction with external stakeholder interests, and therefore it is sometimes referred to as an example of what is called “postnormal” or “postacademic” science (Ravetz 1999). The term “postnormal science” refers to areas where stakes are high while epistemic certainty is low and the (sometimes contested) precautionary principle is invoked to guide policy.

In climatology and other areas of environmental science friction may also arise from the internal dynamics of multidisciplinary research as for example in the interplay between glaciologists (or for that matter forest ecologists) out in the field concerned with empirical work “on the ground” and (the relatively more “visible”) general circulation modellers in central laboratories. This occurs when the latter “dictate” what parameters and data formats the former should prioritize for incorporation into complex simulation models. In such cases gaps between “ground truth” in the field and proxies in the models as well as the existence of different research cultures (and epistemologies) at work in the field and in the lab, respectively, may also lead to misunderstanding and conflict.

Paradigm shift in science policy regimes

A recent periodisation of science policy “regimes” proposed by Philip Morowski and Esther-Mirjam Sent (2008) places successive policy doctrines in the broad context of geopolitics. The authors argue that traditional analysis has tended to focus mainly on features that are characterized in terms of national innovation systems and policy advisory arrangements, how these are funded and comparisons of different priorities over time but with little reference to driving forces in broader economic and political terms. They exemplify by pointing to the distinction that became so popular during the 1990s (cf. Gibbons and others 1994) between academic disciplinary research (so-called Mode 1 scientific knowledge production) as different from research in the context of application (called Mode 2), where the latter was mainly characterized as interdisciplinary, multidisciplinary and transdisciplinary (i.e., including interaction with users of scientific knowledge in society at large, particularly industry and government sponsored agencies – cf. Elzinga 2008).

Parallel to this there was the notion of a Triple Helix of an intensified interlacing of academic, industrial and governmental relations to promote innovation (Etzkowitz and Leydesdorff 2000). In both cases, the two authors say, one gets an apparent ideal typical picture of two phases in the development of science in society in historical time, first one centred on academic institutions and thereafter a second one when applied research and stronger links with socially relevant objectives dominate. It is as if academic research in an earlier period, for example in the 1920s and '30s lacked external institutional motives of societal relevance.

They go on to show how the concept of Mode 2 and with it the idea that the late 1980s marked a break in the traditional “social contract of science” and the advent of a new social contract in the 1990s in fact constitutes a misrepresentation of the longer-term historical record for contemporary tacit political purposes. The term “social contract” applied to science policy did not exist before. It has only entered the science policy discourse in connection with NPM thinking and its language that now in retrospect has “discovered” past contractual relationships between science and the state where - in comparison to the situation today – these were never really articulated as such before. The overriding interest in the new policy discourse has been to legitimate a major economic and political turn that was taking place around in the 1980s, whence principal agency and contract-theory rhetoric comes in handy – thus the claim that a “new science-society contract” had emerged and stabilized. It was not a matter of conscious conspiracy but rather a rethinking in science policy circles that reflected broader realignments and stakeholder interests in society (Elzinga 2002). Hying descriptions of research in the context of application without looking for causal factors that go beyond references to the internal dynamics of innovation systems, in other words, with Mirowsky and Sent’s critically reflexive lens, can be seen to play into a broader political agenda. From the vantage point of their meta-theoretical approach these two authors therefore insist that one also has to consider the globalizing political economy.

To counteract the narrow myopic perspective they identify global events like the advent of the Cold War and later its termination with the collapse of the Soviet Union as important factors to take into consideration. Consequently they identify, apart from an older regime beginning in the late 19th century, *two* newer ones, namely: first “WWII to 1980” as the Cold War regime, followed by the “1980-? //and onward//” the Globalized privatization regime” of science organization. It is in the latter context that we find the increasing use of market mechanisms as instruments to hitch academic institutions more systematically to the carriage of a policy for innovation, thereby depleting policies for science.

“The academic profession is the loser”

The German sociologist, Uwe Schimank recalls how Burton Clark once described the “traditional” German university as a combination of political regulation by the state and professional self-control by an “academic oligarchy” (Schimank 2005: 363). He then goes on to make an interesting point. Analyzing the entry of NPM, Schimank finds that a *reduction* in academic self-governance is one of its goals.

NPM strengthens hierarchical management by rectors and deans, as well as by state authorities and external stakeholders – including industry – while implying deregulation in budgeting and personnel management, and the approval of programmes. This is what Government usually means when it promises greater institutional – not professional autonomy to universities (Ibid., 365).

It is to open universities up for external stakeholder influence. Deeply mistrustful, policy makers dismiss as ivory tower intellectuals those who give vent to nostalgic fear of a loss of professional autonomy. These policy makers read professional “autonomy” as “irresponsibility” and therefore substitute local institutional autonomy to reduce old style professional autonomy. Schimank argues that we have to be alert when the word autonomy is uttered and ask ourselves what kind of autonomy, for whom and why. Seen in this light the audits that are becoming part of daily life in academe are disempowering academic professionals.

Schimank goes on to say,

In this game, the academic profession is the loser. External interests, university leaders, and especially the Government seem to be the winners ... To achieve competitiveness, universities deregulate, create new leadership, and accept a greater measure of public intervention. Spelled out in this way, it becomes clear that NPM is not just a bundle of loosely coupled changes, but rather an integrated approach, seeking an overall redirection of the entire system. Its message: replace the old regime, dominated by a state regulated profession, with a new regime, dominated by a market- and state-driven organization” (Ibid.).

This is what Mirowki and Sent also have in mind when referring to the new situation in science policy as a privatized globalization regime.

Regarding the distinction already alluded to, between science for policy and policy for science (Elzinga and Jamison 1995) it is nowadays sometimes misleadingly suggested that the former is largely evidence based whereas the latter domain of policy is anecdotal.

The so-called evidential base in the case of science for policy is actually still heavily reliant on stories and metaphors as instanced by the schematic a-historical stories advanced by the advocates of Mode 2 and Triple Helix narratives (for a critical review see Elzinga 2004). Also there has been a tendency where science policy analysts lose their arms length analytical distance when they enter into very strong symbiotic relationships with science-policy makers. In the case of European countries this has been seen to occur when leading science policy analysts with an academic base simultaneously serve as consultants in the corridors of the European Commissioners in Brussels (Elzinga 2002 and 2004). This has led to a co-production of some key notions that are then provided with an academic stamp of legitimacy and authority by university-based researchers and then circulate (recycle) throughout the worlds of research policy makers, administrators and university leaders in various countries.

Entry of New Public Management (NPM) into science policy

It is here in the wake of neo-liberalism and in the context of the globalization regime that NPM comes in as a handmaiden. The shift to the new regime of knowledge production characterized by Mirowski and Sent is administratively helped by it. NPM as shorthand for applying private sector or market-based techniques to public services is played out in a specific way in the new globalized privatization regime in the realm of science policy.

As already noted NPM does not necessarily imply a kind of global homogenization process. There is considerable variation in the way countries adopt its logic. Further, one finds variations between policy domains. In the realm of science policy bibliometric methods are increasingly used to measure performance in the research and higher educational landscape, therewith

influencing and changing academic cultures of accountability. The kinds of bibliometric measurements (metrics) that are developed nevertheless can vary considerably from one country to another. At the same time there is considerable cross-fertilization between evaluation units, a process that together with the advent of the first European graduate school in the field of scientometrics will probably contribute to further standardization (Gorraiz and others 2010). To a certain extent the actual approach opted for by policy-makers at the national level is influenced by the topography of the research landscape and the power and coherence of various scientific lobbies within it. Thus a country with a strong tradition of scholarship in the humanities and social sciences but less developed areas in the engineering and natural sciences or medicine will probably shy away from strict bibliometric approaches that only rely on international databases whereas the reverse may be the case when humanities and social sciences are weaker and overshadowed by the dominant “hard” sciences.

Goals, priority setting and incentive systems – from ideology to scoreboards

Research policy in its classical sense has to do with setting goals and priorities for research and development (R&D) (Ziman 1994). The question of allocating resources to R&D activities is one of science policy’s most classical components. For the most part it is a matter of allocating resources out of the public purse at the national level, but it can also concern investments on regional or local municipal levels decided by actors at these levels to stimulate new and emerging technologies (high tech) with an eye to the region’s and the nation’s future competitive advantage in a global marketplace.

It appears then that much of the current discussion about suitable models for resource allocation to research has to do with a demand of accountability in terms that are quantitative and measurable. Helga Nowotny, whom I shall cite more below, has noted that even if the word “evidence” is seldom used in science policy the philosophy nevertheless lies behind the attempts to construct metrics for fine-grained evidence regarding performance as a basis for decisions on continued financing in a part of the public sector. Although the use of quantitative measures to evaluate research performance has increased enormously there is a dearth of literature with a reflexive take on the subject.

To begin with it is useful to make a distinction between two kinds of data or evidence used as a basis for policy-decisions. Let us call them a first and a second order of evidencing.

The first order of evidence concerns situation-descriptors that are needed if one wants to tighten policy audits and facilitate future comparisons between conditions before and after the implementation of a policy or package of affiliated instruments.

A second order of evidence concerns evaluation of science policy and its instruments. It calls for reflexivity of a kind one does not often find in planning and funding agencies. This imbalance perhaps explains why so much of the energy of planners today seems to be directed towards constructing and experimenting with fine-grained metrics that may be incorporated into decision-making related to annual budgetary allocations of resources to publicly funded R&D at universities and other institutions. This is evident in several countries where performance indicators are used as a basis for changing funding flows between universities as well as between faculties and research units within them.

The policy discussion on accountability as influenced by NPM has to do with evidence of the first order, i.e., benchmarking and descriptions of states of affairs within publicly funded units with the help of numbers as well as qualitative assessments (e.g., using review panels) in a

landscape that is subject to policy orchestration by incorporating cultures of compliance. Evidence-basing of policy in the sense one finds in medical clinical work or units responsible for social work would require a second step. Such a second step would entail systematic evaluations of the efficacy of actual policy measures and instruments in order to determine to what degree intended aims or goals are achieved, for example that the prioritized profiled areas of R&D progress in a suitable manner when seen from the combined point of view of quality enhancement and relevance.

Another example would be assessing the actual impact of policies aimed at increasing a country's international economic competitiveness, or intensification of collaboration between industry and universities in priority areas. In my reading of science policy documents I have not been able to find many credible examples of evidence-basing in such a (second order) meaning.

In order to be able to compare the result of an implementation of specific policy instruments and established goals one needs to start with a much more precise description of the existing situation that is the baseline from which one has to start. This is probably the reason why much of the talk about "best practices" is still rather loose while a lot of emphasis is given to introducing "benchmarking" that will give clearer points of reference to compare initial conditions before and final conditions after the implementation of a policy package.

The presidential election campaign in the U.S. a couple of years ago, generated some signals in the academic world suggesting that Barack Obama was a strong advocate of evidence-based science policy (Bhattacharjee 2008), meaning better "science for policy". Some of his statements, made even long before the campaign were contrasted to the practice of the Bush Administration that by comparison was seen to be ideological, particularly with regard to U.S. climate policy, but also in the field of health care (Oberlin BlogSpot 2008).

The Obama Administration's appointment of John Holdren as Science Adviser has confirmed the ambition of basing research policy more systematically on the accumulation of a viable knowledge base. Both the National Science Foundation (NSF) and the National Institutes of Health (NIH) have been given clear mandates to intensify efforts to develop relevant performance indicators that can be used in decision-making relating to both science for policy (innovation and health care, respectively) and policy for science (high quality curiosity oriented or fundamental research).

A new acronym for science based policy (SciSIP)

Some of the current efforts afoot in order to develop a robust indicator-based knowledge base for decision-making in science policy now fall under the heading of a new acronym, "SciSIP" (NSF 2007) which rejuvenates the crystallographer and Marxist John Desmond Bernal's old dream of a science of science to steer the growth of science (Bernal 1939; Elzinga 1988).

John Marburger III, science adviser to President Bush, expressed the initial rationale for such efforts some years ago. Marburger tied it back to strong demands of public accountability regarding tax dollars and the generally diffuse policy of enhancing economic competitiveness of the nation in a global market. In the wake of the launching of the President's American Competitive Initiative (ACI) Marburger (2006) proposed an agenda for a *Science of Science and Innovation Policy*, a characterization that became a new phrase in policy documents. He argued as follows:

Science policy makers tend to rely on economic models and data when they exist, but also employ ad hoc surveys and opinions offered by panels of experts. Science policy implementers are usually government employees and elected officials whose information comes from a variety of sources of varying degrees of visibility, with advocacy groups on the high end and science policy technocrats somewhere near the bottom. I would like to change this. I would like to have science policy tools that are so credible that their products are embraced by the advocates as well as the technocrats. I do not expect tools that approach the credibility of Newton's laws or quantum mechanics, but I believe we can move the standards of science policy making and implementation closer to what already exists in the world of economic policy.

And further:

I am emphasizing models because they are essential for understanding correlations among different measurable quantities, or metrics. (Cf. also OECD 2006 and Marburger 2007).

Under the Obama Administration the task of presidential science adviser has been mantled by a major league scientist, John Holdren, previously a Harvard professor of environmental policy and director of the Woods Hole Research Center. However this has not changed the basic rationale; if anything, it may have gained greater credibility, being tied in part to a different philosophy in the Whitehouse regarding the role of science in society.

Emphasis is still on measuring outcomes and impacts. The intention is also to link micro and macro data sets and make indicators more directly science policy relevant. This is also the ambition that originally prompted the NSF Science Metrics initiative.

The eventual aim is to create a cadre of scholars who can provide science policy makers with the kinds of data, analyses and advice that economists now provide to various government institutions. (Mervis 2006: 347 cited in Nowotny 2007:482).

On the basis of a wide call for proposals a variety of scholars two years ago were granted funding for projects under a program within the Directorate for Social, Behavioural & Economic Sciences (SBE). The three areas targeted are: analytical tools, model building, and data development & augmentation (NSF 2008, and 2007). By 2010 fifty projects located at various universities across the USA had been awarded grants in these areas and a new call for proposals was issued October 1, 2010.

The director of the SciSIP program, Julia Lane, in her statement before Congress in Washington September 23, 2010 summarized the rationale and approach further:

The program has three major aims: advancing evidence-based science and innovation policy decision making; developing and building a scientific community to study science and innovation policy; and developing new and improved datasets. The overarching goal in this effort, however, is to conduct basic research that creates new objective models, analytic tools, and datasets to inform our nation's public and private sectors about the processes through which investments in science and engineering research may be transformed into scientific, social and economic outcomes. (Lane 2010b; see also Lane 2010a).

STAR METRICS (Science and Technology in America's Reinvestment — Measuring the Effects of Research on Innovation, Competitiveness and Science) is the name of a broad program that is

already operational and gathering data. It represents a major step in the direction of generating indicators for a second order of evidence in policy-making, i.e., one that relates to societal impacts.

Too much “trust in numbers”?

The focus on bibliometric methods and other performance indicators of quality and relevance is also apparent in other countries. As so often is the case in science policy it is through the channels of the OECD that new approaches started in the USA are mimicked in other countries. Still, the approaches developed are different in different countries, again owing to differences in some of the legal, political and cultural factors mentioned before. There are also mixed approaches that combine traditional peer review panels with so-called “objective” methods of bibliometrics and other science indicators (scientometrics).

In the UK a system of review panels was for a long time used to evaluate and rank performance of universities and departments within them every five years. In December 2006 the government announced that a new system for the assessment and funding of research to replace that framework after the Research Assessment Exercise (RAE) of 2008 (on the principles underpinning the RAEs see Wright and Williams Ørberg 1997). The new framework as part of its approach was mandated to produce robust UK-wide indicators of research excellence for all disciplines. To some extent this is now in place. It represents a move away from the old “subjective” review panel approach to RAEs towards more “objective” methods based on publication counts and citation measures to gauge quality and impact, plus statistical counts of external research income and postgraduate student activity at universities.

The new framework operates with fewer and broader subject divisions than the RAEs. The full set of indicators for the science-based disciplines was produced for the first time during 2009 and has begun to influence funding allocations from 2010-11. The indicators are based on data averaged across several years. For the arts, humanities, social sciences, mathematics and statistics the operation is being phased in more gradually, initially complemented by information from peer review panels. This is because publication patterns in these areas do not match those of the science-based disciplines, as the relevant international databases on publications and citation frequencies do not give a representative picture. Additionally there has been quite a lot of critique from scholars in the humanities and social sciences as well as mathematics.

The idea nevertheless seems to be that reviews and summations of relevant performance indicators will fully inform funding from 2014 onward (HEFCE 2007). The aim is to try and enhance the overall relative international level of performance (in comparison with other countries) represented by the country’s research base. As Helga Nowotny has pointed out concerning the report that pushes in this direction,

Although the report does not carry the word ‘evidence’ in its title, it is yet another example of evidence-based policy intended to replace the RAE... By devising systems to compare ‘best practices’ at national, European and international levels, self-generating, performance-enhancing mechanisms are created. Their function is to orient towards goals and objectives that readily can generate ever new performance targets and changing objectives by absorbing existing performances... (Nowotny 2007:482).

Nowotny who has first hand experience of policy making and science advice at the highest level within the EU and is currently President of the European Research Council (established only a

TD, 6(2), December 2010, pp. 307-332.

few years ago to try and counterbalance the pull of applied research funding programs in the EU), for her own part, expresses scepticism and is critical of the science metrics approach. She warns against fastening in a reification of numbers and the associated myth of a “trust in numbers” on which it rests; therefore she calls for other, competing constructions of “policy rooms” distributed throughout the science and innovation systems. Perceptive users of bibliometrics and research performance indicators have also warned of inadvertent consequences inherent in too much trust in numbers (Weingart 2005).

Generating quantitative data: peer review based evaluations and foresight

A leading centre in Europe where bibliometric methods have been developed is the Centre for Science and Technology Studies (CWTS) at Leiden University, led by Anthony van Raan who is currently also the editor of the journal *Research Evaluation*. Another is the so-called Leuven-group, the Steunpunt O&O Statistieken (abbreviated SOOS), a Flemish inter-university consortium located at the University of Leuven and directed by Wolfgang Glänzel who also has longstanding affiliations with the Information Science and Scientometric Research Unit (ISSRU) at the Library of the Hungarian Academy of Sciences which is the co-publisher (with Springer Verlag) of the journal *Scientometrics*.

Bibliometrics has been found to be a useful tool for generating a quantitative core of data for peer review based evaluations and foresight exercises (Van Raan 1996). In other words the methodology is held to provide useful starting point if complemented by the activities of traditional peer review panels, hearings, on site visits and the like. Additionally, informed discussion (reflexivity) regarding measurement technologies and their use in science policy will benefit the process (Woolgar 1991). It must be remembered that bibliometric methods and citation analysis convey pictures of how visible various authors and their associated institutions are in the literature. Strictly speaking *visibility* is not the same as quality even though citation counts in journals with high impact factors are sometimes used as proxies (Cozzens 1989). Fields that are inherently transdisciplinary are often difficult to fit into the predominant categories that are used and therefore have a systemic disadvantage (Bondjers and others 2008).

The services of the CWTS unit at Leiden have frequently been used in bibliometric studies at Swedish universities. Medical faculties have also engaged Grant Lewison who has developed bibliometric competence at the Wellcome Institute in the UK. In Sweden Olle Persson and his Information Research Group (started in 1975) at the Sociology Department of Umeå University has been working with bibliometric methods since 1985, publishing empirical and theoretical papers and engaged operationally in various commissioned evaluation exercises. More recently Ulf Sandström who is affiliated with Linköping University has over the years developed various bibliometric skills in connection with research evaluations and policy, among other at the policy unit of the Swedish Research Council (VR), and as expert consultant to the Resources Inquiry that was led by Dan Brändström. The Resources Inquiry in its report *Resurser för kvalitet* (SOU 2007:81) proposed a new model for the allocation of funding through direct appropriations to Swedish universities. In January 2009 the Swedish Research Council (Vetenskapsrådet, VR) was given the mandate to begin this process of introducing a system for correlating allocation of funds to universities with their outputs measured in numbers of publications and amounts of external funding they attracted. Measurement of related citation frequencies is taken as a proxy for scientific quality, an assumption that has however been severely criticized by researchers in a number of different fields who point out that what one gets is only estimates of the “visibility” of various universities and departments within them.

Even before it was endorsed by a parliamentary decision the new approach already prompted a flurry of activities at universities to speed up work in developing their own capacities to do computer-aided evaluations of research performance. These activities continue now to satisfy national policy demands and they mostly engage bibliometricians connected to the university libraries. Leading universities like Uppsala and Lund have also implemented their own comprehensive evaluations and other universities are still in the process of doing so, oftentimes using the database of performance indicators that is generated also as an “evidential” base for controversial decisions regarding some major organizational changes that bring with them more power on the part of management and less in terms of traditional collegial patterns of decision-making.

Strict bibliometrics contested

In Scandinavia there has been discussion regarding what is called the “Norwegian model” for linking state decisions regarding budgetary allocations to university research to systematic reviews of performance. One of the architects behind the Norwegian model is Gunnar Siverstsen (NIFU/STEP in Oslo)² who has helped devise a system whereby researchers at universities and colleges report relevant information about their publishing activities into what has become a national database managed by bibliometricians employed at university libraries. The model is quantitative and bibliometric. An advantage is that since researchers report in their own data to the data base they develop a sense of “ownership” regarding the system and therefore may find it more palatable than the so-called “objective” system of relying only on international citation databases.

A scientific publication is defined by four criteria each of which has to be satisfied: 1) presentation of a new insight; 2) presented in a form that makes the results testable and possible to use in new research; 3) expressed in a language and via a channel of distribution that makes it accessible for most researchers who might take an interest in it; and 4) the publication channel (scientific journal, report series, book, etc.) that it appears in must incorporate peer review procedures (Sivertsen 2008).

Publications distributed through local channels (if more than 2/3 of publications in a publication series coming from the same institution) or non-scientific channels (lacking peer review) are excluded.

Publication channels are divided into two levels: Level 1: a category that covers “normal” quality, where one usually finds 80% of publications in a discipline; Level 2: a category that covers the other 20% where one finds the most significant or highest quality publications, e.g., high impact international scientific journals. In case of a multi-author article the publication is

² The acronym is Norwegian; the English name of the center is Norwegian Institute for Studies in Innovation, research and Education - <<http://www.nifustep.no/English/Pages/default.aspx>>.

divided into corresponding fractional parts attributed to the respective authors' home institutions. A point system is used to give weights to different kinds of publications, and publication points are calculated annually by multiplying the author-fraction affiliated with an institution times the appropriate vectors for publication form (species & levels).

In addition to publication points three other indicators are combined in the result-based decisions for redistributing funds amongst universities and colleges with an eye to their final quality measure expressed on a national scoreboard. The other three indicators are first the number of doctoral degrees, secondly EU-funding attracted, and thirdly Norwegian Research Council funding attracted over the same period. The evaluation/measurement exercise was incorporated into the state budget for universities 2005 and has since then been extended to include the national research institutes sector and health-related enterprises.

Swedish experience and the proposal of a more “objective” model

Interested parties in Denmark, Finland and Sweden have studied the Norwegian model as an alternative to “strict” reliance only on international Science Citation Index data when incorporating quantitative and bibliometric information as a basis for state budgetary allocations of funds to university and other forms of research in the public sector. As already indicated several Swedish universities have been developing bibliometric functions to provide university boards with instruments to monitor quantity and quality of publications.

The most comprehensive report until recently has been one from the University of Uppsala where consultants from the CWTS/Leiden were employed to do an extensive evaluation, *Quality and Renewal. KoF07 Report* (see Uppsala University home page). At the national level the White Paper called the Resource Inquiry (or *Brändströmska utredning*) on financing forms for universities' activities (SOU 2007:81) reflects a certain enchantment with the British experiences with the RAE-system and the discussions regarding its replacement, at least in part, by “robust” indicators.

This Swedish White Paper criticizes the Norwegian model for being too costly and cumbersome because it involves local university based staff to manage researchers self-reporting of publications into a national database that needs to be continually upgraded and validated. To circumvent this “subjective” element the “Swedish model” therefore uses a quasi-objective mode of measuring performance at an aggregate level for comparison of individual universities in this country. The idea is to only make use of existing information regarding publication counts, relative performance levels above or below a world norm calculated for a large number of different classes of journals registered and indexed in the Thomsen/ISI database Web of Science (WoS). The characteristics of the model are:

1. The numerical value the model pins on a university is obtained by focusing on publications in given disciplinary areas;
2. Calculating productivity by translating the actual number of publications to a virtual number of middling level of averagely productive researchers that for each disciplinary-specific area would be required to produce the same number of publications;
3. Calculating the citation-value by looking at the average number of citations received by the publications in question and dividing this by the expected compiled value (number of citations one would expect the corresponding number of middle-level averagely productive researchers to receive (field normalization));

4. Multiplying the productivity value with the citation-value for each disciplinary area for the various universities reviewed.

A matter of cutting costs

It has been argued (Sandström & Sandström 2007b) that the advantage with the “strict” model is that one does not need to collect raw data from the universities. One gets the raw data directly from the database of the ISI/WoS. This procedure, it is argued, is much less costly than having to rely on universities’ databases that require competent staff to provide and manage local inputs that are constantly upgraded and validated. The other advantage emphasized is that application of techniques to achieve field-normalized indicators allows comparisons to be made across different disciplinary areas like technological science, medicine, natural sciences, social sciences & humanities as well as between sub-classes within these broad areas, something the Norwegian model is purported unable to do.

The foregoing line of argument also appears in some work of the aforementioned bibliometrics group at the Australian National University (ARC Linkage Project 2005). In 2007 two of this team revisited and evaluated the political science portion of the 2001 UK/RAE. The outcome of the original review panel assessment (RAE 2001) was compared with the results of a new evaluation for the year 2001 carried out by the Australian bibliometricians using only quantitative indicators accessible in international databases.

The authors of the Australian report (Butler & McAllister 2007: 14-15) write:

Our findings presented here suggest, unequivocally, that a metrics-based model, using objective, transparent indicators drawn from a range of readily-available measures, will yield results which are very close to those of a peer-based evaluation model /using review panels/. Such a stronger reliance on quantitative indicators, of which bibliometrics is a central measure, will, most importantly, obviate the need for a large peer review committee and the consequent indirect biases that it introduces into the system. And not least, such an approach would help to reduce one of the most oft-quoted criticisms of the RAE, namely the cost in university resources and academic staff time.

In the Swedish case, in practice for policy purposes the “objective” model (above) is not employed to the full and some adjustments were arbitrarily introduced before it was put in place to govern the budgetary allocations for the period 2009-2012. One such adjustment in order to compensate for a systematic bias against humanities disciplines is that the number of publications in this area is multiplied by a factor of two. Also, for this period as a start only 10% of the block grants flowing from the Ministry of Education for research at individual universities is allocated on the basis of competition. The 10% incentive key furthermore is broken down into two parts, where one half (5% of the total allocation) goes up or down depending on a given university’s relative ability (*vis-à-vis* that of the country’s other universities) to attract research funding from external sources and the other half varies on the basis of the outcome of the bibliometric measure of publications counts (again relative to the performance of other universities on the same measure). Those that do well in the performativity assessments get more resources, those that do not do so well get less.

At the next level, within each university the leadership of that university can decide how to further decide about the distribution of its budget using a more radical distribution key for funding flows down to the level of the faculties where further performativity measures may be

employed before the funds actually reach individual institutions. Since the whole process is quite new it is too early yet to see the consequences of the new national policy.

Concluding remarks

The purpose of the present paper is to fill an important gap in science policy studies and the literature on the problem of the impact of New Public Management. As such the paper elaborates an important and novel angle in science policy analysis. Furthermore it summarizes at some length various characteristics associated with NPM, one might say its “anatomy” and dynamics, and explicates its nexus with methods of research evaluation now being introduced in academic institutions. Arguments for and against the application of “strict” bibliometric performativity assessments as a policy tool are reviewed. In line with the foregoing, and based on the paper’s comprehensive review of relevant literature, a number of significant policy conclusions may now be stated.

First of all we should not lose sight of significant differences between disciplines as well as differences between major areas of knowledge. While humanities, social sciences and some applied as well as interdisciplinary and clearly transdisciplinary areas (Bondjers and others 2008) are having difficulties to adjust, robust research institutions in most natural sciences and mainstream medical areas probably will have no difficulty under stricter accountingization regimes to turn new number-grinding exercises to their own advantage (Herbert and Müller-Hill 1997). It will become a natural ingredient in reputation management. Still, even in such instances diversity in academic science landscapes may turn out to be the loser.

Secondly, there is a paradox that emerges both in natural and social sciences. “Control of control” as Michael Power calls it tends to become a ritual when where trust gets lost. Institutions contracted by a principal to perform services go up a learning curve and develop skills in formatting reports on outcomes in terms meant to satisfy the principal. These function as signals to assure the principal that objectives and rules are being followed to such and such a degree. The suspicious principal may mistrust the reports and require further evidence to be reassured that controls are in place. This leads to the introduction of a second order control, one wherewith the task becomes one of verifying if a system of control is in place (supposedly internalized in the working environment of the agent). At this stage one is no longer concerned with the actual detail or content of the performance per se, but rather in the existence of second order routines as a proxy (see Table 2).

Table 2. The ritual of Audit Society (according to Michael Power)

| Control of control |
|--|
| Management organizes>> internal audit in order to satisfy>> external audit incorporated in.....>> state inspection |
| Response = Symbols of compliance (assurances)....>> sent back to state authority = system of |
| <<..... internalizing Delegated control that is then subjected to topdown control by the external controller (as proxy to make sure) |

Thirdly NPM as an instrument of *Audit Society* introduces changes in the social relations and cultures of major societal institutions that are evaluated. Academic institutions are no exception. Overtly there is a move toward a *culture of compliance* while underneath it all researchers still seek to maintain agendas of their own. The cynical view is the one expressed by Daniel Greenberg (2007a, also cf. 2007b) in his article in the journal *Science* where he suggests that the ultimate name for a typical university running with the trend of the times ought to be “The University of Avarice” (the article is illustrated with a nice cartoon). It is a question of “campus capitalism” as a handmaiden of industrial and financial capitalist agendas in society at large. External relevance pressures, marketization, commoditization of public goods and subsequent accountingization bring with them cultural changes in the hallowed halls of academe, to be sure. In sum:

- A culture of compliance
- A culture of profiling, corporate branding, trademarking and conscious reputational management and identity management

A fourth point is that apart from orchestrated self-regulation via the inducement of a compliance culture one also by extension gets practices of *reputation management* whereby institutions try to enhance their image. To do well in the universe of citation indicators, publication behaviour in research communities, it has been speculated, may also undergo change to adapt to new computer-aided models for redistributing funds between competing universities and departments within them (Debackere & Glänzel 2004: 273-274). Weingart (2005) speaks of “inadvertent consequences”. Bornmann (2010) has coined the phrase “mimicry in science” to draw attention to a significant factor that is already changing publication patterns. For some universities image may become more important than content.

Reputation in other words becomes a key asset on which providers trade. In our universities we see manifestations of reputation management in the practice of trade-marking and cultivating brand names (*branding*), and even in disputes and endless discussions about such a simply matter as to how the university’s traditional logotype should be modified to bring it in tune with the times. According to identity consultants even a logotype is much more than a pattern on a paper or letterhead, it is a symbol that embodies a metaphysical means to unify actors around a mission as well as attract customers.

A fifth conclusion, more generally, is that the new initiatives to develop performativity metrics in the spirit of NPM in the higher educational and research sector are indicative of changes wrought in science policy during the past couple of decades as we have moved out of Cold War regime to the *new policy paradigm* of privatized globalization. The new initiatives have undoubtedly been important for improving the evidence-base for policy-making but they should not simply be adopted uncritically and taken at face value.

Further conclusions that can be drawn from the foregoing review of the entry of NPM in discourses of governance or orchestration of academic research is that academic research tends to be the loser. *Bibliometric methods on their own are not trustworthy* and may induce unwanted skews into research landscapes; therefore they need to be combined with self-reporting academic databases, traditional peer review panels, on site visits and case studies that are undertaken independently and at arms-length from policy- and decision-making. Short term interests related

to tangible utility is furthermore detrimental to strategic needs of diversity in basic research and scientific knowledge production for the long-term future.

For this and other reasons it is evident that a greater amount of reflexivity needs to be brought into discussions about accountability and the role of NPM, i.e., there is a need for critical *metatheoretical reflection* on both NPM and the various metrics that have been and are still being constructed. This is particularly relevant considering the mismatch that exists, for example, *vis-à-vis* research fields in the humanities and also regarding transdisciplinary modes of research and publication patterns. Comparison of the role of NPM methods and pertinent accountability technologies associated with it in other policy domains may also prove useful.

Finally, traditional historical and newer types of oral history case studies and reviews that analyze longer term trends in various academic disciplines and specialities continue to be important sources of *tacit knowledge* that needs to be articulated. In this connexion the specialized field that goes under the heading of history, philosophy and social studies of science stands challenged to become more policy relevant.

Bibliography

- ARC Linkage Project (2005), *The Strategic Assessment of Research Performance Indicators – a Literature Review*. Canberra: The Australian National University, Research School of Social Sciences, Research Evaluation and Policy Project (REPP) Discussion Paper 05/1, 84 pp. <repp.anu.edu.au/Literature%20Review3.pdf>
- Almqvist, Roland (2006), *New Public Management – om konkurrens, kontrakt och kontroll* (Malmö: Liber).
- Barzelay, Michael (2001), *The New Public management. Improving research and policy dialogue* (Berkeley: University of California Press).
- Bernal, John D. (1939), *The Social Function of Science* (London: Routledge & Sons Ltd.).
- Bhattacharjee, Yudhijit (2008), “New Focus – Democrat: Barack Obama”, *Science*, Vol. 319, No. 5859, pp. 28-29.
- Bondjers, Göran, Harri Vaino and Dag Thelle, *The Karolinska School of Public Health – potential, visions, and implementation*. Stockholm: Karolinska Institutet March 2008).
- Bornmann, Lutz (2010), “Mimicry in science?”, *Scientometrics* 2 June 2010 (On line: <http://www.lutz-bornmann.de/icons/Mimicry.pdf>)
- Butler, Declan (2007), “Academics strike back at spurious rankings”, *Nature* Vol. 447, pp. 514-515 (31 May 2007).
- Butler, Linda and Ian McAllister, “Metrics or Peer Review? Evaluation of the 2001 UK Research Assessment Exercise in Political Science” http://repp.anu.edu.au/papers/2007_ukresearchassess.pdf.
- Christensen, Tom and Per Laegreid (eds.) (2005), *New Public Management. The transformation of Ideas and Practice* (Hampshire UK: Ashgate Publishers).
- Cope Stephen, Frank Leishman and Peter Starie (1997), “Globalisation, new public management and the enabling State. Futures of police management”, *International Journal of Public Sector Management*, Vol. 10, No. 6, pp. 444-460.

- Cozzens, Susan E. (1989), "What do citations count? The rhetoric-first model", *Scientometrics*, Vol. 15, Nos. 5-6, pp. 437-447.
- Debackere, Koenraad and Wolfgang Glänzel, "Using a bibliometric approach to support research policymaking: the case of the Flemish BOF-key", *Scientometrics*, Vol. 59, No. 2, pp. 253-276.
- Drechsler, Wolfgang (2005), "The Rise and Demise of the New Public Management", *Post-autistic Economic Review*, Issue no. 33, September 2005 (article 2), pp. 1-9.
- Elzinga, Aant "Science Policy in Sweden, sectorization and adjustment to crisis", *Research Policy*, Vol. 9, No. 2 (April 1980), pp. 116 – 146.
- Elzinga, Aant (1988), "Bernalism, Comintern and the Science of Science: Critical Movements Then and Now", in Jan Annerstedt & Andrew Jamison eds., *From Research Policy to Social Intelligence* (MacMillan Press, London 1988), pp. 87-113.
- Elzinga, Aant (2002), "Making STS relevant for Technology Policy", *Social Studies of Science* Vol. 34, No. 46, pp. 949-956.
- Elzinga, Aant (2004), "Models, metaphors and reifications in science policy discourse", *Science as Culture* Vol. 13, No. 2, pp. 105-121.
- Elzinga, Aant and Andrew Jamison: "Changing Policy Agendas in Science and Technology", in S. Jasanoff et al., (ed.): *Handbook of Science and Technology Studies*. Sage, London, 1995.
- Elzinga, Aant (1996), "Shaping Worldwide Consensus: the Orchestration of Global Change Research", in Elzinga, Aant & Catharina Landström (ed.), *Internationalism and Science*, (London and Los Angeles: Taylor Graham), pp. 223-255.
- Elzinga, Aant (2004), "The New Production of Reductionism in Models Relating to Research Policy", Symposium paper presented at Nobel Symp. at Sw. Royal Ac. of Sc., Nov 2002, in Karl Grandin, Nina Wormbs and Sven Widmalm, eds. *The Science-Industry Nexus. History, Policy, Implications* (Sagamore Beach, MA: Science History Publications/USA), pp. 277-304.
- Elzinga, Aant (2005), "Scientific Community: Development of Science Policy", In Sal Restivo (ed.), *Encyclopedia of Science, Technology and Society*. (Oxford: Oxford University Press).
- Elzinga, Aant (2008), "Participation", in Hirsch Hadorn, Gertude et al. (eds), *Handbook of Transdisciplinary Research*. (Bern & Zürich, Springer Verlag + Business Media B.V.), pp. 345-360.
- Etzkowitz, Henry and Loet Leydesdorff (2000), "The dynamics of innovation: from National Systems and "Mode 2" to a Triple Helix of university–industry–government relations", *Research Policy*, vol. 29, pp 109–123.
- Fulltext:
http://www.csi.ensmp.fr/WebCSI/4S/download_paper/download_paper.php?paper=elzinga.pdf
- Gibbons, Michael; Camille Limoges, Helga Nowotny, Simon Schwartzman, Peter Scott, & Martin Trow (1994). *The new production of knowledge: the dynamics of science and research in contemporary societies* (London: Sage).

- Gilbert, G.N. and S. Woolgar (1974), "Essay review of the quantitative study of science: an examination of the literature", *Science Studies*, Vol. 4, pp. 279-294.
- Godin, Benoit (2003), "Measuring science: is there 'Basic research' without statistics?", *Social Science Information*, Vol. 42, No. 1, pp. 57-90.
- Godin, Benoit (2005), *Measurement and Statistics on Science and Technology, 1920 to the Present*. London: Routledge.
- Godin, Benoit (2006), "The Linear Model of Innovation. The Historical Construction of an Analytical Framework", *Science, Technology and Human Values*, Vol. 31, pp. 639-667.
- Godin, Benoit (2007), *What's Science? Defining Science by Numbers 1920-2000*. Montreal: Project on the History and Sociology of S&T Statistics at the Université du Québec à Montréal. Working Paper No. 35.
- Gorraiz, Juan and others (2010), "esss 2010: A review of the inaugural European Summer School for Scientometrics in Berlin", *Scientometrics*, 18 September 2010 online.
- Greenberg, Daniel S. (2007a), "Science and Society: On the Road to Academic Greatness – A Parable", *Science* Vol. 317, Issue 5843 (7 September), pp. 1328-1329.
- Greenberg, Daniel S. (2007b), *Science for Sale. The Perils, Rewards and Delusions of Campus Capitalism* (Chicago: The University Press).
- HEFCE (2007), (http://www.hefce.ac.uk/pubs/circlets/2007/c106_07/).
- Herbertz, H. and B. Müller-Hill (1995), "Quality and efficiency of basic research in molecular biology: a bibliometric analysis of thirteen excellent research institutes", *Research Policy*, Vol. 24, pp. 959-979.
- Hood, Christopher (1991), "A Public Management for All Seasons?" *Public Administration*, Vol. 69, No. 1, pp. 3-19.
- Hood, Christopher (1994), *Explaining Economic Policy Reversals* (Buckingham: Open University Press).
- Hood, Christopher (1995), "The 'New Public Management' in the 1980s: Variations on a Theme," *Accounting, Organizations and Society*, Vol. 20, Nos. 2/3, 93-109.
- Kilby, Partick (2004), "Is empowerment possible under a New Public Management environment? Some lessons from India?" *International Public Management Journal*, Vol. 7, No. 2, pp. 207-225.
- Lane, Jan-Erik (1994), "Will public management drive out public administration?" *Asian Journal of Public Administration*, Vol. 16, No. 2, pp. 139-151.
- Lane, Jan-Erik (1995) *The Public Sector: Concepts, Models and Approaches* (London: Sage 2nd rev. edn.).
- Lane, Jan-Erik (2000), *New Public Management* (London: Routledge).
- Lane, Julia (2010a), "Let's make metrics more scientific", *Nature*, vol. 464 (25 March 2010), pp. 488-489).
- Lane, Julia (2010b), Statement in Testimony before the U.S. Congress September 23, 2010: http://www.nsf.gov/about/congress/111/jil_policyscience_092310.jsp
- Marburger, John H. III (March 29, 2006):< <http://commerce.senate.gov/pdf/marburger-032906.pdf>>

- Marburger, John H. III (April, 19 2007): <http://commerce.senate.gov/public/_files/Testimony_JohnMarburger_OSTP_FY08JHMSenateCommerceTestimony_FINAL.pdf>
- Martin, B.R and J. Irvin (1983), "Assessing basic research. Some partial indicators of scientific progress in radio astronomy", *Research Policy*, Vol. 12, pp. 61-80.
- McCann, Eugene J. (2001), "Collaborative visioning or urban planning as therapy? The politics of public-private policy making", *Professional Geographer*, Vol. 53, No. 2, pp. 207-218.
- Mervis, Jeffrey (2006), "NSF begins a push to measure societal impacts of research", *Science*, Vol. 312, No. 5772, p. 347.
- Mirowski, Philip and Esther-Mirjam Sent (2008), "The Commercialization of Science and the Response of STS", in Edward J. Hackett et al. (eds), *The Handbook of Science and Technology Studies*. Third Edition (Cambridge Ma.: The MIT Press), 635-689.
- Nowotny, Helga (2007), "How many policy rooms are there? Evidence-based and other kinds of science policies", *Science, Technology and Human Values*, Vol. 32, No. 4, pp. 479-490.
- NSF (2008), (www.nsf.gov/pubs/2008/nsf08520/nsf08520.htm/)
- NSF (2007), Science of Science and Innovation Policy FY08 <<http://www.nsf.gov/pubs/2008/nsf08520/nsf08520.htm>>.
- Oberlin BlogSpot (2008) "Barack Obama: Pushes evidence-based science policy". <<http://oberlinsciencelibrary.blogspot.com/2008/01>>
- OECD Blue Sky II (2006): <<http://www.oecd.org/dataoecd/48/14/37483994.pdf>>
- Osborne, David and Ted Gaebler (1992), *Inventing Government* (Reading: Addison-Wesley Publ. Co.).
- Pielke, Roger A. Jr. (2007), *The Honest Broker. Making Sense of Science in Policy and Politics* (Cambridge: University of Cambridge Press).
- Power, Michael and Richard Laughlin (1992), "Critical theory and accounting", in M. Alvesson & H. Wilmott (eds), *Critical Management Studies* (London: Sage).
- Power, Michael (1997), "From Risk Society to Audit Society", *Soziala Systeme*, Vol. 3. H. 1, pp. 3-31.
- Ravetz, Jerome (1999), "What is post-normal science?", *Futures*, Vol. 31, pp. 647-654.
- Ryan, N (2001), "Reconstructing Citizens as Consumers: Implications for New Modes of Governance", *Australian Journal of Public Administration*, Vol. 60, No. 3, pp. 104-109.
- Sandström, Ulf and Erik Sandström (2007a), "A metric for academic performance applied to Australian universities 2001-2004." <<http://forskningspolitik.se/DataFile.asp?FileID=137>>
- Schimank, Uwe (2005), "'New Public Management' and the Academic Profession: Reflecting the German Situation", *Minerva* (2005) 43: 361-376.
- Shore, Cris and Susan Wright (1997), *Anthropology of Policy: Critical perspectives in governance and power* (London/(New York: Routledge).
- Shore, Chris and Susan Wright, (1999) "Audit Culture and Anthropology: Neo-Liberalism in British Higher Education", *Journal of the Royal Anthropological Institute*, Vo. 5, No. 4, pp. 557-575.

- Sivertsen, Gunnar (2008), "Den norske modellen", *Forskningspolitikk*, No. 1/2008, pp. 14-15.
- Slaughter, Sheila and Larry L. Leslie (1997), *Academic Capitalism: Politics, Policies, and the Entrepreneurial University* (Baltimore: Johns Hopkins University Press).
- Stokes, Donald E. (1997), *Pasteur's Quadrant: Basic Science and Technological Innovation*. (Washington: DC: Brookings Institution Press, 1997).
- Van Dijk, Meine Pieter, *Urban Management and Institutional Change. An integrated approach to achieving ecological cities*. Institute for Housing and Urban Development Studies, Rotterdam. IHS Working Paper 16/2008.
- Van Raan, A. F. J. (1996), "Advanced bibliometrics as quantitative core of peer review based evaluation and foresight exercises", *Scientometrics*, Vol. 36, pp. 397-420.
- Weale, Albert (2009), "Metrics versus peer review?", *Political Studies Review*, Vol. 7, Issue 1, pp. 39-49.
- Weingart, Peter (2005), "Impact of bibliometrics upon the science system: inadvertent consequences?", *Scientometrics*, Vol. 18, No. 14, pp. 117-131.
- Woolgar, Steve (1991), "Beyond the citation debate: towards a sociology of measurement technologies and their use in science policy", *Science and Public Policy*, Vol. 18, pp. 319-326.
- Wright, Susan and Jakob Williams Ørberg (2000), "Autonomy and control: Danish university reform in the context of modern governance", *Learning and Teaching*, Vol. 1, Issue 1, pp. 27-57.
- Ziman, J. M. (1994), *Prometheus Bound: Science in a Dynamic Steady State*. Cambridge: Univ. Press.