CHAPTER 1

RICHARD WHITLEY

CHANGING GOVERNANCE OF THE PUBLIC SCIENCES

The Consequences of Establishing Research Evaluation Systems for Knowledge Production in Different Countries and Scientific Fields

INTRODUCTION

Many recent discussions of state science and technology policies in OECD economies since the end of the Second World War have distinguished a number of distant phases or 'paradigms' (Ruivo 1994) in state-science relations, which reflected both different perceptions of the role of scientific research in industrialised societies and changes in the size and complexity of the public science systems (see e.g. Brooks 1990; Freeman and Soete 1997: 374-395; Martin 2003). From viewing academic and other research on fundamental processes as producing knowledge that would contribute to the achievement of economic and social goals in a predominantly linear fashion, many political and bureaucratic elites have come to see the connections between scientific research and public purposes as being more complex and interactive, and to believe that realising social and economic benefits involves a more active steering of research than was commonly thought in the 1950s and 1960s (Braun 2003).

Increasingly, research conducted in universities and similar organisations concerned with understanding basic processes in the material, life and social worlds is being viewed by politicians, business elites and other influential actors as a strategically important activity for national and regional development that can and should be managed for public policy purposes. Not only have the public sciences, i.e. those reputationally governed work organisations producing knowledge for publication in peer reviewed journals and books (Dasgupta and David 1994; Whitley 2000), become too large and expensive to be left to scientific elites to manage, but their outputs have become strategic assets in the international competition to dominate technological development in new industries and ensure continued economic growth. As such, their development, organisation and governance are now

R. Whitley and J. Gläser (eds.), The Changing Governance of the Sciences, 3–27.
© Springer Science+Business Media B.V. 2007

the focus of state policies and procedures in most OECD countries, with many states establishing ministries of science and technology and specific agencies and programmes to steer research in desired directions (Drori et al. 2003).

This steering often involves the incorporation of public policy goals and evaluation standards into the selection and monitoring processes of research funding agencies, such that project proposals have to show how the research contribute to these objectives if they were to be funded (Braun 1998). Problem focused or 'strategic' funding agencies, such as the British Medical Research Council and the US National Institutes of Health, are often charged with supporting the best research that will contribute to their overall strategic objectives, such as solving specific medical problems, so that researchers are encouraged to study phenomena that will generate useful results. Much of this work falls into what Stokes (1997: 71-74) has termed 'Pasteur's quadrant', i.e. research seeking to understand the basic mechanisms producing particular effects that can be used to deal with practical problems, such as, in the case of Pasteur, spoilage in fermentation processes.

The role of these kinds of agencies, and similar attempts to steer current research goals and careers towards public policy uses, has become more significant as the provision of state funding for research stabilises or declines relative to the number of researchers needing external resources to conduct projects, and alternative sources are few in number and/or limited in their resources (Braun, 1998). Most of these attempts have focused on the conditional selection and monitoring of research projects according to their likely contribution to strategic objectives as claimed in project proposals. They have relied on the prospective evaluation of new proposals and their applicants to guide research in particular directions, as well as maintain high quality through competitive bidding for resources allocated by variously organised peer review procedures.

In most countries, state investment in this prospective evaluation of applications for funds and guiding resource allocation procedures towards the achievement of public policy goals has been considerably greater than that provided for the retrospective evaluation of the results. As Cozzens (1990: 283) puts it:

> The US funding system places overwhelming emphasis on pre-performance evaluation (in the form of the peer review system) and gives minimal attention to programme evaluation. Thus major US funding agencies devote most of their efforts to giving out grants, not to finding out what happened once they were made.

However, as Cozzens presciently suggested in the same article (ibid.: 282): "Institutions and nations that are not currently evaluating, regularly and with all the skill they can muster, will probably be doing so ten years hence", partly to ensure the high quality of research outputs, and partly to satisfy potential and actual demands for increased political accountability of the greatly enlarged public science system (see also Cozzens' and Kneller's chapters in this volume).

As developments in the 1990s and the contributions to this volume attest, retrospective evaluation systems for assessing the quality of research conducted in universities and similar research organisations have indeed proliferated across many OECD economies, such that the *ex ante* evaluation of proposals for external project

funding in many sciences and countries is increasingly being complemented by the *ex post* evaluation of project results and publications through various national research evaluation systems.

Not only have many states shifted support for basic research away from block grants to universities to research councils and specialised funding agencies more amenable to state guidance (Block 1990), thus facilitating the short-term selection of research priorities through project funding, but they are increasingly evaluating research outputs in terms of qualitative and quantitative criteria, sometimes in order to reallocate funds to 'excellent' research organisations and away from less competitive ones. More generally, the previous direct institutional control of universities as state organisations that was prevalent in Japan and many continental European countries is being widely replaced by various forms of administrative decentralisation and increased reliance on output evaluations, as indicated by many contributions to this volume.

These sorts of changes in state-science relations have, though, occurred at different rates and in different ways in different countries. Block (1990: 40), for instance, shows very clearly how research and research students in US universities have been much more dependent on external project based funding than most OECD countries in the 1970s and 1980s. Despite the broad tendency in many societies to make universities and other research organisations formally separate from the state, major divergences in the funding, organisation and control of universities remain, as the chapters by Gläser and Laudel, Kneller, and Engwall and Nybom in this volume demonstrate (see also Campbell 2003; Gaughan and Robin 2004; Liefner 2003).

Furthermore, the nature and use of state evaluation systems varies considerably between OECD countries, as do the critical characteristics of public science systems that both structure the kinds of evaluation systems that are developed and mediate their consequences. Particularly important in this respect are the nature of the funding regime for academic research, the strategic capabilities of research organisations, the structure of scientific elites, and the segmentation of universities, and research institutes and of scientists' labour markets (Whitley 2003a). Consequently, we would expect the effects of these attempts by state agencies to steer, monitor and evaluate scientific research on the organisation of knowledge production, and on how new formal knowledge differs and is used, to vary considerably between differently structured public research systems in path dependent ways.

The institutionalisation of different kinds of research evaluation systems in many national academic systems is, then, an important component of more general changes in the governance of the public sciences that have developed over the past two decades or so. As the contributors to this volume make clear, we cannot adequately understand how and why these changes occurred, or identify their consequences for the organisation and direction of knowledge production, without locating them in their specific societal contexts, particularly the changing nature of political and economic institutions and interconnections.

It is additionally important to consider how the significant differences in the intellectual and social organisation of scientific fields (Whitley 2000) are likely to affect, and be affected by, the implementation of such evaluation systems and

broader changes in state-science relationships. Accordingly, as a contribution to the comparative analysis of the role of research evaluation systems in the changing governance of the public sciences, and their consequences for knowledge production, in this paper I suggest a framework for understanding how both variations in the nature of public research systems, and differences between the sciences, can be expected to influence the development and impact of different kinds of retrospective evaluation systems.

I first discuss the major characteristics of these evaluation systems and how they differ between countries and over time, and then consider the sorts of broad effects they are likely to have on research strategies and innovation. To simplify the analysis, it is useful to combine many of the more consequential aspects into an idealised 'strong' form of evaluation system in order to highlight the sorts of impact that they seem likely to have. The next section of the paper examines the key features of public research systems that are likely to mediate these consequences in particular countries and outlines how they should do so. Finally, I discuss the key differences between scientific fields that can be expected to influence the outcome of introducing such strong evaluation systems on particular sciences.

CHARACTERISTICS OF RESEARCH EVALUATION SYSTEMS AND THEIR CONSEQUENCES

Research evaluation systems (henceforth RES) are organised sets of procedures for assessing the merits of research undertaken in publicly-funded organisations that are implemented on a regular basis, usually by state or state-delegated agencies. As the chapters in this book exemplify, their nature varies considerably between countries, and also changes over time. In particular, they differ in how they are organised and governed, and in their implications for resource allocation decisions. Considering first their governance and structure, we can distinguish between evaluation systems in terms of their frequency, formalisation, standardisation and transparency. For example, the British Research Assessment Exercise (RAE) has varied in its periodicity from three to seven years, and the Australian formula-based measurement procedure occurred every year. Similarly, the extent to which evaluations are conducted according to formally specified procedures differs considerably between the many carried out informally at the invitation of individual universities or departments on a largely ad hoc basis and those organised by central bureaucracies with systematic rules.

Standardisation here refers to the use of common evaluation procedures and practices across the sciences and over time. This varies considerably between Australia, the Netherlands and Spain as well as between other countries, as the chapters by Gläser and Laudel, van der Meulen, and Cruz-Castro and Sanz-Menendez demonstrate. It is worth emphasising that while these may be formally prescribed and intended to function in the same way between fields in some evaluation systems, actual practices may in fact vary considerably between review panels, just as journal refereeing and grant allocation practices can differ between fields and reviewers, as shown by Langfeldt (2001). Such variations seem

especially likely when peer review judgements of quality are being made in different kinds of science, as evident in the British RAE's struggles to combine formal standardisation of procedures and consequences with appropriate and meaningful judgements of research quality in very contrasting fields of enquiry.

Research evaluations also vary greatly in their public transparency. In highly transparent systems, evaluations are carried out by panels that have been appointed according to formal, published procedures and base their judgements on publicly specified criteria and ways of working. Usually, these are published and form the basis of public rankings of the relative performance of individual departments and universities. In other systems, judgements are made by small groups of colleagues appointed informally, who decide their own working procedures and report their results to universities and, sometimes, state agencies, in confidence.

These characteristics tend to occur together. Evaluations of research achievements in many Continental European countries, for instance, have been organised relatively informally on a disciplinary basis, with neither the processes involved nor results being published, although some of these have become more systematically organised around standard protocols as in the Netherlands (van der Meulen in this volume). Many also have been primarily intended to improve research capabilities, and so can be described as 'formative', as in Schiene and Schimank's account of evaluation in Lower Saxony in this volume, rather than arrive at summated judgements that rank the quality of individual departments.

In contrast, other countries have developed much more frequent, formal, standardised and transparent RES that have been managed by funding agencies or ministries who have published their results. The British RAE has probably become the most well known of these, but other countries have also begun to introduce similar schemes on a regular basis. These kinds of evaluations assess the relative quality of research outputs from particular university departments or units of assessment over the previous four or five years, and are intended to rank them on a standard scale that is publicly available. Judging panels are selected in highly formal ways, often with continuing membership, and have to follow systematic procedures that are published and standardised across the sciences and universities. The governance of research evaluations here is nationally centralised and their organisation is highly rule governed, standardised, relatively permanent and publicly transparent.

In the case of the UK, the RAE has become progressively more formalised with written procedures being widely discussed and publicised, and outcomes analysed in the press and elsewhere in terms of prestige rankings. Not least because of legal challenges and threats of such actions, panel procedures and judgemental criteria are formally prescribed and publicly justified. As a result, the RAE has become a very important process for researchers and their employers in terms of their public reputations, as well as financially.

Another aspect of the organisation of RES that can affect their impact on knowledge production concerns the ways that the units of assessment are constructed. As van der Meulen and Schiene and Schimank emphasise in their chapters in this book, this can vary considerably between funding regimes and academic systems with significant consequences for the coordination of research in

universities. In the Netherlands the primary unit of evaluation is the research programme, and universities have had to construct these organisational entities in order to gain research funding. In Lower Saxony, the individual Chair remains the key unit of research assessment, but evaluating teams have strongly encouraged collaboration between professors and the formation of research teams that could tackle 'big' problems, thus potentially generating more collective organisational research capabilities.

Elsewhere, they may either cover entire disciplines such as History or Philosophy in ways that are isomorphic with traditional university departmental boundaries; focus on more specialised fields such as development studies; or cover wide ranging areas of concern such as environmental studies. How such boundaries are drawn, and by whom they are implemented, affect the ways that judgements of intellectual significance are made, especially in fields that cross traditional academic boundaries such as economic history, innovation studies and mathematical finance.

Finally, a major difference between RES concerns their consequences for the funding of research in the public sciences, both their direct impact on resource allocation and also in the proportion of employers' incomes that is affected by them. In quite a number of states, especially where they directly control resource allocation to individual universities and departments as in many continental European research systems, evaluations have not so far been directly linked to funding decisions (Campbell 2003 and the contributions of Cruz-Castro and Sanz-Menendez, Engwall and Nybom, and van der Meulen to this volume). In contrast, in Australia and the UK they have had significant direct effects on resource allocation outcomes, albeit to different degrees across the sciences and between individual research organisations, as Gläser and Laudel show in their account of the Australian RES.

Various discussions have suggested that implementing RES in the public services can be expected to have a number of effects on researchers and their employers (See Martin and Geuna 2003; Tabil 2001). These include:

a) heightened awareness of the significance of public reputations for high quality research, especially internationally, and hence;
b) greater investment in publicising research achievements nationally and internationally and gaining scientific recognition; and
c) increased competition for publication in leading journals and for gaining project funding for research on currently significant topics.

Systematic and transparent RES are also likely to increase the visible stratification of researchers, universities and journals and encourage the standardisation of research and publication styles across the sciences and across local and national scientific schools. In some cases they may additionally discourage scientists from investing in long term, highly risky and interdisciplinary research, and from challenging current orthodoxies, as Schiene and Schimank suggest could happen in Lower Saxony if professors are forced to collaborate in research teams.

Such consequences seem especially likely when evaluations are centrally organised with rankings published in the press, conducted regularly, and have direct significant consequences for the allocation of public funds to research organisations on the basis of the quality of their research. In considering how different kinds of

RES are likely to affect the organisation and conduct of research in different public science systems and fields, it is useful to contrast two ideal types of weak and strong evaluation systems.

Weak RES can be identified in the following terms. They are typically organised informally by funding agencies and/or a consortia of universities with little standardisation of procedures or criteria. They rarely publish their conclusions, which are usually intended to encourage organisational improvements, and so be formative, rather than to arrive at summative judgements. Whilst these assessments focus on the overall effectiveness and quality of research groups and departments, they do not rank them according to an overall scale of international excellence, nor do they have direct financial consequences, although they may have some indirect ones in the medium term. While such RES may encourage scientists and universities to consider their research strategies more explicitly and to focus more on producing research that is recognised internationally as significant, their impact on current arrangements is likely to be incremental rather than radical, although they may lead to some shifts in the organisation of research, as perhaps is the case in Lower Saxony (Schiene and Schimank in this volume).

Strong RES, in contrast, institutionalise public assessments of the quality of the research conducted in individual departments and universities by scientific elites on a regular basis according to highly formalised rules and procedures. These assessments are usually ranked on a standard scale and published so that the relative standing of universities and departments can be readily ascertained. In most cases, they are organised around existing disciplines and scientific boundaries. Such peer-review-based evaluations directly affect funding decisions, often on a significant proportion of research organisations' income, and so can have a considerable impact on the management of universities and similar organisations.

The impact of developing and implementing research evaluations on knowledge production is likely to be especially noticeable when these are relatively 'strong' in this sense. Five major consequences of institutionalising such systems can be summarised in the following terms.

Firstly, by focusing attention on evaluations of the outputs of their work, researchers are likely to become more aware of the need to compete with others to gain recognition from scientific elites and coordinate their projects with those of others. This means that they will seek to contribute to the collective goals of their field as understood by current elites and so research in general should become more integrated around these goals as evaluation systems become more influential.

Secondly, as evaluators in these peer-review-based RES are forced to judge the relative merits of research outputs, they will develop and apply standard criteria of quality and intellectual significance for the field as a whole, thereby centralising judgements across individual researchers, universities and other research organisations. As they continue to do this on a regular basis, these standards and goals will become institutionalised as dominant in the field, and so the level of strategic task uncertainty, i.e. the degree of uncertainty about the intellectual importance of particular research strategies and outputs for collective goals (Whitley 2000: 123-124), should decline.

Thirdly, this centralisation and standardisation of research goals and evaluation criteria throughout scientific fields means that the diversity of intellectual goals and approaches within sciences should decline over time, especially where they challenge current orthodoxies. As evaluations become more important for both researchers and their employers, the costs of pursuing deviant strategies increase, and pressures to demonstrate how one's work contributes to dominant disciplinary goals will grow. These are especially strong for junior researchers who need to show the merits of their research as assessed by current disciplinary priorities and standards in order to gain employment and promotion.

Fourthly, such reinforcement of disciplinary standards and objectives is likely to inhibit the development of new fields and goals that transcend current intellectual and organisational boundaries by increasing the risks of investing in research projects that do not fit within them. Increasing competition for reputations and resources based on them, resulting from strong evaluation systems heighten the risks of moving into novel areas and adopting techniques and frameworks from other fields. Intellectual innovations will therefore tend to be focused on current sciences and their concerns. Radical intellectual and organisational innovation is thus less likely in societies that have strong, institutionalised research evaluation systems because these reinforce conservative tendencies in determining intellectual quality and significance.

Finally, the standardisation, formalisation and publication of quality rankings intensify the stratification of individual researchers, research teams and employer organisations. By regularly conducting and publicising such judgements, strong evaluation systems heighten awareness of one's relative position in the science system and encourage both individual and organisational strategies to enhance them. This stimulates the scientific labour market and, over time, is likely to concentrate resources and the most valued skills in elite universities, as the UK RAE seems to have done. However, such effects will differ between academic systems organised in different ways as well as between different kinds of scientific fields. I now turn to consider how the consequences of different evaluation systems seem likely to vary between differently funded and organised public science systems.

THE CONSEQUENCES OF STRONG RESEARCH EVALUATION SYSTEMS IN DIFFERENT PUBLIC SCIENCE SYSTEMS

National public science systems do, of course, continue to differ in a wide variety of ways but there are six major features that can be expected to affect the operation of research evaluation systems and their impact on the organisation of research and knowledge production.

Firstly, despite a general tendency for states to increase the conditionality of their research funding by making more resources dependent on competitive bidding for research projects, one of the most important ways in which academic systems vary remains the extent and periodicity of discretion granted to scientists and their employing organisations. In particular, the more resources are allocated as block grants to research organisations, usually on the basis of existing commitments to

departments and laboratories but sometimes according to the number of students or, later, graduates (Hansen and Borum 1999; Liefner 2003), the more discretion researchers have over the selection of research topics and approaches and the more time they have to pursue ambitious intellectual goals. In block grant funding regimes, then, the ability to continue to conduct research does not greatly depend on scientists' performance in the short to medium term.

In contrast, highly project-based funding regimes, where scientists have to gain new funds for each new project from external sources (usually state agencies) on a competitive basis, shorten the feedback loop between research performance and resource allocation and increase the costs of project failure. As many states have reduced the proportion of funds for research in the public sciences allocated as block grants, in favour of external project funding, they have increased the importance of performance-based funding and reduced the time horizons of researchers' discretion, as well as facilitating direct state guidance of research goals through making funding dependent on projects' contributions to public policy goals (Braun 2003). This feature of research funding regimes can be termed the frequency and significance of project performance review.

Secondly, another important feature of funding regimes that has become more significant as states seek to use scientific research to deal with public policy issues and improve the competitiveness of national economics, is the incorporation of state and other public policy objectives - such as the development of new technologies and industries - into funding procedures (Ruivo 1994; Braun 2003). Whether this is attempted through formal contracts, through the use of public problem solving criteria as additional selection filters for evaluating competent project proposals, or by involving user groups in the development of research programmes, it encourages the institutionalisation of multiple and often different research goals and standards for allocating funds that may legitimate and support use-oriented research careers and novel areas for study.

While such incorporation of public policy goals into project funding procedures can result more in ritualistic invocation of their desirability than substantive redirection of research activities towards problem solving goals – especially where cohesive disciplinary elites are able to control standards of research competence and significance – it can also limit the concentration of control over intellectual objectives and approaches exercised by such elites and lead to increased pluralism within cognate sciences. This is more likely when overall research funding in the public sciences is stable or declining in real terms and the proportion allocated for problem or use-oriented research is substantial, as Braun (2003) suggests. Overall, we can term this feature of funding regimes the importance of varied public policy goals in funding decisions, as distinct from those focused on intellectual purposes.

A third important feature of research funding systems concerns the variety of funding agencies and of their objectives. As Gläser and Laudel make clear in their contribution to this volume, a significant feature of the Australian public science system is the high level of dependence of scientists on a single agency, the Australian Research Council. In Britain by contrast, as well as in Sweden and other European states, there are often a number of different kinds of organisations providing funds for a variety of scientific research with diverse purposes, so that

scientists have more choice in raising money for their work. In the USA, of course, the variety of funding sources is even greater, including between different federal and local state agencies (Stokes, 1997). In general, the more varied the goals of different funding organisations with significant resources at their disposal, the less dependent researchers are likely to be on particular groups of disciplinary colleagues and agency staff, and so are the more able to pursue diverse intellectual purposes with novel research approaches. This feature can be described as the degree of external research funding diversity.

The effects of these three features of national funding regimes on the organisation of knowledge production are complicated by a fourth major difference in the organisation of public research systems: the organisational independence and capabilities of universities and other research organisations. While many states that organised universities as part of the central or local bureaucracies are now moving to a more decentralised system, their autonomy from central direction still differs considerably between, say, the UK and Spain. This means that universities' ability to exercise strategic independence and develop distinctive organisational capabilities in allocating resources, monitoring performance and pursuing distinctive purposes remains nationally variable, as is emphasised by Weingart and Maasen in their discussion of the 'enterprising' university in this volume.

In states where scientific staff and other employees are civil servants employed by national and regional governments on state-based terms and conditions, and the budgets of universities are largely determined centrally, their ability to develop specific competences through employment contracts in a comparable manner to private companies will be severely limited. In many ways, research organisations in much of continental Europe and Japan have, at least until very recently, been administrative shells in which most financial and managerial decisions are taken by ministries of education and science, while most decisions about research and teaching are made by senior academics, both individually and through faculty committees, with appointments and promotions being decided jointly (Clark 1995; Coleman 1999; Muller-Camen and Salzgeber 2005). As a result, the organisational autonomy, cohesion and distinctiveness of universities and other research organisations in such states remain relatively restricted, especially with respect to reallocating resources between fields, creating new departments and closing others down.

As Liefner (2003) suggests, the combination of general block grant funding of universities with high levels of state co-ordination of finance and administration has been associated with incremental organisational innovation and relatively slow rates of structural change. Once established in state budgets, department and institutes were largely left to pursue their own intellectual goals in many European universities, with little state or organisational steering or evaluation. This relatively high level of intellectual autonomy from bureaucratic and university pressures enjoyed by incumbent researchers – or at least by heads of research groups – was accompanied by limited flexibility in accommodating demands for new scientific fields and approaches, especially when funds became restricted. Because universities and similar organisations had limited resources that they could control

themselves, and often were constrained by incumbent professors, their ability to innovate organisationally in this situation was quite weak.

Where - in addition - authority structures within research organisations are highly stratified and administrative boundaries are coterminous with units of knowledge production – as in the German Institute System (Clark 1995; Muller-Camen and Salzgeber 2005) – researchers become highly constrained to follow the research goals and methods of their organisational superiors. Furthermore, if career mobility additionally occurs more within universities than between them, as in much of the post-war Japanese national academic system (Coleman 1999; Kneller in this volume; Sienko 1997), both intellectual and organisational innovation seems likely to be limited.

In other countries, notably the UK and USA, but also in Australia and Canada, universities have historically had much greater autonomy in allocating resources, even when the bulk of their resources have come from the state, and directly employed staff, sometimes on varied terms and conditions. The significance of university administrations has been correspondingly greater than in the previous situation, particularly in the USA, and each organisation has been more able to develop distinctive collective competences than in the more centralised academic systems. Internal resource allocation, appointments and promotions, opening and closing departments and similar decisions are here matters for individual research organisations and so they are able to cross subsidise research groups and respond to state policies in more autonomous and varied ways.

This feature of public science systems obviously depends on the overall nature of the state funding regime. Where, for example, university funds for research become increasingly dependent on external project based finance from one or two state agencies, and they have declining control over both the amount they receive for teaching and how it is allocated, their actual strategic autonomy will be limited even if they have considerable formal independence from the state. In Spain, for instance, Cruz-Castro and Sanz-Menendez, in their chapter, suggest that the inability of universities to fund research has made them incapable of acting strategically. Also, in Australia the considerable formal independence of universities has become greatly reduced in practice by both funding cuts and high dependence on project grants and evaluation-based funding, according to the chapter by Gläser and Laudel in this volume (see also Marginson and Considine 2000). It remains to be seen whether current changes in many European countries and Japan will similarly combine increased formal autonomy with considerable de facto dependence on state policies (Engwall and Nybom; Kneller, and Weingart and Maasen in this volume).

Two further features of public science systems that are likely to influence the consequences of introducing strong research evaluation systems are: a) the cohesion and prestige of scientific elites in each society – and general social standing of scientists and scientific knowledge – and, b) the segmentation of universities and other research organisations in terms of their funding, objectives and labour markets (Whitley 2003a). The first reflects the ability of scientific elites to organise themselves effectively as an interest group and to control the ways in which intellectual prestige and resources are allocated between institutions and fields. It is considerably assisted by a concentration of prestige and resources in a few elite

universities and other research organisations at the apex of a relatively stable intellectual and social hierarchy, such as that dominated by the universities of Cambridge and Oxford in the UK and those of Tokyo and Kyoto in Japan.

The organisational segmentation of research purposes, careers and labour markets in public science systems refers to the strength and stability of the organisational separation of research organisations dedicated to the pursuit of different kinds of knowledge production, such as theory driven knowledge, use-oriented research and technology transfer. In highly segmented research systems, scientists do different kinds of work, are evaluated on different criteria and have distinctively separate career paths. They are also discouraged from moving between research organisations and rarely compete in the same labour markets. As a result, relatively few are willing to leave the university sector to establish their own companies exploiting the knowledge produced by their research because of the difficulty of returning to academic posts in the event of their firm failing, especially in very high risk sub-sectors such as therapeutic biotechnology (Casper 2000; Casper and Whitley 2004).

In contrast, where organisational segmentation between universities, applied research organisations and private companies is lower, as in much of the USA, mobility tends to be greater. Additionally, US universities seem to be able to incorporate a greater variety of intellectual purposes and activities than do those in other countries, and historically they have been much more responsive to market opportunities (Clark 1995; Kenney 2000; Casper 2006; Mowery et al. 2004). On the whole, the weaker is such segmentation, the easier it should be to establish new scientific fields, especially those falling within 'Pasteur's quadrant', such as computer science, when the state or other organisations are willing to invest in them (Mowery 1999).

Variations in these six features of the organisation of public science systems are likely to affect the ways in which, and the degree to which, the implementation of strong research evaluation systems will affect knowledge production. Their reduction or reinforcement of the five main kinds of consequences outlined above are summarised in table 1 and will now be discussed in a little more detail.

Beginning with the frequency and importance of assessments of research goals and performance through competitive bidding for research project grants, the introduction of strong RES into academic systems that require researchers to compete intensively for external grants is likely to produce many of these consequences, especially where funding is concentrated in one or two agencies and project evaluations are organised around similar intellectual boundaries as those used by assessment exercises. The importance of peer review – and hence of elite conceptions of significant problems and appropriate ways of tackling them – will be greatly enhanced by such assessment systems in this situation. Competition for disciplinary and specialty reputations will be intense and scientists highly constrained to coordinate project goals and approaches with those of national and foreign colleagues. Research organisations (ROs) will also compete intensively for reputations and funds, and the stratification of researchers, departments and universities be further encouraged.

Table 1. Mediating influences of public science system characteristics on the consequences of strong RES

Characteristics of Public Science Systems	*Major Consequences of Strong RES*				
	Increasing Research Organisation Stratification	*Intensification of Reputational Competition and of Coordination of Research Goals*	*Strengthening of Central Disciplinary Standards and Priorities*	*Reduction of Intellectual Diversity and Pluralism*	*Increasing Constraints on Establishing New Fields and Approaches*
High level of project-based funding and frequency of performance review	Reinforced	Reinforced	Reinforced	Reinforced	Reinforced
High level of public policy and use-oriented programme funding	Reduced	Reduced	Reduced	Reduced	Reduced
Varied, diverse funding agencies and goals	Reduced	Reduced	Reduced	Reduced	Reduced
High level of research organisation strategic autonomy	Reduced in elite universities with control over their own resources	Reduced in elite universities with control over their own resources	Reduced in elite universities with control over their own resources	Depends on significance of RES funding	Reduced in elite universities with control over their own resources
Strong and cohesive scientific elite implementing RES	Reinforced	Reinforced	Reinforced	Reinforced	Reinforced
Segmentation of research organisation goals and careers	Reinforced in academic ROs	Reinforced in academic RO's	Reinforced in academic RO's	Reinforced in academic RO's	Reinforced in academic RO's

Such competition for favourable evaluations from scientific elites for both project funds and the outputs from research activities seem likely to limit the ability of universities and similar organisations to implement distinctive research strategies

that challenge current orthodoxies. The more dependent they become on both *ex ante* and *ex post* peer reviewed judgements, the less intellectual autonomy they are able to exercise and the more focused they will become on achieving highly favourable assessments, especially when rankings of individual departments and fields are published. While some cross subsidy of less central scientific fields and unorthodox approaches may be feasible for elite organisations in the short term, this seems unlikely to be tenable as a long-term commitment for most research organisations in this situation.

The mutual reinforcement of project based finding and strong retrospective evaluation systems can be expected, then, to limit university independence in pursuing deviant goals with unorthodox methods, and scientists will be discouraged from pursuing research careers that involve working on problems deemed uninteresting or developing theories that are regarded as wrong by the current scientific consensus, such as the theory of continental drift in the 1920s and 1930s (Hallam 1973). Even in public science systems that formally grant considerable autonomy to universities, then, this complementary impact of pre- and post-research evaluations would make it difficult for them to avoid implementing their results more or less directly and discourage the establishment of wholly new scientific fields that challenge dominant scientific ideals. While perhaps not completely realised in any funding regime as yet, this combination could become established in the UK if, as the Royal Society has advocated, all public research funds are allocated through the Research Councils and there is no block grant funding of academic research in UK universities.

Where researcher dependence on project based research funding is lower, the effects of introducing strong evaluation systems may be less marked, especially where universities have greater levels of organisational autonomy and distinctiveness and there are a variety of funding agencies. Where they can control employment policies and financial allocations, research organisations are, in principle, able to invest in different kinds of scientific fields and switch resources between topics in response to changing opportunities. They should therefore be more organisationally innovative than in centralised state co-ordinated academic systems, and more diverse in terms of the fields they research and the kinds of approaches used. Some intensification of reputational competition and of national and international coordination of research goals can be expected to result from the implementation of strong evaluation systems in these kinds of public science systems, thus reducing the potential diversity of research styles and length of projects, but this should be incremental rather than radically transforming.

Considering next the role of major state and/or other funding bodies' provision of project funds for strategic research programmes that contribute to public policy goals and problems, this seems likely to mitigate the centralising and standardising effects of strong evaluation systems, especially if resource allocation decisions involve non-scientific groups and criteria. Conflict between assessments based on the goals and standards of disciplinary elites and those tied more to the objectives of project funding agencies focused on public policy problems could result in increased differentiation within and across universities between groups and departments focused on primarily disciplinary research goals and those more concerned with

use-oriented ones. It could also reduce the attractiveness of research careers based on contributions to use-oriented research programmes.

While strong evaluation systems may well intensify reputational competition for contributions to disciplinary goals and encourage intellectual coordination of research goals and approaches in such contexts, this seems less likely where assessments are implemented by problem or use-oriented scientific groups as well as by disciplinary ones, and substantial project resources are available from agencies following public policy goals that cross disciplinary boundaries. Here, their impact should be correspondingly weakened, and research organisations more able to exercise strategic autonomy in developing distinctive research goals and allocating resources between groups and approaches.

More generally, the more diverse and varied are funding agencies' goals and associated criteria for assessing project proposals, the weaker we would expect the effects of implementing strong RES to be, except perhaps for their impact on universities' research management. This is because scientists' dependence on peer review judgements of the quality of their research will be limited by their access to different funding sources that judge the worth of intellectual contributions in contrasting ways. Thus, scientific elites will not be so able to dominate the allocation of resources by monopolising assessment criteria and procedures.

In particular, greater pluralism of research agency goals and evaluation standards should enable scientists to develop new fields and approaches, as well as careers based on their contributions to these, which may not be highly valued by current discipline-based elites controlling quality judgements. Depending on the balance of resource allocation between research assessment-based funding and that provided by agencies with diverse objectives and standards, researchers may be able to establish careers in more problem-oriented fields despite the development of strong, discipline-elite dominated evaluation systems.

As the previous discussion has emphasised, the actual strategic autonomy of universities and other research organisations depends greatly on how funding is organised, but it is worth considering how high levels of organisational autonomy might influence the impact of strong evaluation systems on knowledge production as a separate mediating factor. Where such independence reflects elite status and control of significant resources, through for example endowments and control over other kinds of resources that do not depend on state support or peer review, we might expect universities to pursue distinctive strategies and support more diverse intellectual goals and approaches than would be viable for lower status organisations.

While they may not always do so, such prestige and independent control of significant resources could enable them to invest in long term and risky projects, as well as recruiting and supporting scientists who do not wish to follow established approaches. At the very least, elite research organisations should be able to continue to employ staff researching unfashionable topics that are outside the intellectual mainstream of the field. They should additionally be in a position to support research that transcends established disciplinary boundaries and contributes to the development of new scientific fields. Thus, elite organisations that have some autonomy and independently controlled resources could be expected to limit many of the general consequences of introducing strong RES, although many may choose not to do so.

Considering next the cohesion and prestige of scientific elites, this affects the likelihood that they will be able to influence the organisation and direction of evaluation systems. Powerful elites are likely to design them to favour the best research in terms of their conceptions of quality and contribution to their central intellectual objectives. They can be expected to prefer evaluation systems based on peer review judgements since these enable disciplinary elites to set assessment standards. As a result, where elites are cohesive and able to influence the development and implementation of evaluation systems, we would expect such assessments to intensify further reputational competition and coordination of research goals around elite intellectual goals, as well as reinforcing current disciplinary and specialty boundaries.

This combination of powerful scientific elites and strong evaluation systems would inhibit the establishment of new scientific fields based on radically novel goals and approaches, as well as limiting the attractiveness of use-oriented research careers. Whatever the public genuflections to technological and public policy goals that sometimes accompany the implementation of research evaluation systems, strong scientific elites should, in practice, be able to establish their performance criteria as the dominant ones, as seems to have happened in the UK RAE and in the evaluation system developed in Lower Saxony.

When, in addition, universities and other research organisations are organised into a strong and stable hierarchy of intellectual and social prestige, such elite influence over the design and operation of research evaluation should reinforce their stratification, as leading scientists would tend to be based in the top organisations and quality based evaluation should encourage concentration of resources in them, because the elite established the standards being used. This is even more likely when research funds are concentrated in a small number of agencies that rely on peer review for allocating resources. In strongly hierarchical academic systems, this will reinforce existing distinctions and criteria for evaluating research outputs, as seems probable in Japan (Kneller in this volume). In contrast, where there is a plurality of different kinds of research support agencies pursuing diverse intellectual and policy objectives, research groups and universities will be more able to gain resources for different kinds of purposes and the stratification of research organisations may not be so stable, as is arguably the case in the postwar USA (Stokes 1997).

Turning next to consider the likely mediating effects of organisational segmentation of research purposes, careers and labour markets on the effects of strong evaluation systems, high levels of such segmentation can be expected to reinforce many of these in the major research universities. This is because the organisational and labour market separation of ‘pure’ research from more use-oriented work limits the feasibility of pursuing varied intellectual objectives and obtaining funding for diverse kinds of research within academic research organisations. Since strong research evaluation systems focusing on peer review assessments of quality are likely to be implemented by leading academic scientists in highly segmented public science systems, their standards and goals will dominate the evaluation system and so intensify reputational competition amongst university researchers.

Insofar as new scientific fields and approaches are encouraged by use-oriented research programmes, and multiple funding agencies pursuing varied objectives do exist in a public science system, then, high levels of organisational segmentation will tend to ensure that these pressures for greater diversity will be channelled into specialist research organisations rather than impinging greatly on the leading research universities, as perhaps has happened in postwar Germany. This suggests that establishing new research areas that cross disciplinary boundaries and involve the use of new kinds of intellectual approaches will be difficult in highly segmented public research systems, at least in the major universities.

In general, then, many of the expected consequences of implementing strong RES should be considerably reduced in public science systems that combine considerable diversity of research funding agencies and foundations with high levels of university autonomy and strategic capacity based on control of their own resources. Especially for elite research organisations in societies that accord them considerable social prestige and independent access to finance, the largely intellectually conservative implications of adopting strong evaluation systems may be restricted. Conversely, where research funding is available from only one or two state agencies and private funds are in short supply, and elite universities have limited social and political support, these consequences can be expected to be quite marked, as Gläser and Laudel suggest has been the case in Australia in their chapter in this book.

THE EFFECTS OF STRONG RESEARCH EVALUATION SYSTEMS ON DIFFERENT KINDS OF SCIENTIFIC FIELDS

Just as strong research evaluation systems are likely to reinforce many conservative tendencies in national public science systems, so too they can be expected to encourage greater elite coordination of research goals and standards within established scientific fields, especially where intellectual elites have to rank the quality of research outputs publicly. By institutionalising and standardising such judgements across researchers in universities and other research organisations, national research evaluation systems force leading scientists to set intellectual priorities for disciplines and research specialties that are likely to lead to increased integration of research strategies and projects within these fields. However, the intellectual and social organisation of different sciences varies considerably, and changes over time, in ways that can be expected to affect these and other consequences of implementing research evaluation systems (Whitley, 2000).

In particular, scientific fields differ in: a) the diversity of agencies and organisations funding research and of legitimate audiences for research results; b) the extent to which researchers are able to obtain externally provided project-based funding; c) the extent of local organisational variability of research goals and frameworks; d) the tendency of researchers to concentrate on a limited number of central problems and approaches as opposed to studying a wide range of different topics and problems; e) their centrality and prestige in the broad hierarchy of the sciences; and f) the extent of elite cohesion and consensus on intellectual goals,

research methods and the value of results. How these characteristics are likely to mediate the effects of strong RES, and their possible change as a result of implementing strong RES, are summarised in tables 2 and 3 and will now be further discussed.

Table 2. Likely Effects of Different Characteristics of Scientific Fields on the Consequences of Implementing Strong RES

Characteristics of Scientific Fields		*Mediation of Impact of Strong RES*
Diversity of Funding Sources and Legitimate Audiences for Research Reputations	High	Reducing
	Low	Increasing
Variability of Research Goals and Approaches	High	Reducing
	Low	Increasing
People to Problem Ratio	High (urban sciences)	Increasing
	Low (rural sciences)	Reducing
Elite Structure	Cohesive and agreed on core goals and approaches	Increasing
	Heterogeneous	Reducing
Centrality and Prestige of fields	High	Increasing
	Low (peripheral fields)	

Table 3. Likelihood of strong RES affecting characteristics of scientific fields

Characteristics of Scientific Fields		*Changes in characteristics due to strong RES*
Diversity of Funding Sources and Legitimate Audiences for Research Reputations	High	None
	Low	
Variability of Research Goals and Approaches	High	Diminishes because of increased coordination of research goals
	Low	Reinforced
People to Problem Ratio	High (urban sciences)	Reinforced
	Low (rural sciences)	Growing because of reduced reputations for work in marginal areas

Characteristics of Scientific Fields		*Changes in characteristics due to strong RES*
Elite Structure	Cohesive and agreed on core goals and approaches	Reinforced
	Heterogeneous	Increasing fragmentation of fields and control of resources
Centrality and Prestige of fields	High	Further increased because of strengthening stratification of fields; reinforce elite control and ideals of science
	Low (peripheral fields)	Further diminished because of strengthening stratification of fields; imitation of dominant research styles and norms

Considering first the research funding characteristics that also apply to public science systems as a whole, we would expect similar consequences for individual sciences as those for national research systems. Fields where researchers are able to obtain research funds from diverse organisations and have opportunities for seeking legitimate intellectual reputations for contributions to varied goals and audiences are less likely to be greatly affected by strong evaluation systems than are those where scientists are highly dependent on funding based on the judgements of such assessment systems. Thus, sciences where there is a diversity of goals and frameworks that are supported by different funding agencies and organisations will be less tightly coordinated around disciplinary objectives as a result of strong evaluation systems being implemented than will those with more limited sources of support. In particular, fields that are able to gain funding for use-oriented research, such as engineering, medicine and many of the policy and practice-oriented social sciences, should be less affected by the establishment of formal and consequential research evaluation systems than fields more dependent on a small number of state funding agencies.

Additionally, fields where scientists are able to gain most of the resources required for their work from external agencies such as national academies and research councils are less likely to be greatly affected by the implementation of strong RES than are those more dependent on RES-based assessments. While many of the referees and members of decision making committees in the former may well also be involved in the operation of RES, and so the criteria and standards of the latter will often reinforce those used in project selection, the impact of strong evaluation systems is likely to be incremental and limited since the proportion of total research support governed by them in this situation is relatively low.

Conversely, the impact of strong evaluation systems on scientists' research activities should be more marked in sciences where major resources – especially those funding the time of academics – depend on their judgements and where researchers have not hitherto been subject to regular assessment through peer review. Here, the effects of introducing systematic evaluations of research quality that affect universities' budgets directly are likely to be much more visible, resulting in qualitative increases in reputational competition and in the level of intellectual coordination of research goals and approaches across research sites. At least in terms of public debate and the selection and organisation of research projects, they can be

expected to have stronger consequences in these kinds of scientific fields than in those where scientists are more able to gain resources from external funding bodies.

Similarly, such assessment systems should significantly increase the degree of intellectual coordination and integration of research around common topics and concerns in sciences where goals and approaches have tended to differ between groups in different universities, such that separate schools of thought have developed and been reproduced in different research institutes. Especially where local traditions in, say, philosophy, have been important sources of intellectual variation and loyalty, the establishment of systematic, public and consequential evaluations of research quality are likely to encourage standardisation of research styles and strategies, as well as increased agreement about important research problems and contributions. This seems to have happened in Dutch philosophy when funding became linked to evaluations of international excellence (Van der Meulen and Leydesdorf 1991). In fields that already manifest considerable national and international coordination of research objectives and styles, on the other hand, the introduction of systematic research evaluation systems may simply reinforce this incrementally rather than leading to major shifts in problems and approaches.

This more marked influence of strong RES on sciences with relatively low levels of intellectual integration can also be expected to apply to more 'rural' fields that have low researcher to problem ratios. As summarised by Becher and Trowler (2001: 106-107), these kinds of science tend to be more diverse in the range of problems tackled by researchers and less concentrated in terms of the proportion of scientists working on the same or closely related issues. In contrast, 'urban' fields are those in which researchers focus intensively on a limited number of central problems and compete for reputations on the basis of their contributions to dealing with these.

Implementing strong RES in which disciplinary elites are forced to make judgements about the merits of different kinds of problems and contributions to them seems quite likely to reduce the diversity of topics and approaches deemed to be within the discipline in highly rural sciences, and so encourage more intellectual coordination and competition as researchers respond to this hierarchical arrangement of problems. Over a series of such evaluations, research on certain issues will come to be seen as more important and rewarded than that on others and so competition between scientists for the significance of their contributions to these core problems will grow. In fields that are already quite urban in this sense, on the other hand, strong RES should intensify competition and coordination.

The extent of this effect will, though, depend on the cohesion of scientific elites and their consensus on the central problems of the discipline. Where this is considerable, then we would expect strong RES to reinforce their authority and their control over the intellectual agenda of the field, especially where RES-based funding constitutes a major proportion of the resources needed for carrying out research. Since they will typically implement the RES, they will be able to determine the quality standards and intellectual priorities for judging research outputs and so intensify competition for reputations based on contributions to the central problems of the discipline. Research that deviates from current orthodoxies is unlikely to be regarded as worthwhile, or even competent, in these circumstances, and so intellectual pluralism will decline further.

In contrast, where elites are more differentiated, if not fragmented, in their intellectual goals and approaches, and there is limited consensus on the key problems of the field, the development of a strong RES may help to reproduce such separation of specialisms and topics rather than intensifying competition and coordination across the discipline. As long as no school or group is able to dominate the process and centralise control of key resources, the leaders of the major subfields could tacitly or explicitly agree to recognise each others goals and standards as equivalently high quality and so rank the 'best' research outputs dealing with each of these as excellent for the purposes of the evaluation exercise. Essentially, the RES would here increase coordination within each subfield and intellectual school but not strengthen integration between them, thus confirming the existing intellectual power structure in the discipline. This outcome is especially likely if the concentration of control over key research resources is low so that each group can gain access to them relatively easily.

Finally, just as we would expect the response of universities and other research organisations to strong research evaluation systems to vary according to their prestige and control over resources, so too the relative prestige and centrality of scientific fields should make a difference to their impact. Since such evaluation systems are likely to reflect the conceptions and standards of leading researchers in dominant scientific fields, and to be organised and implemented in ways that are complementary to current elites' perceptions of the best kinds of work, strong RES are more likely to reinforce their views, and the degree of elite control over research agenda and priorities, than to follow the research styles and goals prevalent in peripheral sciences. As they impinge upon funding decisions and university research strategies, they will therefore encourage the less prestigious fields to imitate the procedures and approaches – at least superficially – of the more central ones.

The impact of strong RES on marginal fields in terms of increasing competition and coordination should, then, be more visible than that on more central ones, especially when scientists depend greatly on RES-based funding. Hierarchies of journals and devaluation of alternative publication media for research results, formalisation of ranking systems, and increased awareness of 'international' standards will become more institutionalised in these kinds of fields, largely in imitation of the more successful sciences and current perceptions of their norms. Again, these might be expect to be especially marked in the less prestigious and autonomous research organisations, whose managers feel compelled to follow the 'rules of the game' established by the leading organisations and agencies, and are unable to insist on the significance of differences between the sciences. Reliance on bibliometric indicators of intellectual impact seems likely to reinforce this narrowing of research styles as well as having other consequences as discussed by Gläser and Laudel in this volume.

This may lead to changes in prevalent research styles and skills in the more peripheral sciences as traditional elites are challenged by more 'scientific' colleagues who seek to institutionalise common goals and standards as means of integrating research and demonstrating the scientific status of the field. In some cases, this search for scientific respectability has led to calls for the establishment of Kuhnian-style paradigms to deliver intellectual progress, apparently regardless of

their merits – any candidate will do (Pfeffer 1993). Depending on the cohesion and strength of current elites in a field, as well as the importance of evaluation system-based funding and reputations, this means that the establishment of strong research evaluation systems could result in qualitative shifts in elite membership and purposes in the less prestigious sciences.

CONCLUSIONS

This discussion has emphasised a number of points about the development and effects of establishing RES in the public sciences that can be summarised as three main conclusions. First, such retrospective evaluation systems are highly variable and path dependent in the sense that both how these institutional innovations in state-science relations became institutionalised and changed in different academic systems, and how they are affecting the direction and organisation of scientific research, strongly reflect the differing organisation of the public science system in each country, as well as the overall structure of professional labour markets and certain characteristics of the state. Differently organised groups of politicians, elite civil servants, business elites and other influential policy advisers have developed contrasting RES in varied ways in different kinds of state systems, and their effects have been mediated by varying characteristics of national academic systems, as the chapters on Australia, Germany, the Netherlands, Spain and Sweden in this book demonstrate.

Thus, although imitation and borrowing of particular aspects of RES have occurred between many of the OECD countries, sometimes to a considerable extent, as many contributions to this book demonstrate, their implementation in practice has been greatly affected by the continued substantial variations in how state agencies allocate funds for scientific research and higher education, how scientific elites are organised and manage relations with politicians and the state bureaucracy, and how the employment and careers of researchers are organised.

Second, a crucial factor in examining the development and consequences of RES in the public sciences is the changing nature and behaviour of universities and other research organisations in many of the societies where they have been introduced. How these have reacted to, and influenced, the implementation of different kinds of evaluations have had major effects on the organisation of scientists' careers, and professional labour markets more generally, as well as on the organisation and direction of research groups and fields. As many RES have been established as part of a general process of delegating some financial and administrative authority and resources to universities and similar organisations, they have become more significant collective actors in the development and organisation of scientific knowledge, and mediated the consequences of state policies, such as the introduction of RES, to a much greater extent in many countries than hitherto.

Although the extent of this process of financial and administrative decentralisation has varied considerably between countries in Europe and elsewhere, and in some is more apparent than real, to the extent that it has produced qualitative shifts in resource control and strategic autonomy from the state to research organisations, then how this autonomy has been used by different groups of

managers and researchers becomes more important than when they largely implemented the decision of ministries and faculties. This is especially so if some develop a more independent research policy that does not simply reproduce the results of peer review judgements in the RES and project proposal assessments.

How universities and other research organisations develop different kinds of research management capabilities as separate employment organisations becomes much more significant for the organisation and development of the public sciences as they gain more independence and control their own resources. While unlikely perhaps to generate such distinctive, organisation-specific competences through authority relations and managerial hierarchies as private companies in market economies are seen to do by adherents to the resource-based theory of the firm (see e.g. Metcalfe and James 2000; Penrose 1959; Whitley 2003b), the combination of RES and managerial delegation to research organisations makes the management and behaviour of universities and research institutes more consequential for scientific research than before in many countries. However, the actual autonomy and strategic capabilities developed by many universities and other research organisations in Continental Europe and Japan seems rather limited so far.

Third, while there are some general implications of adopting strong forms of RES for the sciences in general, especially where they are organised around established disciplinary boundaries and elites, their impact can be expected to vary considerably across different kinds of scientific fields, often in ways that are interdependent with particular features of national public science systems. Whatever kind of RES is implemented and changed in a country, it will complement and be more appropriate for some kinds of sciences and research styles than other ones, and so will potentially change some fields more than others.

Since the sciences vary in many ways, as well as changing over time as some fields become more iconic than others and come to dominate implicit and explicit hierarchies of scientific prestige, any evaluation system that tries to standardise assessments of the productivity and quality of scientific research is bound to have differential effects across the sciences. Whether this will in time lead to greater standardisation of research styles, publication practices and the organisation of scientific work depends, as I have suggested, on the nature and importance of the RES, how it is implemented, and on the nature of the public science system and behaviour of leading universities and other research organisations in each country.

REFERENCES

Becher, Tony and Paul Trowler (2001), *Academic Tribes and Territories: Intellectual enquiry and the culture of disciplines*, Buckingham: Open University Press and the Society for Research into Higher Education. Second edition.

Block**,** Hans-Juergen (1990), 'The University Systems in Transition: Possibilities and limitations of universities in the 'steady state', in Susan E. Cozzens, Peter Healey, Arie Rip and John Ziman (eds.), *The Research System in Transition*, Dordrecht: Kluwer, pp. 35-50.

Braun, Dietmar (1998), 'The Role of Funding Agencies in the Cognitive Development of Science', *Research Policy,* 27, 807-821.

Braun, Dietmar (2003), 'Lasting Tensions in research Policy-Making – A delegation problem', *Science and Public Policy,* 30, 309-321.

Brooks, Harvey (1990), 'Lessons of History: Successive challenges to science policy', in Susan E. Cozzens, Peter Healey, Arie Rip and John Ziman (eds.), *The Research System in Transition*, Dordrecht: Kluwer, pp. 11-22.

Campbell, David F. J. (2003), 'The Evaluation of University Research in the United Kingdom and the Netherlands, Germany and Austria', in Philip Shapira and Stefan Kuhlmann (eds.), *Learning from Science and Technology Policy Evaluation. Experiences from the United States and Europe,* Cheltenham: Edward Elgar, pp. 98-131.

Casper, Steven (2000), 'Institutional Adaptiveness, Technology Policy and the Diffusion of New Business Models: The Case of German Biotechnology', *Organization Studies,* 2, 887-914.

Casper, Steven (2006), 'How Do Technology Clusters Emerge and Become Sustainable? Social network formation and inter-firm mobility within the San Diego biotechnology cluster', Working Paper: Keck Graduate Institute of Applied Life Sciences.

Casper, Steven and Richard Whitley (2004), 'Managing Competences in Entrepreneurial Technology Firms: A comparative institutional analysis of Germany, Sweden and the UK', *Research Policy,* 33, 89-106.

Clark, Burton R. (1995), *Places of Inquiry: Research and advanced education in modern universities*, Berkeley: University of California Press.

Coleman, Samuel (1999), *Japanese Science: View from the inside*, London: Routledge.

Cozzens, Susan E. (1990), 'Options for the Future of Research Evaluation', in Susan E. Cozzens, Peter Healey, Arie Rip and John Ziman (eds.), *The Research System in Transition*, Dordrecht: Kluwer, pp. 281-294.

Dasgupta, Partha and Paul A. David (1994), 'Toward a New Economics of Science', *Research Policy*, 23, 487-521.

Drori, Gili S., John W. Meyer, Francisco O. Ramirez and Evan Schofer (2003), *Science in the Modern World Polity: Institutionalization and globalization,* Stanford: Stanford University Press.

Freeman, Christopher and Luc Soete (1997), *The Economics of Industrial Innovation*, London: Pinter.

Gaughan, Monica and Stephane Robin (2004), 'National Science Training Policy and Early Scientific Careers in France and the United States', *Research Policy*, 33, 569-581.

Hallam, A. (1973), *A Revolution in the Earth Sciences: From continental drift to plate tectonics*, Oxford: Oxford University Press.

Hansen, Hanne Foss and Finn Borum (1999), 'The Construction and Standardization of Evaluation: The case of the Danish university sector'. *Evaluation*, 5, 303-329.

Kenney, Martin (ed.) (2000), *Understanding Silicon Valley: The anatomy of an entrepreneurial region*, Stanford: Stanford University Press.

Langfeldt, Liv (2001), 'The Decision-Making Constraints and Processes of Grant Peer Review, and Their Effects on the Review Outcome'. *Social Studies of Science*, 31, 820-841.

Liefner, Ingo (2003), 'Funding, Resources Allocation, and Performance in Higher Education Systems', *Higher Education*, 46, 469-489.

Marginson, Simon and Mark Considine (2000), *The Enterprise University: Power, governance and reinvention in Australia*, Cambridge: Cambridge University Press.

Martin, Ben (2003), 'The Changing Social Contract for Science and the Evolution of the University', in Aldo Geuna, Ammon J. Salter and W. Edward Steinmuller (eds.), *Science and Innovation: Rethinking the rationales for funding and governance*, Cheltenham: Edward Elgar, pp. 7-29.

Martin, Ben and Aldo Geuna (2003), 'University Research Evaluation and Funding: An international comparison'. *Minerva,* 41, 277-304.

Metcalfe, J. Stanley and Andrew James (2000), 'Knowledge and Capabilities: A new view of the firm', in Nicolai Foss and Paul Robertson (eds.), *Resources, Technology and Strategy: Explorations in the resource based perspective*, London: Routledge, pp. 31-52.

Mowery, David C. (1999), 'The Computer Software Industry', in David C. Mowery and Richard R. Nelson (eds.), *Sources of Industrial Leadership*, Cambridge: Cambridge University Press, pp. 133-168.

Mowery, David C., Richard R. Nelson, Bhaven N. Sampat and Arvids A. Ziedonis (2004), *Ivory Tower and Industrial Innovation: University-Industry technology transfer before and after the Bayh-Dole Act,* Stanford, California: Stanford University Press.

Muller-Camen, Michael and Stefan Salzgeber (2005), 'Changes in Academic Work and the Chair Regime: The case of German business administration academics', *Organization Studies*, 26, 271-290.

Penrose, Edith (1959), *The Theory of the Growth of the Firm*, Oxford: Blackwell.

Pfeffer, Jeffrey (1993), 'Barriers to the Advance of Organization Science'. *Academy of Management Review*, 18, 599-620.

Ruivo, Beatriz (1994), ''Phases' or 'Paradigms' of Science Policy?', *Science and Public Policy*, 21, 157-164.

Sienko, Tanya (1997), *A Comparison of Japanese and U.S. Graduate Programs in Science and Engineering*, Tokyo: National Institute of Science and Technology Policy, Discussion Paper no. 3.

Stokes, Donald E. (1997), *Pasteur's Quadrant: Basic science and technological innovation*, Washington, D.C: Brookings Institution Press.

Tabil, Ameen Ali (2001), 'The Continued Behavioural Modification of Academics since the 1992 Research Assessment Exercise', *Higher Education Review*, 33, 30-46.

Van der Meulen, Barend and Loet Leydesdorff (1991), 'Has the Study of Philosophy and Dutch Universities Changed under Economic and Political Pressure?', *Science, Technology, and Human Values*, 16, 288-321.

Whitley, Richard (2000), *The Intellectual and Social Organization of the Sciences*, Oxford University Press. Second edition.

Whitley, Richard (2003a), 'Competition and Pluralism in the Public Sciences: The impact of institutional frameworks on the organisation of academic science', *Research Policy,* 32, 1015-1029.

Whitley, Richard (2003b), 'The Institutional Structuring of Organisational Capabilities: The role of authority sharing and organisational careers', *Organization Studies*, 24, 667-695.