

Measurements and Distortions

Susan Wright

► **To cite this version:**

Susan Wright. Measurements and Distortions: A review of the British system of research assessment. Working Papers on University Reform, Jun 2008, Aarhus, Denmark. <hprints-00308631>

HAL Id: hprints-00308631

<https://hal-hprints.archives-ouvertes.fr/hprints-00308631>

Submitted on 31 Jul 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Working Papers on University Reform

Working Paper 9:

Measurements and Distortions

- A review of the British system of research assessment

By Susan Wright

Danish School of Education, University of Århus

June, 2008

Working Papers on University Reform
Series Editor: Susan Wright

This working papers series is published by the research unit "Transformations of universities and organizations" at the Danish School of Education, Århus University. The series brings together work in progress in Denmark and among an international network of scholars involved in research on universities and higher education.

"Transformations of universities and organizations" aims to establish the study of universities as a field of research in Denmark. The field has three components: International and national policies to revise the role of universities in the knowledge economy; the transformation of universities as organizations; and changes in teaching and learning in higher education. The research unit aims to understand each of these components in the context of the others and explore the links between them. The central questions are: How are different national and international visions of learning societies, knowledge economies, and new world orders spurring reforms to the role and purpose of universities and to the policies and practices of higher education? How do university reforms introduce new rationalities of governance, systems of management and priorities for research and teaching? How do managers, researchers, and students negotiate with new discourses, subject positions and forms of power arising from university reforms? What kinds of changes are students, academics and university managers themselves initiating, and how do these interact with reforms coming from governments and international organizations? This interaction is central to the work of the research unit.

The unit draws together ideas and approaches from a range of academic fields and collaborates internationally with other higher education research environments. Currently the unit's main activity is a research project, 'New management, new identities? Danish university reform in an international perspective' funded by the Danish Research Council (2004-2007). The unit holds seminars and there is a mailing list of academics and students working in this field in Denmark and internationally.

Members of the unit include professor Susan Wright, Danish University of Education, associate professor John Krejsler, Danish University of Education, assistant professor Jakob Krause-Jensen, Danish University of Education, Ph.D. student Gritt Bykærholm Nielsen, Danish University of Education, and research assistant Jakob Williams Ørberg, Danish University of Education.

Further information on the research unit and other working papers in the series are at <http://www.dpu.dk/site.asp?p=5899>. To join the mailing list, hold a seminar or have material included in the working paper series please contact professor Susan Wright at suwr@dpu.dk or at the Danish School of Education, Århus University, Tuborgvej 164, 2400 Copenhagen NV, Denmark.

**Measurements and Distortions
- A review of the British system of research
assessment**

SUSAN WRIGHT

**Danish School of Education, University of Århus
Tel.: + 45 8888 9179
Email: suwr@dpu.dk**

Copyright: Susan Wright

Contents

Introduction	1
The purpose and importance of the RAE in the context of changes in British higher education.....	2
Principles underpinning RAE	6
RAE method	7
‘Collateral damage’	10
Future of RAE.....	13
Metrics and international rankings.....	17
Skewing Effects	18
<i>Peer review – whole process</i>	19
<i>Effect on disciplines as a whole – Research Funding Council’s disciplinary benchmarking</i>	19
<i>Corporate costs</i>	21
Conclusion	23
References	26
Notes.....	29

Introduction

Many European countries, and notably those in Scandinavia, are in the process of devising systems for assessing universities' research output and for differentiating funding for universities according to the results. From the autumn of 2007, debates in Denmark about ways of measuring and rewarding academic research began gathering steam.ⁱ The Ministry of Science, Technology and Innovation had already been working for a year on a model for allocating 5 per cent of the budget for universities competitively between institutions on the basis of a 'quality barometer'.ⁱⁱ Universities had entered into contracts with the Ministry committing them to allocate a part of their government grant internally on the basis of research quality. For example, Copenhagen University undertook to develop a method for allocating 10 percent of its research funding on the basis of indicators of international quality.ⁱⁱⁱ Faculties also committed, in their contracts with the Rector, to establish systems for grading the publications of each department based on the impact of the journals in which they were published, and for establishing 'star research programmes'. From 2008, faculties would allocate 10 percent of their budgets between departments accordingly.^{iv} The Debate about proposed systems exploded into the press. Meanwhile the Ministry continued to explore different models.

There is no shortage of such models for 'steering by numbers' as a recent report for the OECD showed (Frølich 2008). Britain had introduced the Research Assessment Exercise (RAE) in 1986 in order to measure the quality of the research of each department and concentrate government grant on the best departments. In 1993, Hong Kong modified the RAE model, adapting it to its smaller sector (8 universities). Hong Kong's method was still based on peer review, but used the individual academic, rather than the department, as the unit of assessment. Meanwhile, Australia went in another direction with its Institutional Grants Scheme (IGS). Rather than peer review, the IGS was based on performance indicators. If the compliance costs were lower than for the RAE, the IGS suffered from heavy flaws in the data used, the unreliability of some indicators, and the skewing effects of inappropriate weightings on research activity and its quality - notably the emphasis on the volume of external research income. As the weaknesses of models based purely on performance indicators became evident, in 2003 New Zealand in its Performance-Based Research Fund (PBRF) opted for peer review of each individual's evidence portfolio in a quality evaluation of research outputs (60 percent), supplemented by two proxies for quality: research degree completions (25 percent) and external research income (15 percent). Both the Australian and New Zealand systems have been evaluated and have now gone through a second reform, focusing more on research quality and impact (Bakker et al. 2006; Sheehan 2006). Meanwhile, the British authorities have been exploring ways of reducing the costs of the RAE, shifting from measuring quality to impact, and trying to find bibliometric measures, such as that developed by Leiden University, to reduce the cost of peer review. The Danish Ministry first considered the RAE as a

model, and then a model to measure the quantity of publications in 'core' journals being developed in Norway.

The British RAE has been repeatedly studied and adjusted, but the literature is quite dispersed: one of the aims of this paper is to provide a review to draw out the strengths and problems of this peer review system and its bibliometric and other alternatives. A second aim is to examine the ways that, according to the literature, the RAE distorts academic research. There is no doubt that such distortions occur. The Funding Council for England itself stated that 'Any assessment process, particularly one as important to its subjects as the RAE, will distort the very thing it intends to measure' (quoted in House of Commons 2004: 36). The third aim is to go beyond this to identify ways that such assessment systems skew university activities much more fundamentally than so far recorded in the literature. Three things are identified in particular: skewing of the conditions for critical dialogue among academics, which is arguably fundamental to processes of research; the effects of skewing the distribution of certain core disciplines across the country, and especially the social inequalities that result from 'regional deserts'; and inefficiencies in the use of public funds, when the competition between universities to improve their rankings results in their ever-increasing use of public funding to subsidise the profits of the four big commercial journal publishers.

The paper draws both on academic studies (Adams and Smith 2006; Elton 2000; McNay 1996, 1999; Velody 1999; Whittington 1997) and on official reports. Roberts (2003) conducted a review and public consultation after the 2001 RAE to inform the shape of future RAEs. The House of Commons' Select Committee on Science and Technology (House of Commons 2004) conducted its own inquiry and response to Roberts. The British Academy (2007) established a panel to report on research evaluation through peer review and metrics specifically in the humanities and social sciences. Running through all these reviews, reforms, and re-reforms are questions about whether to measure output, impact or quality? How costly are the systems? And in what ways do they distort academic work? These questions are addressed in this paper, with the aim of contributing to international debates an analysis of the operation and effects of the British Research Assessment Exercise.

The purpose and importance of the RAE in the context of changes in British higher education

As the ways academic work are organised and assessed vary in detail from country to country, it is worth providing some background to the British situation and the context within which the Research Assessment Exercise was introduced. Although most academics are expected to allocate part of their time to research and part to teaching, not all employment contracts state the hours that an academic is to work each week, let alone the proportions they are to spend on research and teaching. In many old

universities there are ‘traditions’ such as 40 percent teaching, 60 percent research. Often universities calculate teaching allocations using a currency invented for the purpose (variously referred to as work points or ‘beans’) but which has no fixed rate of conversion to hours or income. This means work loads can be continually inflated, without upsetting the fiction that the rising teaching load still only constitutes 40 percent of the week’s work. A second fiction is that all the remaining 60 percent is available for research, whereas administration also has to be carried out in this time. Some formal tasks, like head of department or teaching convenor, are allocated a number of work points, but the burgeoning record keeping and reporting systems that all academics have to undertake for audits of both teaching and research have meant an enormous increase in administration. This administrative load tends not to feature in the official fictions about academic work, so that academics are meant to fit it in somehow, but still produce a research output in keeping with 60 percent of their time.

This fictional way of allocating time may have suited academics when their workload was stable but from the 1980s onwards, their work has intensified. There has been a shift from elite to mass higher education. Participation of the 18-25 year-old cohort expanded from 5 percent in 1960s, to 15 percent in 1980s, and nearly 50 percent in 2007. Not only have student numbers risen steeply, but in the 20 years to 1997, government funding per student declined by 40 percent (Dearing 1997). The introduction of competitive student fees in 2004 (capped at £3000 per annum) was only ever expected to cover half of the annual £2 billion shortfall in government funding of teaching (Wright 2004). In the 1980s, a Teaching Quality Assessment system was introduced to reassure government that they were justified in their view that there had been so much slack in the system that teaching loads could be increased and funding reduced without affecting the quality of teaching. Indeed, the method of audit combined with the professionalism of academics resulted in consistently high scores for teaching quality (Shore and Wright 2000). The danger from government’s and managers’ points of view, was that academics would concentrate their efforts on the teaching quality audit and neglect their research. The periodic Research Assessment Exercise was effective in turning academics’ attention to research because, and as the government’s research grant became increasing differentiated according to RAE scores, maintaining or improving this score became the over-riding priority for departments. The Research Assessment Exercise was therefore introduced in the context of a steep intensification of academic work, and itself further contributed to that process.

The RAE is a national system, standardised across all disciplines, which has been held every 3-7 years (1986, 1989, 1992, 1996, 2001, 2008). The 2001 RAE involved 50,000 researchers in 2,598 submissions from 173 universities. The results of the 2008 RAE will be announced shortly. Up till now, the system has always been based on peer review of researchers’ publications, supported by some numerical indicators. The RAE grades each department on a 7 point scale from 5* at the top to 1 at the bottom.^v Universities decide how many of their staff to submit to the RAE, and the results also include an alphabetical ranking from A (95-100% academic staff submitted) to F

(below 20% of academic staff submitted). The median in 1996 was D (40-59% submitted) (Whittington 1997: 182). The Higher Education Funding Councils for England, Scotland and Wales use the numerical and alphabetical RAE grades to allocate the annual research funding for each university for the period until the next RAE. The Funding Council's allocation of a block grant to each university covers teaching as well as research and is also based on each department's registered student numbers. The university then decides how to allocate this block grant between research and teaching and across departments. Universities have increasingly adopted devolved budgeting systems which pass rewards to departments which earned more income for the university through high RAE grades, and punish those which did not do so. HEFCE's total budget was £5,993m in 2004-5, of which £1,081m was for research and £3,826m for teaching (House of Commons 2004: 6). In addition to the block grant, seven Research Councils, each covering a specific disciplinary area, allocate project funding on a competitive basis, now calculated at 'full economic cost'. The combined budget of the Research Councils in 2003-4 was £1,900m (House of Commons 2004: 6). Both parts of this 'dual-support system' allocate funding competitively through processes heavily reliant on peer review. The government published a 10 year investment framework for science and innovation in July 2004 and confirmed that the dual-support system will continue.

In the first RAE, in 1986, only 14 percent of the Funding Councils' budget for research was allocated according to the results. This increased to 30 percent in 1989, and 100 percent from 1992 onwards (Welch 2007). From 1985 the government also decided that the funding for teaching would continue to be allocated on the basis of each registered student rather than selectively on the basis of the results of the Teaching Quality Assurance system. In the context of expanding demand and declining resource, research funding often subsidised the costs of teaching and the RAE result was the one and only variable that could make or break the economy of a department.

In effect, departments were set against each other to compete over the division of a fixed-sized cake. The RAE is a punitive system because the very top departments gain the lion's share at the expense of lower tiers of departments. These unfortunate departments are still very good, but they compete from a very unequal basis. The system has gradually differentiated the sector, by concentrating research resources on selected universities and leaving others to focus mainly on teaching. This concentration of research funding on fewer departments had a blip in 1992. When the polytechnics were made into 'new' universities, they entered their strongest departments in the 1992 RAE and even if they got low grades, this yielded them a research bonus as they had never had basic funding for research before. This temporarily spread the funding out again over the enlarged university sector.

The following picture emerged from McNay's study of the 1992 RAE:

1. 'Assured' departments – top 'old' universities that gained top grades and increased funds.
2. 'Anxious' departments – 'old' universities with good grades but reduced research funding and declining teaching funding – squeezed.
3. 'Aspirant' departments – 'new' universities with a little basic funding for some departments' research for the first time.

By 2001 crunch point was reached. The 2001 RAE showed that the quality of research had increased enormously, but the government refused to increase the Funding Councils' research budget to reward the improved departments. In 1996, 31 percent of the academics included in the RAE worked in 573 departments rated top (5 or 5*). By 2001, 55 percent of researchers included in the RAE were in 1,081 departments rated top (5 or 5*) and 80 percent of academics were in departments in the top three grades (4, 5 or 5*) (House of Commons 2004: 34; Roberts 2003: 4). This astounding increase was checked internationally: all judgements about 5 and 5* departments were validated by 290 overseas experts, who only rejected 9 judgements (House of Commons 2004: 13). If the Funding Councils were to fund all departments according to their improved grades, their budget needed to increase by £206m. Instead the Ministry only gave them an additional £30m. (House of Commons 2004: 7). The Funding Councils used this crisis further to concentrate funding in the top departments: instead of spreading what was effectively a cut in funding evenly across all grades, the Funding Councils maintained the funding formula for the 'assured' 5 and 5* departments at the same level as before, reduced the funding formula for the 'anxious' grade 4 departments and eliminated basic research funding for departments below that grade, many of which had up to that time had been the 'new' universities' 'aspiring' research centres. The way this concentrated research resources in top departments can be seen in the following table. Whereas the top departments' received approximately 50 percent of research resources in 1996, they got nearly three quarters of the research resources after 2001.

Table 1. Percentage allocation of overall research funding to differently ranked departments in England and Scotland

Research Assessment Exercise Ranking	1996	2001
Top departments (Grade 5 and 5*)	54.3	74/70.2*
Middle departments (Grade 4)	16	18
Bottom departments (Grade 1-3a)	29.3	10/11.7*
Totals	99.6	102

* England and Scotland respectively

Source: La (Manna 2004)

The situation now resembles the football premier league in which the clubs at the top of the table get the overwhelming share of the television rights and the bottom clubs hardly receive enough crumbs to survive. During discussions about the 2008 RAE, the Treasury made known that it wanted to go further – to create a super league.^{vi} It favoured concentration of research resources on the top 10 universities, and within

those universities to focus funding further on science, technology, engineering and medicine (Fazackerley 2006) even though these STEM subjects already gain the lion's share of the available project funding too. The House of Commons Select Committee (2004: 10) has argued against this policy of concentrating resources at the top and in favour of a system of resource allocation that also rewards development and improvement. It is unpredictable how this selective funding will interact with the market in higher education, created in the 2004 Higher Education Act, but only (probably) coming fully into effect in 2009.

Principles underpinning RAE

The RAE has always had two elements: first, peers assess the quality of each department's research nationally and internationally, then they give the results to the Higher Education Funding Councils, which decide how to allocate basic research funding selectively to departments. The basic principles underpinning the RAE method are that it should have the confidence and consent of the academic community (not just the university leaders) and should be fair, transparent and efficient (Roberts 2003: 6). While these principles may accord with academic values, the RAE is extremely unpopular among academics – in a survey of 1,500 academics by the union UCU in 2006 only 17 percent felt positively that the RAE 'promotes healthy competition and guarantees [that] excellence is rewarded'. 87 percent thought that it had negative effects on research, and 41 percent wanted to scrap it altogether. But, if an RAE had to be retained, 81 percent opposed a metrics system and 57 percent preferred to stay with peer review (Baty 2006).

This complex response may be explained by the different ways the two halves of the RAE operate. It seems that the principles of fairness and transparency are largely upheld in regard to the first element of the RAE, the peer assessment of the quality of each department's research nationally and internationally. For the 2008 RAE this work is being done by 67 subject panels, coordinated by 15 main panels (HEFCE 2006). On the whole, the panels' peer review tends to be widely respected and to have academics' confidence and consent. In regard to the second half of the process, when the Higher Education Funding Councils use the RAE results to allocate basic research funding selectively to top departments, the system does not have the confidence of the academic community (including university leaders) and it is not thought to be fair, transparent and efficient. The Funding Councils do not announce in advance how the RAE results will be used to calculate funding. The effects of this were disastrous in 2001, when, as explained above, many departments found that, as a result of maintaining or improving their grade, their funding was cut. The House of Commons' Select Committee report makes strong statements on this issue: the Funding Councils are not giving universities any idea in advance how their 2008 RAE scores will translate into funding, they are requiring universities 'to develop investment strategies with no basis for calculating potential return', which is asking universities 'to play the

game' 'blindfolded' and 'without a rule book', as well as leaving the system open to 'retrospective manipulation' (House of Commons 2004: 3, 33).

The RAE is consistently referred to in the UK as a system of 'peer review' but following Foss Hansen, it is more accurately a 'modified' peer review (Hansen and Borum 1999: 308). In peer review, the peers not only assess their colleagues' work, but the decision making and outcomes also lie in their hands. While this is not perfect, and there is room for bullying and cronyism, such negative tendencies are tempered by the constant awareness that others may do to you, what you do to them. This induces an ethic of fairness and responsibility, which, it is widely accepted, peers exercise whilst setting the RAE grades. One of the important modifications in the RAE system is that peers do not make their own decisions on the basis of the RAE evaluation. They give the RAE grades to the Funding Councils and have absolutely no influence on the way the RAE grades are used in decision making thereafter. In effect, the RAE grades they create are an intermediate currency, and the peers have no idea how they will be translated into money. The government or the Funding Councils can inflate the RAE currency at will. A grade 4 in the 1996 RAE was a good grade; after the 2001 RAE funding decisions it could be a death sentence for a department. The RAE appropriates peer review and incorporates it into decision making based on quite other principles which are not fair or transparent, and, as argued below, has effects which are far from efficient.

RAE method

Most universities have appointed a pro-Vice Chancellor and a central management staff to steer each department's RAE submission. As soon as one RAE is over, the central management 'support' each head of department in setting up a departmental strategy for the next RAE. The head of department also has to make his or her academic staff's various research activities look like a coherent strategy with focused priority areas. Each individual needs to be made aware of the output that is expected of them, the number of books and articles to produce, and which journals to aim for. They also have to be discouraged from doing anything else (for example teaching development or public communication of research) that would distract from putting their maximum effort into publication. Consultants are brought in to conduct a 'dry run' of the submission and advise where the weaknesses are or which areas of research could be maximised. Several universities have developed investment strategies to buy in a professor or a research team to try and secure a 5*. For example Liverpool University appointed 40 new 'star' professors who were world leaders in their fields to raise the RAE results, and invested £1m in teaching support to release existing staff to concentrate on their research (THES 18 July 2005). The 'transfer market' now applies to 'star' academics as much as to footballers. Notably the strategy of the Vice Chancellor of the fused University of Manchester included recruiting 5 Nobel prize winners by 2015 (Tysome 2006). Such stars will officially be on the university's pay roll, but most only work part-time in the UK and may not be

very much in evidence. Sometimes an individual is moved internally to boost another department's RAE submission; sometimes departments are fused to strengthen a bid, especially as the RAE seems to favour large departments (House of Commons 2004: 7. 22). Some universities which were established in the 1970s on the basis of interdisciplinary units have gone through wholesale reorganisation so that their departments match the disciplinary map of RAE panels and so as to avoid the danger of falling down the cracks between RAE panels.

In the run up to preparing the RAE submission, a regular check is kept on each department to make sure they are performing as required. Importantly, decisions are made about which academics in each department to include in the submission. As the funding formula calculates not only the grade but the number of people in the department who have contributed to gaining that grade, it is a fine judgement to work out if the department will earn more money by sending in just the top staff and aiming for a high grade, or by aiming for more people to be calculated at the rate for a lower grade. In the parlance, the question is whether 'to cut off the tail'. This decision involves designating some people 'not research active', which means they will not earn the department any basic research funding in the next period and will be given a much higher teaching load. This is a designation from which it may be almost impossible to recover, and can spell the end of a research career.

The materials to be submitted have varied over time. The first RAE in 1986 assessed the whole published output of each academic included in the review.^{vii} This, it quickly became clear, meant that quantity 'counted' as much, if not more, than quality. To counter this dynamic and emphasise that the RAE was about the quality, not just the volume, of research, in the 1992 RAE each academic was asked to submit their two best pieces of work for review plus aggregate numbers of publications. Still this statistic was not found helpful to assess quality, so from 1996, each academic was asked to submit the four best pieces of work that were published in the period under review. Departments still have had to submit other information and in 2001 this included: academics' CVs, their total research output, details of research students, research income, the department's research strategy, and any exonerating 'staff circumstances'.

Each panel decides the 'Criteria for Assessment' in its subject, but the general understanding is that theoretical articles written for an academic audience carry most weight. Academics soon learnt, often with guidance from management, to become 'accountable selves' and to adjust their academic practices to focus on 'what counts' (Shore and Wright 2001). The House of Commons' report refers to such practices as 'gamesmanship'. For example, in order to ensure that researchers have four publications out in time, instead of consolidating the results of a piece of research in one article or book, academics are encouraged to 'salami-slice' them into numerous small articles, which they send to different journals. Alternatively, sometimes an academic repackages an article by giving it another title and changing the order of the paragraphs, in order to publish very nearly the same material more than once.

Academics also have to rush to press as soon as they have any results emerging from a project, rather than making a more mature and considered interpretation of the whole piece of work. Even more seriously, the Dean of the Business School of Durham University has recently been suspended and investigated for plagiarism in his PhD thesis and a journal article (Tahir 2007). Whereas in theory academics still espouse unwritten codes of integrity and quality in scholarly publication, they can become entrapped when, as Calabrese and Roberts point out in their study of academic cheating in peer review systems in the USA, policy processes and management systems establish a ‘theory in use’ of ‘publish or be damned’ which condones skirting the rules of the game and deviation from these standards (Calabrese and Roberts 2004: 330, 337).

Whittington (1997) provides an account of the processes the accountancy panel followed in reaching its judgements. Each department’s submission was read by all panel members, but was read in most detail by two panel members who ‘double marked it’ and reported their judgement to the panel. Even though panels’ judgements are expressed in a 7 point scale, only the top half is really used – 3a, 4, 5, 5*. There has been considerable criticism of the ‘cliff edges’ between these grades, where a marginal decision about whether a department is a grade 4 or grade 5 makes an enormous difference to their funding and their future. To solve this problem, the RAE 2008 will produce a ‘graded profile’ for each department (Figure 1), based on three main criteria: research outputs (still primarily 4 best publications per person), research environment, and indicators of esteem. The new grading scale is for research that is world leading (grade 4), internationally excellent (grade 3), internationally recognised (grade 2) and nationally recognised (grade 1). This system is designed to highlight pockets of excellence within a department, but by the same token, it also identifies the presence and size of a department’s ‘tail’. It is still unclear how these ‘graded profiles’ will be used in national and international league tables - and arguably a large part of this effort is directed towards securing the ranking of a country’s universities in the international league tables. It is equally unclear how the Funding Councils will turn the graded profiles into funding formulas – universities are still playing the game blindfolded and without goal posts.

Figure 1. Quality profiles

Unit of assessment A	Full-time equivalent research staff submitted for assessment	Percentage of research activity in the submission judged to meet the standard for:				
		4*	3*	2*	1*	Unclassified
University X	50	15	25	40	15	5
University Y	20	0	5	40	45	10

Source: (RAE2008 2007 <http://www.rae.ac.uk/aboutus/quality.asp> accessed 25.1.2007)

‘Collateral damage’

Whilst, as quoted above, the Funding Council for England recognised that the RAE distorts the very thing it intends to measure, the House of Commons’ Select Committee’s report went further, and referred to the ‘collateral damage’ caused by the RAE (House of Commons 2004: 36). This section reviews four kinds of collateral damage identified in official documents.

First, the House of Commons’ Select Committee’s report finds that the RAE in its dominance over what goes on at universities, distorts the balance of academic work. Elton’s review of the 1992 and 1996 RAEs also found that instead of combining research with teaching and other forms of dissemination, departments were using casual labour (PhDs, Teaching Assistants, part-time lecturers) to teach undergraduates, so that permanent staff could focus on publishing (Elton 2000: 278; see also McNay 1996: 152). Teaching has become increasingly divorced from research in some top universities. Those academics who do not accept such a division of labour and who try both to perform for the RAE and to maintain a mix of research, teaching, application and dissemination that they associate with an academic ideal, or those in the ‘anxious’ departments whose funding gives them no choice but to dance on all three legs at once, are exhausted from juggling activities and working long hours and are extremely stressed. In 2006 levels of stress among academics were greater than among traditionally high-stressed occupations (Kinman et al. 2006) and the Health and Safety Executive (the government agency responsible for work environment and safety) targeted universities in an audit of stress levels and a campaign to implement new management standards (Baty 2005).

Second, the RAE distorts research itself. There is no doubt that research output has increased considerably but McNay’s survey of the 1992 RAE found heads of department sceptical about whether this increase was attributable to RAE or was happening in spite of it. They thought the results might be presentational rather than a sign of real change (McNay 1996: 123). Two very senior academics who gave evidence to the House of Commons’ Select Committee attributed the measured improvement in UK research quality to ‘game playing’ so that the results ‘represent a “morass of fiddling, finagling and horse trading”’ and ““are starting to lack credibility”’ (House of Commons 2004: 21).

Even more importantly, the RAE has been attributed with changing the nature of UK research and discouraging longer term ‘blue skies’ research. Responsible academics and colleagues, accountable in the terms required by the RAE, focus on ‘what counts’ to maximise their department’s income. They set themselves short term research goals, which will be sure to generate publishable results within the next four-year RAE period. There is also an unwillingness to cross disciplinary boundaries that mean

results will not fit with RAE panels, and a reluctance to explore adventurous ideas that may not lead to publishable results (House of Commons 2004: 36-7). Charlton (Charlton and Andras 2008; Charlton 2008) takes this point further and demonstrates that the period since the first RAE in 1986 is marked by an increase in what Kuhn called 'normal science' - an incremental extrapolation of established ways of thinking. In the same period there has been a decline in what Kuhn called 'revolutionary science' that addresses big problems that are intractable to normal science and, if successful, will produce new paradigms, theories or discoveries. Charlton uses two measures of 'revolutionary' scientists: the number of Nobel laureates, and the number of 'highly cited' academics in the Thomson Web of Science - that is, the most highly cited people in their field. Whereas the UK used to have more laureates per capita than the USA, from 1987 to 2006, the UK had 9 and the USA 126 (including 5 British brain-drainers). To take his analysis further, Charlton listed the number of Nobel laureates and highly cited academics in each university in the USA and in the UK. By these criteria, Oxford University was equivalent to the University of Minnesota - a big and successful university in normal science, but third tier by US standards. As the University of Oxford still appoints the highest quality scientists in the world, Charlton asks why they are not producing the highest quality, revolutionary science of which they are capable? His answer is that scientific ambition to tackle big question is a sign of uncollegial selfishness, because under the RAE, research which will not predictably and quickly produce publishable results has penalties not just for the scientist in question but for all those around him or her. Charlton argues that Oxford academics have curbed their ambitions and became superbly effective at winning large grants and generating large volumes of papers, which score high in the RAE, but are not revolutionary science.

The RAE is also considered to have distorted research by making 'pure', theoretical work count more than applied. Although some RAE panels have said that they will give consideration to applied work, nobody believes them! As the House of Commons' report put it

If there is a perception that panels will not give parity to pure and applied research then departments will be disinclined to include applied research outputs in their submissions and ultimately to conduct this research at all (House of Commons 2004: 16).

The RAE's focus on theoretical research for an academic audience creates a considerable disincentive to applied research, practice-based research, collaboration with practitioners, industry or users. The Lambert report argued that the RAE should recognise excellent research conducted with industry to be on a par with academic research (Lambert 2003). But forms of output other than articles in top journals, for example reports for users, are 'iffy' and communication with other audiences, for example, through a television programme, or professional or popular publication, 'don't really count'. A Royal Society survey revealed that 64 percent of young scientists felt they had to concentrate on their research and had not time to

communicate to the public (THES 30 June 2006). Overall the RAE creates a disincentive to engage with the surrounding society: deflection of energy into such activities is a threat to the department's research excellence. This is of significance to countries such as Denmark, where the government both wants to fund research differentially in a way that privileges top disciplinary and theoretical journals and at the same time expects academics to engage with a range of stakeholders and make their research useful and significant for the surrounding society.

The third kind of collateral damage results from the way the system is coercive. It evokes a 'competitive, adversarial and punitive spirit' (Elton 2000: 275) and in the words of the President of an Oxford College who is a beneficiary of the system, its 'funding dirigism' leads to a 'narrow competitiveness within general conformity and compliance' (Smith 1998 quoted in Elton 2000: 280). It is made clear to everyone that there is a direct link between their individual performance and their department's RAE result. Everyone has to play the game, and 'perform' according to what counts, or else, as mentioned above, be designated 'not research active'. The overall scale of the resulting waste in research talent is unknown, but it is clear that it is women who suffer especially under the system (AUT 2004). Only 46 percent of women academics, as against 64 percent of men were included in RAE submissions (THES 18 August 2006). The submission includes a section for exonerating circumstances, to explain why someone has not achieved 4 top quality outputs in the previous period, but the assumption is that everyone has elastic hours and no other commitments. For those who are also parents, as an opinion piece in the Times Higher Education Supplement recently pointed out, the first question on the RAE form should be 'Have you reproduced in the last seven years?' Having a child means that the author is unlikely to have had the periods of uninterrupted research time on which the whole exercise is predicated (Scurr 2007). The House of Commons' Select Committee expressed concern over the 'under-representation of women in the highest-rated departments and that women have been disproportionately excluded from the RAE' (House of Commons 2004: 24).

The assessment system is also punitive at the level of the department. The funding for the sector is not elastic so the best are competing to deprive their colleagues. The differential funding rewards the successful and punishes even excellent departments that have not hit the top grade. Even though universities have the power to distribute their block grant in any way they choose, and could put funding into the development of lower rated departments, that could be seen as threatening the continued status of the top departments. There is a lack of systematic research on how universities allocate their funding between departments and the extent to which they replicate the Funding Council's punitive approach in their own internal allocation of funding. But in those subjects where the funding system for teaching does not cover actual costs, a low RAE score may spell financial ruin for the department. In recent years there has been a spate of closures of departments of the STEM subjects (chemistry, physics, engineering, medicine), mathematics, linguistics, and 'minority' languages, especially Asian and Middle Eastern languages. This has caused alarm as such subjects are

arguably crucial for Britain's standing in the purported global knowledge economy. Universities now have to give the Funding Councils a year's notice if they intend to kill off a threatened species. Even in subjects where weaker departments are not facing closure, the Funding Councils' concentration of funding on top departments in a handful of universities is seen as damaging to research and as endangering the production of the next generation of academics. A submission to the House of Commons' Select Committee from geography argued that they needed a full range of departments, including those graded 3 and 4, to be active in research, as they are an important bedrock and training ground for the discipline (House of Commons 2004: 37). The House of Commons' Select Committee's report says several times that there should be development funding for weaker departments based on a department submitting a 'business plan' to achieve a higher grade next time (House of Commons 2004: 7, 10). Such a suggestion, based on good educational values of encouragement and improvement, runs entirely counter to the intentions and dynamic of the system.

Future of RAE

There is currently a debate in Britain about the future method of allocating research funding. All agree it should be on the basis of research quality. The Treasury seems determined to replace peer review with a cheaper system based on citation indexes. Against this is the argument that citation indexes can measure output and possibly impact, but not quality. The Funding Councils argue that the existing system is efficient because, they say, the direct costs of the RAE are only 1 percent of the research budget that is to be distributed in the ensuing period. The Funding Councils' direct costs were £5.6m in 2001 and they are expected to be £10m in 2008 (House of Commons 2004: 27). They calculate the cost of the RAE for universities was £30-37m in 1996 (so maybe double that in 2008). But it is almost impossible to calculate the staff time involved and the opportunity costs incurred. There have been recurrent attempts to find a less costly system, but a major review of the RAE by Roberts concluded:

Some of us believed, at the outset of the process, that there might be some scope for assessing research on the basis of performance indicators, thereby dispensing with the need for a complex and labour-intensive assessment process. Whilst we recognise that metrics may be useful in helping assessors to reach judgements on the value of research, we are now convinced that the only system which will enjoy both the confidence and the consent of the academic community is one based ultimately upon expert review. We are also convinced that only a system based ultimately on expert judgement is sufficiently resistant to unintended behavioural consequences to prevent distorting the very nature of research activity (Roberts 2003: 6-7).

From June to October 2006 there was a consultation with 300 responses and in December 2006 the Minister announced that the next RAE after 2008 will use indicators of research income and quality combined with advice from experts including users. He intended that this would be administratively lighter for universities (HEFCE 2006).

In November 2007 a further proposal went out to consultation. This proposed two systems, a Research Excellence Framework (REF) for STEM subjects (science, technology, engineering and medicine) starting in 2009 and a 'light touch' (undefined) Research Assessment Exercise for the arts and humanities starting in 2013. The REF model for STEM subjects consists of three elements:

- Research income
- Postgraduate numbers
- Bibliometric indicator of research output

The Research Funding Councils had commissioned a study from Leiden University to produce a model for the bibliometric indicator (THES 2007). Their proposal is to count (not read) every paper produced by every academic and place them in a band according to the impact factor of the journal in which they are published. The number of publications against the impact factor would create a spectrum for each department and the RAE could then be held as frequently as the government decides, even annually. However, a report commissioned by Universities UK (the association of vice chancellors) argued that quality should be measured by the number of times each article is cited, rather than by the 'impact factor' of the journal in which it is published (Corbyn 2007).^{viii}

Initially, it seemed academics in STEM subjects would be prepared to accept the use of bibliometric measures of either citations or journal impact factors. But then some leading scientists tried out the system that would be used for counting citations, which is based on Thomson's Web of Science database. Many science disciplines have an open source database, in which scientists worldwide place their articles as soon as they are accepted by a journal, and which they use as a near-comprehensive source of publications on their subject. The most highly cited astronomy and space scientist in the UK, Carlos Frenk, found that the Thomson Web of Science recorded 5,000 fewer citations for his articles than the open source Astrophysics Data System – a loss of 18 per cent. His colleague, Nigel Glover, found that nearly 1,600 citations for his articles on the particle physics database, Spires, did not register on the Web of Science – a loss of 37 per cent (Corbyn 2008b). This put the validity of the bibliometrics in question, and led to calls for STEM subjects to retain an element of peer review combined with the quantitative measures.

Problems with metrics had already been recognised as making the REF model unsuitable for the humanities and social sciences. An Expert Group set up by the Arts

and Humanities Research Council and the Funding Council for England reported that, while there was no objection to metrics in principle,

There is no single metric that is appropriate to measuring the performance of arts and humanities research. Of the metrics available, some are well established, while others are being developed. The metrics chosen as part of the assessment framework should reflect the multiplicity of peer review systems which are already in place and are an integral feature of academic life (AHRC and HECE 2006).

A report from the British Academy (2007) reinforced this view. In principle, metrics could be a good way to supplement peer review, but there are no existing metrics which capture the quality of publications in the humanities. The report argued that the effects of a move to an inappropriate metrics would be very serious because the funding allocated as a result of the RAE constitutes a much higher proportion of the total dual support funding for the humanities (85 percent) than for medicine and the natural sciences (50 percent), which rely to a much greater extent on grants from the Research Councils.

In the natural sciences, the main form of publication is through journals. Articles are shorter, and the material within them is taken up by the research community very quickly. The British Academy report argues that this sets up a dynamic in journals publication in these disciplines which can be described as a virtuous circle resulting in a 'high impact factor'. The more the articles in a journal are cited, the more authors want to publish in that journal, so the more demanding becomes the refereeing process, and the more demanding the selection criteria, the better the articles will be, and the more they will be cited.

The British Academy identified how the publishing dynamic in the arts and social sciences is quite different from the sciences:

- Publications are evaluated not just in terms of research 'results', but the strength of the argument developed. To set out an argument in depth, books are much more important than edited volumes and journal articles.
- Thomson Scientific measures impact every two years and this reflects the practices in the natural sciences and medicine where specific results are picked up quickly in the literature. Publications in the arts and social sciences often take longer than two years to make an impact (as seen in the cases of many Nobel prize winners).
- Humanities journals publish fewer, longer articles. This makes them vulnerable to small-number problems and a journal's ranking can change erratically and significantly from year to year, often as the result of citations to only one article.

- The ‘best’ articles may not just appear in the biggest, English language journals. There are many more, smaller journals addressing specialist audiences and language communities, which do not meet the ‘international authorship’ and other criteria for inclusion in the Thomson Scientific journals list on which the citation index is based.
- The citation indexes, which add up the number of times a journal article has been cited in other journals, do not include all the forms of citation relevant to the arts and social sciences. Where a book is cited in a journal article, this is included, but not the lists of references found in books, which cite either journal articles or other books. This makes citation indexes extremely faulty for the arts and social sciences (British Academy 2007).

The British Academy report argued that it is not possible to make inferences from journal rank to author quality and any journal ‘impact factor’ would have to take account of all the above points. None of the available rankings of journals or citations indexes do this. The European Science Foundation has tried to overcome some of these difficulties by compiling its own database of scholarly journals in the humanities - a European Reference Index for the Humanities (ERIH) – but the British Academy’s working group is very critical of this initiative. While claiming to be a simple list, it smuggles in quality ranking on unclear criteria, it still gives preference to English journals and widespread distribution rather than scholarly standing, and putative journal esteem seems to be based on the views of very small numbers of people. The report concludes that the ERIH is not a reliable way to construct metrics of peer-reviewed publications (British Academy 2007: 36).

Initially it seemed that the humanities and social sciences were pleased not to be included in the STEM subject’s REF model, but then worries grew that if there were two assessment systems, there might be two funding systems, and it would be harder to keep a check on government shifting funding towards STEM subjects. Meanwhile STEM subjects, doubting the validity of commercial citation indexes, began arguing for a ‘light touch’ peer review informed by metrics, similar to that for the arts and humanities. This was not what the Treasury wanted to hear. Then, just as the Research Funding Council’s consultation was closing, the Secretary of State himself announced that he had ‘thrown a rock into the pond’ because he wanted a fourth measure to be included in the REF model – a reward for academics who provide policy advice to government (Corbyn 2008a). As mentioned above, the RAE itself had always prioritised ‘pure’ over ‘applied’ research and the REF dependence on citations in ‘core’ journals narrowed the definition of ‘excellent’ research further. There are no available measures for applied work, public engagement or policy reports. Shortly after this debacle, the senior civil servant responsible for developing the REF model left the Research Funding Council for a job in university management. The shift from the RAE’s peer review for assessing quality to REF’s metrics model for measuring impact seems to be temporarily in disarray.

Metrics and international rankings

Why is there such pressure in the UK and elsewhere, to govern by numbers? What is the attraction of converting academic work, via a series of proxies, into a number? Especially when this process makes academics focus on fundamental research in academic publications, at the expense of government's other aims for academics to engage with the surrounding society in industrial collaborations, applied research, professional development, and public debate?

One answer is that numbers have a spurious commensurability. They make 'quality' in the natural sciences and in the humanities and social sciences look as if it is equivalent and comparable so as to provide an 'objective' basis for government decisions about funding allocations. Another answer is that grades are easy to convert into league tables and these are not just for national consumption: they are to pit universities against each other in the global market of international students, or they provide a basis for universities to find partners of comparable standing, so as to market themselves collaboratively through international consortiums. Such systems of grading and ranking boil all the work going on in an entire institution down into a single number which is meant to guide people with such diverse interests as businessmen wanting a research partner, to international students wanting a welcoming and supportive environment.

There are now two international league tables and 15 national league tables, as well as numerous specialist league tables (on MBAs, medical degrees, etc). Usher and Savino's (2006) study shows that each of these international league tables is based on very different kinds of data – surveys of stakeholders, data from governments or research councils (often arising as a bi-product of their administrative procedures) and university sources. The authors of each league table select particular information as proxy indicators, assign each indicator a weighting, and produce a final score and ranking. This process imposes a definition of quality on an institution, with no right to query the basis for its construction.

Usher and Savino's (2006) analysis of these league tables shows that no two agree on what constitutes quality. They conclude that the world's top ten universities^{ix} come out on top regardless of the method, but the position of other universities is a 'fluke'

- One common feature, which maybe explains governments' current penchant for merging universities, is that the indicators are not normalised to account for university size, so that bigger universities score better.
- There is no common definition of quality, and these definitions are not converging.

- They use a wide range of indicators for research, teaching, and reputation. None of the 7 basic categories of indicators and no single indicator is common to all the university ranking systems.

The Jiao Tong University at Shanghai produces ‘Academic Ranking of World Universities’. This gives most weight to research (90 percent). This is based on the Thomson ISI citation index, but the Jiao Tong authors consider that this citation index is accurate only for the hard sciences, so they focus mainly on research in those disciplines and take the social sciences and humanities very little into account. The Jiao Tong University ranking of world universities does not take ‘reputation’ into account at all. In contrast, the THES’ ‘World University Rankings’ places 50 percent of its weighting on ‘reputation’ based on the result of a questionnaire sent to selected academics around the world. If governments’ current focus on numerical scores and league tables is partly in order to get their top universities into these two international rankings of the top 100 universities, then maybe they should take a much more strategic approach and just focus on scoring high on natural science publications and on ‘reputation’. Certainly, many US universities have staff whose job it is to produce information for *US News and World Report* that will put their institution in a favourable light. Gamesmanship maybe, but also survival.

The ways in which assessment models, whether based directly on peer evaluations or indirectly on peer reviewing as part of journal production, get tied into government systems for rewarding excellence and punishing failure, and then get embroiled in world rankings, conflate three activities which are of very different kinds. A quest for world rankings, when the models are fluky, calls for sheer gamesmanship. This is a very different exercise from allocating funding across a sector in a way that will sustain vibrant research activity across the whole range of disciplines that make up a ‘university’. Those political decisions are a separate exercise from peer review which aims to advance a discipline by promoting critical debate and dialogue across differences. The problem is that once research activity is turned into a number, it is very easy for this number to migrate across different systems and be used for purposes which make little sense in terms of its point of origin, but which have enormous impacts on the original activity.

Skewing Effects

This analysis shows that systems of grading and ranking research output rely centrally on three elements: processes of peer review, the survival of disciplines as a whole rather than just top departments, and commercial publication and citation indexes. This last section of the paper will consider how the RAE has had the effect of skewing all of these three basic aspects of the infrastructure of academic work. It is probable that these skewing effects would be even greater in a metric system.

Peer review – whole process

The British Academy's study reveals how important peer review is to the whole process of academic production. Even so, they only focused on primary peer reviews (the evaluation of articles submitted to journals and the assessment of project grant applications sent to Research Councils) and secondary peer reviews (departmental reviews and the RAE). They did not include the evaluation of PhD applications for funding or reviews of applicants for jobs or promotion.^x Nor did they include the practice of informally circulating papers for critical comment before submitting them to a journal, or the continuing debate about articles following publication. Peer review is woven into the whole fabric of academic work, in creating conversations, critical responses, and building on each others' ideas, in which journals act as a node point.

The focus on citation indexes skews this process. The British Academy report refers to universities' instructing their academics to focus only on producing articles and sending them to top journals, and to refuse requests to act as a peer reviewer themselves (articles sent for submission, book reviews, grant application reviews). As an editor I can confirm this. Some universities or departments are instructing academics to withdraw from peer reviewing, so that they can send more articles to journals – and contribute to the rising demand for peer review; other academics are just under too much work pressure to cope with such requests. Academics have to respond to what counts in this system where grading and ranking determines economic survival. If managers and governments want peer reviewing to survive (and so much of their decision making relies on this peer reviewing infrastructure) they have to reward the whole process, not just the product. That is, scores should not just be given for articles published in peer reviewed journals, but for fulfilling requests to act as a peer reviewer or being a journal editor too.^{xi}

Effect on disciplines as a whole – Research Funding Council's disciplinary benchmarking

Official reports have expressed concern that the concentration of research funding on top departments is damaging the ability of disciplines to sustain a vibrant research culture and reproduce themselves. However, the system's effects may be much greater than this: they are not just skewing the organisation of a discipline as a whole, but having much wider-ranging social effects. The mathematics discipline exemplifies this point. There is no doubt that research in the top departments is flourishing, yet the discipline is in crisis and threatening social mobility.

English mathematicians are contributing solutions to problems that have been unresolvable for ages: two of the six most recent winners of the Fields medal (the equivalent to the Nobel prize) have come from the UK, and British-based blue-skies research in mathematics is prominent in 'many of the great technological advances of the 20th century' (London Mathematical Society 2005: 3; see also Stothart 2006).

Figure 2. Mathematics results 2001 RAE

RAE 2001	Pure Maths departments		Applied Math departments	
	No	%	No	%
5*	4	9	6	10
5	25	53	25	43
4	5	11	10	17
3 and below	13	27	17	29
Totals	47	100	58	99

Source: RAE (2001), 'Research Assessment Exercise Results and Overview Reports (for Pure Mathematics and Applied Mathematics in the UK)', *Higher Education and Research Opportunities in the UK*

The 2001 RAE results show that just over half of all mathematics departments gained RAE grades of 5 and 5* (Figure 2) and, in keeping with the national pattern, research funding has been concentrated in these departments. Departments with lower grades, while still recognised as doing research of national significance, are often in financial difficulties as a result of the government's funding mechanisms. Five universities have closed their mathematics departments since 1999 (e.g. at Hull University in 2005) others have merged mathematics with other departments or given mathematics a service function for engineering or other subjects. Eight percent of bachelor degree courses in mathematics have closed in the ten years 1998-2007 (UCU 2006: 4). As a result, there are 'regional deserts' – whole regions of Britain where there are very few mathematics degree courses (London Mathematical London 2005: 4). For example, the provision of single honours mathematics and science degree courses in the east and north-east of England fell by nearly one third from 1998 to 2007 (UCU 2006: 5). This hits the poorest students hardest, because it is these students who can only afford to study if they stay at home. (It is of course these students, from poor and ethnic minority families with no history of going to university, that the government wants to gain access to higher education and professional employment under the 'widening participation' initiative). This is a contributory factor to the shortage of maths school teachers in Britain, reported by a government-funded inquiry as a shortfall of 3,400 in 2004 (Smith 2004). This shortage has travelled through to a dearth of school pupils taking A-level and of university students studying maths, and now a shortage of academics. The 2001 RAE report for pure mathematics stated that 60 percent of new appointments as university lecturers over the previous five years were from Eastern Europe and Germany, and the report for applied mathematics stated that 'most' new appointments were from overseas and there was a 'concern for indigenous talent' (RAE2001 2001).

The effect of the RAE in skewing disciplines, ranking departments, focusing funding on the ‘top’ and threatening the survival of research in the ‘bottom’ is now being taken seriously by the Economic and Social Science Research Council (ESRC). They have instituted a process of benchmarking a whole discipline against international comparisons. They make clear that the aim is to judge the international standard of the whole discipline and not to rank individuals and departments, or pit them against each other, or even convert the discipline as a whole into a number. Their aim is to assess the strengths that should be encouraged and the weaknesses that need support.

The ESRC started this process with anthropology, both because anthropology is of its essence an international discipline, and also because the ESRC asked the renowned anthropologists on the international benchmarking panel and on the UK steering group to devise a qualitative methodology that other disciplinary panels could take up and adapt for their reviews. The ESRC was extremely pleased with the outcome. The report

- highlighted the 14 research areas in which British anthropology is a world leader or significant player, and the four areas where more could be done.
- analysed the discipline’s research capacity – anthropology is not a school subject, so universities have to create their own constituency of students from scratch; the shortage of funding for UK PhDs and the model of doctoral education; enthusiasm about the provision of postdoctoral funding; the strength of the market for anthropology outside the academy.
- Assessed the impact of the discipline on policy and practice
- Looked to the future, regarding funding, infrastructure, the development of assessment in more disciplinarily-tailored ways and 6 recommendations. (Brenneis 2006).

This qualitative analysis of a whole discipline provides academics, managers, and policy makers with clear information on which to base a dialogue and plan developments. It makes possible the maximisation of collective effort. It is especially valuable for a small discipline within the British context, and it might well be a valuable model to use in a country such as Denmark, with a small university sector in a world context. Such a holistic approach, which assesses the strengths and weaknesses of a discipline, makes it possible to take into account not just the strengths of the top departments, but the ability of the discipline as a whole to contribute to the economy and society.

Corporate costs

When university managers and governments adopt matrices based on citation indexes as one of their main steering technologies – and quickly academics learn what counts – how does this affect the economics of publishing? And how does that affect the economics of universities? How is such a move compatible with governments’

injunctions that universities should act more like corporations and measure efficiency in terms of profit?

What, for example, are the corporate costs of the pressure on academics to publish in 'core' or 'high impact' journals? The British Academy's report estimated that an article in the humanities and social sciences takes about 4.5 hours to review, and at a minimum salary level of £100 per hour, this means, if each article is reviewed by two people, academics contribute £900-worth of their time to the journal. This is about equal to the publisher's cost of publishing each article. In other words, universities are subsidising the costs of producing each journal by at least half – and more if the time of the editor her/himself is taken into account. The University of California system calculated the contribution of free labour it makes to one journal publisher alone, Reed-Elsevier. They did not include the cost of faculty time spent in peer reviewing, but they worked out that 150 UC faculty were editors of journals, a further 964 served on editorial boards and 15 percent of the content of the journals had been written by UC faculty.

The extent of this free donation of academic work to commercial publishers can be glimpsed, if it is considered that there are about 20-25,000 peer reviewed scholarly journals worldwide, and growing at a compound annual rate of 3-4 percent. For example, the British Academy's report quotes a study of 174 publishers, which shows they launched 1,048 new journals in 2000-5 (British Academy 2007: 4). The number of articles submitted is growing likewise. This means more academics are using their time as editors and peer reviewers. Of course, academics are not seeking full economic costs for peer reviewing - the British Academy's report also shows the peer review system only exists because of professional commitment to contribute to the academic public good. But universities should be aware of the implications of their emphasis on metrics for the use of academic time.

Penny Ciancanelli (2008) points out that commercial publishers have always been involved in journal publishing, and since the start of the Science Citation Index in the 1960s, they have concentrated on 'core' journals which self-respecting university librarians could not be without. What is new, she argues, is the size of the publishing firms and the corresponding scale of the profit accumulated. Four commercial publishers (Reed-Elsevier, Candoover-Cinven, Thomson, Wiley) now own most of the 'core' academic journals not owned by learned societies. Reed-Elsevier publishes 25 percent of core science publication and is the largest journal publisher in the world. It earned profits of 37 percent on its publishing business in 2005 and the shift to digital access in 2007 increased profitability by 20 percent. Elsevier offers universities portfolio licensing agreements, in which, to gain access to the prestigious, high use titles, universities also have to subscribe to lower end journals that they would prefer not to buy.

According to Ciancanelli (2008), between 1986 and 2002 expenditures by US university libraries increased by 220 percent but the number of journals purchased

increased by only 9 percent. This reflects the market power of a few commercial publishers. The University of California is Reed-Elsevier's second largest client. Its 10 campuses serve 208,000 students, with 121,000 full time academics. In 2002/3 the University of California paid Reed-Elsevier \$8m for digital access to 1,700 journals and \$2m for print copies of journals. This was half of the University of California's journals budget. But these journals accounted for only 25 percent of total usage. Even though a large proportion of what Reed Elsevier sells was created, vetted or enhanced by UC faculty, the university had to buy it back at increasingly unaffordable prices.

This is an example that could be documented for many other universities around the world. In sum, when research assessment systems place over-riding emphasis on publication in peer-reviewed journals, this leads to an exponential growth in commercial publishing of scholarly journals. Universities use their public funding to subsidise these publishers' commercial profits twice over. First universities pay the salary costs of academics to write the articles, edit and peer review them. All this work is given freely to publishers, presumably under the aegis of the university's 'public service'. Second, the universities use their library budgets to buy back this work at highly inflated and fast escalating prices. The university is forced in this instance to act like a corporate customer – publishers offer no reciprocal public service. When governments insist on basing calculations of research impact, or, erroneously, quality, on commercial publishers' citation indexes, this further strengthens the profits of the owners of these indexes and of the 'high-impact' journals (whereas discipline-based open source databases offer a free and potentially more accurate alternative). Governments' metrics-based research assessments systems, as currently conceived, will only increase this double subsidy of the four big publishers' commercial profits, paid for by universities through academic time and library budgets. Even, or especially, in terms of the government's dominant economic rationality, surely that does not meet the basic criterion of 'efficiency' in the use of public resources?

Conclusion

There is a clear need for universities to be accountable to society for their public funding, and to show that they are providing a public benefit. This mandate is becoming skewed, however, when governments (like the UK's Thatcher government in the early 1980s) redefine the purpose of universities as to serve the economy – and claims they have 'failed' the economy. The more recent shift to considering universities themselves to be economic actors in the economy, and indeed 'drivers' of the global knowledge-based economy, shifts their remit further away from 'public service'. Yet this shift has not been accompanied by an overall analysis of universities as economic entities. In the absence of this overview, and in this limbo between public service and profitable corporation, does the RAE improve accountability? The evidence suggests not, but that academic work is becoming skewed by systems that determine 'what counts'.

The Research Funding Council for England itself made the point (quoted above) that any assessment system that is taken seriously by the people it assesses will distort the system. Power (1997) has made the point more generally that audit processes transform organisations into their own image. Official inquiries and reports in Britain have documented the ways that the RAE has distorted academic work, and on the basis of this experience, they can analyse the likely distortions that the proposed metrics-based system would bring about. They identify the effects on research practices, the need to focus on ‘safe’ research which guarantees results that can be quickly published. In the British assessment and funding environment, it poses an unacceptable risk to departments if their members engage in blue-skies ‘revolutionary’ science (Kuhn 1962) that might shift scientific paradigms, but might also yield no immediately publishable results that ‘count’. The RAE has clearly increased the volume of scientific output, but the increase is of ‘normal’ science at the expense of ‘revolutionary’ science (Charlton and Andras 2008; Charlton 2008). Official reports also record that this increase in scientific output is due to changes in publication practices, notably the pressure to ‘rush to press’ so as to create publications whilst the research is being formulated and conducted and before the results are all analysed, and to ‘salami slice’ a set of results into several publications. There is also considerable concern in official publications about the effects of the assessment system and funding regime on the survival of particular disciplines. The system relies so heavily on buying and keeping ‘top’ researchers and sloughing off anyone else who cannot perform continuously at this level. This creates a structural disadvantage for academics engaged in reproduction and parenting, or who through illness or an interest in applied work do not focus all their efforts all the time on the kinds of publications that count. It is extremely hard for such people to get back on the research ladder. By concentrating research funding on the ‘top’ departments, this puts in jeopardy other departments which are also sites of very good research, and are crucial for nurturing talent that will eventually move to the top departments. By concentrating funding on the top universities, this differentiates the sector, between an elite with good research conditions and others which are under such financial pressure they can never aspire to joining the elite. It is also leading to the separation of research from teaching. These distortions are well-documented as side effects of making academics focus on what counts in terms of government policy.

The paper takes the analysis further. It argues that assessment systems such as the RAE and its associated financial regime do not just distort research: they are skewing the very nature of academia itself. First, by pressing academics to focus their time and energy on publishing in peer reviewed journals and not doing anything else – not even engage in peer reviewing themselves – this is skewing the whole process of peer dialogue on which academic work relies. Second, by concentrating research funding on top departments and letting others close or end their degree programmes, this does not just affect research but the whole role of universities in society. The case of mathematics can be documented also for other STEM subjects, especially chemistry and physics, and for West and East Asian languages. ‘Regional deserts’ for such

subjects restrict access to university for especially ethnic minority and socially disadvantaged students who can only afford to study if they stay at home. There sets in a shortage of teachers, a shortage of school leavers with the qualifications and enthusiasm to read those subjects at university, and ultimately, as the mathematics case showed, universities rely on global brain drain to fill their staffing requirements. The ESRC has gone some way to recognise these problems by establishing its new method of reviewing the international standing, health and development needs of whole disciplines. But the government's focus continues to be on ensuring a few elite universities are well placed in global rankings. The way this skews the role of universities in society puts in question the government's other aim of producing a highly educated workforce that will succeed in the global knowledge economy. Third, the focus on citations in high impact journals, most of which are owned by four commercial publishers is skewing university finances and the use of public funds. Such a narrow definition of what counts in both the RAE and the proposed REF leads to university budgets being used to provide a double subsidy for commercial profits. The aim of such assessment systems is to make academics accountable for their public funding, but the effect is a serious skewing of the use of public funding and its diversion into private profits. Academics and universities must be publicly accountable for the resources they receive from society, but systems of accountability need to be based on a much more holistic analysis of their wider economic and social effects. An approach of the kind started here is needed in order to review what really should count if the aim is to generate universities that use public funding efficiently to wide social benefit, based on vibrant and self-sustaining academic practices.

References

- Academy, British (2007), 'Peer Review: the challenges for the humanities and social sciences', (London: The British Academy).
- Adams, Jonathan and Smith, David (2006), 'Evaluation of the British Research Assessment Exercise', in Leon Bakker, et al. (eds.), *Evaluating the Performance-Based Research Fund. Framing the Debate* (Wellington, NZ: Institute of Policy Studies, Victoria University).
- AHRC and HECE (2006), 'Use of research metrics in the arts and humanities. Report of the Expert Group set up jointly by the Arts and Humanities Research Council and the Higher Education Funding Council for England'.
- AUT (2004), 'Academic staff 2002-3 - Gender and research activity in the 2001 RAE', (London: Association of University Teachers).
- Bakker, Leon, et al. (eds.) (2006), *Evaluating the Performance-Based Research Fund* (Wellington: Institute of Policy Studies, Victoria University).
- Baty, Phil (2005), 'Cardiff to take part in audit of stress', *Times Higher Education Supplement*, 1 April.
- (2006), '41% would like next exercise to be halted', *Times Higher Education Supplement*, 13 October.
- Brenneis, Don (2006), 'International Benchmarking Review of UK Social Anthropology: An International Assessment of UK Social Anthropology Research', (Economic and Social Research Council, Association of Social Anthropologists, Royal Anthropological Institute).
- Calabrese, Raymond and Roberts, Brian (2004), 'Self-interest and scholarly publication: the dilemma of researchers, reviewers, and editors', *The International Journal of Educational Management*, 18 (6), 335-41.
- Charlton, Bruce (2008), 'Oxford University = University of Minnesota. What has the RAE ever done for Oxford University?' Improvement in normal science – but decline in revolutionary science.' *Oxford Magazine*, 271, 3-5.
- Charlton, Bruce and Andras, P. (2008), '“Down-shifting” among top UK scientists? - The decline of ‘revolutionary science’ and the rise of ‘normal science’ in the UK compared with the USA', *Medical Hypotheses*, 70, 465-72.
- Ciancanelli, Penny (2008), '(Re)producing universities: Knowledge dissemination, market power and the global knowledge commons', in Deborah Epstein, et al. (eds.), *Geographies of Knowledge, Geographies of Power. World Yearbook of Education* (London: Routledge).
- Corbyn, Zoe (2007), 'New RAE based on citations', *Times Higher Education Supplement*, 9 November.
- (2008a), 'Denham wants policy advice credited in REF', *Times Higher Education*, 24 January.
- (2008b), 'Researchers raise questions over validity of REF citations criteria', *Times Higher Education*, 10 January, p. 11.

- Dearing, Ronald (1997), 'Higher Education in the Learning Society. Summary Report of the National Committee of Inquiry into Higher Education', (Norwich: HMSO.).
- Elton, Lewis (2000), 'The UK Research Assessment Exercise: Unintended Consequences', *Higher Education Quarterly*, 54 (3), 274-83.
- Fazackerley, Anna (2006), 'Treasury focus on top 10', *Times Higher Education Supplement*, 13 October.
- Frølich, Nicoline (2008), 'The Politics of Steering by Numbers. Debating Performance-Based Funding in Europe', (Oslo: NIFU-STEP).
- Hansen, Hanne Foss and Borum, Finn (1999), 'The construction and standardization of evaluation. The case of the Danish university sector', *Evaluation*, 5 (3), 303-29.
- HEFCE (2006), 'Response to consultation on successor to research assessment exercise', (Higher Education Funding Council for England and Department for Education and Skills).
- Hodgson and Rothman (1999), 'The editors and authors of economics journals: a case of international oligopoly?' *The Economic Journal*, 109, 165-86.
- House of Commons, Science and Technology Committee (2004), 'Research Assessment Exercise: a re-assessment', *Eleventh Report of Session 2003-4* (London: House of Commons).
- Kinman, Gail, Jones, Fiona, and Kinman, Russell (2006), 'The well-being of the UK academy, 1998-2004', *Quality in Higher Education*, 12 (1), 15-27.
- Kuhn, Thomas (1962), *The Structure of Scientific Revolutions* (Chicago: University of Chicago Press).
- Lambert, Richard (2003), 'Lambert Review of Business-University Collaboration'. London, Mathematical Society (2005), 'Submission to the [House of Commons] Select Committee on Science and Technology's Inquiry into Strategic Science Provision', (London: House of Commons).
- Manna, Manfredi La (2004), "'Superstar" rating is not the issue', *Times Higher Education Supplement*, 16 April, p. 13.
- McNay, Ian (1996), 'The impact of the 1992 RAE on institutional and individual behaviour in English higher education: the evidence from a research project', (Bristol: HEFCE).
- (1999), 'The paradoxes of research assessment and funding', in Mary Henkel and Brenda Little (eds.), *Changing Relationships between Higher Education and the State* (London: Jessica Kingsley).
- Power, Michael (1997), *The Audit Society. Rituals of Verification* (Oxford: Oxford University Press).
- RAE2001 (2001), 'Research Assessment Exercise Results and Overview Reports (for Pure Mathematics and Applied Mathematics in the UK)', *Higher Education and Research Opportunities in the UK* (<http://www.hero.ac.uk/rae/>), accessed 10 January 2007.
- RAE2008 (2007), 'Research Assessment Exercise. Quality profiles', (<http://www.rae.ac.uk/aboutus/quality.asp> accessed 25.1.2007).
- Roberts, Gareth (2003), 'Review of Research Assessment', *RA Review*.

- Scurr, Ruth (2007), 'When filling out RAE entry forms, the first question should be: Have you reproduced in the past seven years?' *Times Higher Education Supplement*, 24 August.
- Sheehan, Peter (2006), 'A perspective on the proposed Australian Research Quality Framework', in Leon Bakker, et al. (eds.), *Evaluating the Performance-Based Research Fund* (Wellington, New Zealand: Institute for Policy Studies, Victoria University).
- Shore, Cris and Wright, Susan (2000), 'Coercive accountability: the rise of audit culture in higher education', in Marilyn Strathern (ed.), *Audit Cultures. Anthropological Studies in Accountability, Ethics and the Academy (EASA Series)* (London: Routledge).
- (2001), 'Audit culture and anthropology', *Journal of the Royal Anthropological Institute*, 7 (4), 759-63.
- Smith, Adrian (2004), 'Making Mathematics Count, the Report of Professor Adrian Smith's Inquiry into Post-14 Mathematics Education', (London: The Stationery Office Ltd).
- Stothart, Chloe (2006), 'Zeros make heroes', *Times Higher Education Supplement*, 15 December, p. 6.
- Tahir, Tariq (2007), 'Durham don sent on leave', *Times Higher Education Supplement*, 2 November.
- THES (2007), 'All work to be counted in RAE evaluation', *Times Higher Education Supplement*, 12 October, p. 1 and 7.
- Tysome, Tony (2006), 'UK boosts its tally of Nobel laureates', *Times Higher Education Supplement*, 27 October 2006.
- UCU (2006), 'Degress of Decline? Core Science and Mathematics Degree Courses in the UK 1998-2007', (London: University and College Union).
- Usher, Alex and Savino, Massimo (2006), 'A World of Difference: A Global Survey of University League Tables', *Canadian Education Report Series* (Toronto: Educational Policy Institute).
- Velody, Irving (1999), 'Knowledge for what? The intellectual consequences of the Research Assessment Exercise, Special Issue', *History of the Human Sciences*, 12 (4), 111-46.
- .
- Whittington, G. (1997), 'The 1996 Research Assessment Exercise', *British Accounting Review*, 29, 181-97.
- Wright, Susan (2004), 'Markets, corporations, consumers? New landscapes of higher education', *Learnng and Teaching in the Social Sciences*, 1 (2), 71-93.

Notes

- Academy, British (2007), 'Peer Review: the challenges for the humanities and social sciences', (London: The British Academy).
- Adams, Jonathan and Smith, David (2006), 'Evaluation of the British Research Assessment Exercise', in Leon Bakker, et al. (eds.), *Evaluating the Performance-Based Research Fund. Framing the Debate* (Wellington, NZ: Institute of Policy Studies, Victoria University).
- AHRC and HECE (2006), 'Use of research metrics in the arts and humanities. Report of the Expert Group set up jointly by the Arts and Humanities Research Council and the Higher Education Funding Council for England'.
- Bakker, Leon, et al. (eds.) (2006), *Evaluating the Performance-Based Research Fund* (Wellington: Institute of Policy Studies, Victoria University).
- Baty, Phil (2006), '41% would like next exercise to be halted', *Times Higher Education Supplement*, 13 October.
- Brenneis, Don (2006), 'International Benchmarking Review of UK Social Anthropology: An International Assessment of UK Social Anthropology Research', (Economic and Social Research Council, Association of Social Anthropologists, Royal Anthropological Institute).
- Calabrese, Raymond and Roberts, Brian (2004), 'Self-interest and scholarly publication: the dilemma of researchers, reviewers, and editors', *The International Journal of Educational Management*, 18 (6), 335-41.
- Charlton, Bruce (2008), 'Oxford University = University of Minnesota. What has the RAE ever done for Oxford University?' Improvement in normal science – but decline in revolutionary science.' *Oxford Magazine*, 271, 3-5.
- Charlton, Bruce and Andras, P. (2008), 'Down-shifting' among top UK scientists? - The decline of 'revolutionary science' and the rise of 'normal science' in the UK compared with the USA', *Medical Hypotheses*, 70, 465-72.
- Ciancanelli, Penny (2008), '(Re)producing universities: Knowledge dissemination, market power and the global knowledge commons', in Deborah Epstein, et al. (eds.), *Geographies of Knowledge, Geographies of Power. World Yearbook of Education* (London: Routledge).
- Corbyn, Zoe (2007), 'New RAE based on citations', *Times Higher Education Supplement*, 9 November.
- (2008a), 'Denham wants policy advice credited in REF', *Times Higher Education*, 24 January.
- (2008b), 'Researchers raise questions over validity of REF citations criteria', *Times Higher Education*, 10 January, p. 11.
- Dearing, Ronald (1997), 'Higher Education in the Learning Society. Summary Report of the National Committee of Inquiry into Higher Education', (Norwich: HMSO.).

- Elton, Lewis (2000), 'The UK Research Assessment Exercise: Unintended Consequences', *Higher Education Quarterly*, 54 (3), 274-83.
- Fazackerley, Anna (2006), 'Treasury focus on top 10', *Times Higher Education Supplement*, 13 October.
- Frølich, Nicoline (2008), 'The Politics of Steering by Numbers. Debating Performance-Based Funding in Europe', (Oslo: NIFU-STEP).
- Hansen, Hanne Foss and Borum, Finn (1999), 'The construction and standardization of evaluation. The case of the Danish university sector', *Evaluation*, 5 (3), 303-29.
- HEFCE (2006), 'Response to consultation on successor to research assessment exercise', (Higher Education Funding Council for England and Department for Education and Skills).
- Hodgson and Rothman (1999), 'The editors and authors of economics journals: a case of international oligopoly?' *The Economic Journal*, 109, 165-86.
- House of Commons, Science and Technology Committee (2004), 'Research Assessment Exercise: a re-assessment', *Eleventh Report of Session 2003-4* (London: House of Commons).
- Kuhn, Thomas (1962), *The Structure of Scientific Revolutions* (Chicago: University of Chicago Press).
- London, Mathematical Society (2005), 'Submission to the [House of Commons] Select Committee on Science and Technology's Inquiry into Strategic Science Provision', (London: House of Commons).
- Manna, Manfredi La (2004), "'Superstar" rating is not the issue', *Times Higher Education Supplement*, 16 April, p. 13.
- McNay, Ian (1996), 'The impact of the 1992 RAE on institutional and individual behaviour in English higher education: the evidence from a research project', (Bristol: HEFCE).
- (1999), 'The paradoxes of research assessment and funding', in Mary Henkel and Brenda Little (eds.), *Changing Relationships between Higher Education and the State* (London: Jessica Kingsley).
- Power, Michael (1997), *The Audit Society. Rituals of Verification* (Oxford: Oxford University Press).
- RAE2001 (2001), 'Research Assessment Exercise Results and Overview Reports (for Pure Mathematics and Applied Mathematics in the UK)', *Higher Education and Research Opportunities in the UK* (<http://www.hero.ac.uk/rae/>), accessed 10 January 2007.
- RAE2008 (2007), 'Research Assessment Exercise. Quality profiles', (<http://www.rae.ac.uk/aboutus/quality.asp> accessed 25.1.2007).
- Roberts, Gareth (2003), 'Review of Research Assessment', *RA Review*.
- Sheehan, Peter (2006), 'A perspective on the proposed Australian Research Quality Framework', in Leon Bakker, et al. (eds.), *Evaluating the Performance-Based Research Fund* (Wellington, New Zealand: Institute for Policy Studies, Victoria University).
- Shore, Cris and Wright, Susan (2000), 'Coercive accountability: the rise of audit culture in higher education', in Marilyn Strathern (ed.), *Audit Cultures*.

- Anthropological Studies in Accountability, Ethics and the Academy (EASA Series)* (London: Routledge).
- (2001), 'Audit culture and anthropology', *Journal of the Royal Anthropological Institute*, 7 (4), 759-63.
- Smith, Adrian (2004), 'Making Mathematics Count, the Report of Professor Adrian Smith's Inquiry into Post-14 Mathematics Education', (London: The Stationery Office Ltd).
- Stothart, Chloe (2006), 'Zeros make heroes', *Times Higher Education Supplement*, 15 December, p. 6.
- Tahir, Tariq (2007), 'Durham don sent on leave', *Times Higher Education Supplement*, 2 November.
- THES (2007), 'All work to be counted in RAE evaluation', *Times Higher Education Supplement*, 12 October, p. 1 and 7.
- Tysome, Tony (2006), 'UK boosts its tally of Nobel laureates', *Times Higher Education Supplement*, 27 October 2006.
- UCU (2006), 'Degress of Decline? Core Science and Mathematics Degree Courses in the UK 1998-2007', (London: University and College Union).
- Usher, Alex and Savino, Massimo (2006), 'A World of Difference: A Global Survey of University League Tables', *Canadian Education Report Series* (Toronto: Educational Policy Institute).
- Velody, Irving (1999), 'Knowledge for what? The intellectual consequences of the Research Assessment Exercise, Special Issue', *History of the Human Sciences*, 12 (4), 111-46.
- .
- Whittington, G. (1997), 'The 1996 Research Assessment Exercise', *British Accounting Review*, 29, 181-97.
- Wright, Susan (2004), 'Markets, corporations, consumers? New landscapes of higher education', *Learning and Teaching in the Social Sciences*, 1 (2), 71-93.

ⁱ Activities included the conference at which an earlier version of this paper was presented. The conference 'Research Quality and Quality Indicators' was organised by the Danish Social and Economic Research Council and held on 30 October 2007 at FUUH Conference Centre, Copenhagen.

ⁱⁱ 'Orientering om arbejdet med en kvalitetsbaseret model for fordeling af basismidler – status pr. 26. november' Notat from Direktørsekretariatet, dated 26 November 2007
<http://forskning.hum.ku.dk/upload/application/pdf/f51d6748/Kvalitetsmodel%20status%2026%20nov.pdf>

Accessed 22 April 2008.

ⁱⁱⁱ Copenhagen University's Development Contract 2006-8, dated April 2006 page 2

<http://www.ubst.dk/institutioner-og-okonomi/udviklingskontrakter-resultatkontrakter/KU%20udviklingskontrakt%202006-2008.pdf>

accessed 22 April 2008. The university sector is steered by means of a chain of contracts – the Ministry has a performance contract with its own Department for University and Buildings Administration, then there is a contract between the Minister and the Chair of the Governing Board of each university, which sets out performance targets for each of the university's main activities, and similar contracts

between the Rector and each Dean of Faculty, and between the Dean of Faculty and each Head of Department (Wright and Ørberg 2008).

^{iv} See for example, 'Faculty Contract entered into between the Rector and the Social Science Faculty – implementing Copenhagen University's Development Contract for 2006-8' dated November 2006, page 5. http://cms.ku.dk/upload/application/pdf/16c15579/fakultetskontrakt_SAMFI11.pdf accessed 22 April 2008.

^v Initially the scale was 1 to 5 but in the 1996 RAE a 5* was introduced for particularly internationally strong departments and the middle grade was divided into 3a and 3b (Whittington 1997: 181).

^{vi} SPRU says there is no evidence that further concentration of resources on a few large departments in large universities will result in any superior efficiency Science and Technology Committee House of Commons, 'Research Assessment Exercise: A Re-Assessment', *Eleventh Report of Session 2003-4* (London: House of Commons, 2004). P. 37)

^{vii} This is similar to the system currently proposed for Denmark.

^{viii} The 'impact factor' of journals is notoriously difficult to measure because Thomson's and other databases only include certain journals in their reckoning, whereas articles may be cited much more widely, including in non-English language journals. A study of economics journals revealed that the 'ranking' of a journal was largely a projection of the elite status of the editors' universities, rather than an objective measure of the quality of the ideas published, see Hodgson and Rothman, 'The Editors and Authors of Economics Journals: A Case of International Oligopoly?' *The Economic Journal*, 109 (February 1999), 165-86. I am grateful to Penny Ciancanelli for these points.

^{ix} Oxford, Cambridge, Harvard, Yale, Princeton, MIT, Stanford, Peking, Tsinghua, Toronto.

^x These reviews in Britain are less formalised processes than tenure review in the US or the evaluation of publications of applicants for jobs in Denmark.

^{xi} *The British Medical Journal* (BMJ) now issues certificates for academics to use as proof of their role as a reviewer.

Previous Working Papers in this series:

1. Setting Universities Free? The background to the self-ownership of Danish Universities, Jakob Williams Ørberg, July 2006.
2. Trust in Universities - Parliamentary debates on the 2003 university law, Jakob Williams Ørberg, October 2006.
3. Histories of RUC - Roskilde University Centre, Else Hansen, November 2006.
4. An Insight into the History of the 2003 University Law and some of the Ideas that Surrounded It – With a special focus on changes to the management structure of Danish Universities, the changing role of universities in society, and the relationship between the different political initiatives that were introduced, Peter Brink Andersen, November 2006.
5. Who Speaks for the University? - Legislative frameworks for Danish university leadership 1970-2003, Jakob Williams Ørberg, May 2007
6. 'After Neoliberalism'? - The reform of New Zealand's university system, Cris Shore, June 2007
7. Women in Academia - Women's career paths in the social sciences, in the context of Lund University and Swedish higher education, Susan Wright, October 2007
8. Will market-based ventures substitute for government funding? - Theorising university financial management, Dr. Penny Ciancanelli, May 2008