

## Chapter 4

### **Impacts of performance-based research funding systems: A review of the concerns and the evidence**

*by*

*Linda Butler*

*Australian National University and University of Newcastle, Australia*

*This chapter examines the impacts of performance-based research funding systems. It discusses the difficulties of measuring impacts, particularly in distinguishing between intended and unintended consequences and in establishing whether outcomes are desirable or not. It also explains why presenting an evidence-based assessment is a challenge.*

## Introduction

Assessing the impact of performance-based research funding systems (PRFS) is a fraught exercise, which perhaps explains the paucity of broad authoritative texts on the subject. The literature is full of words like “likely”, “potential”, and “possible”, but contains relatively few concrete examples that examine the impact of PRFS in detail, either through investigative data analysis or well-structured survey/qualitative investigations.

This deficit is acknowledged in *The Changing Governance of the Sciences* (Whitley and Gläser, 2008), which aims to:

“... set the stage for sociological research on RES [research evaluation systems] by outlining what is, and what should be, known about this major institutional innovation in the science policies of many countries. The contributions propose what is sociologically important about these RES and sketch the ways in which these phenomena could be investigated.” (Gläser, 2008, p. 246)

Policy analysts and researchers face a number of conundrums when attempting to identify responses to PRFS:

- How does one attribute causality? (Gläser *et al.*, 2002, p. 20)
- Is it the specific system that leads to particular responses, or simply the existence of an assessment regime, *any* assessment regime? Both peer review and formula-based systems can lead to increased output, institutional strategising, etc. (Gläser, 2008, p. 257)
- Are the changes in behaviour positive or negative? This can depend on the stakeholder group, the discipline, etc. For example, the debate on “transfer markets” in leading researchers just prior to each Research Assessment Exercise (RAE) cycle can be regarded either as game playing, or as driving the sector to provide salaries that are just rewards for achievements.
- What are the mediating effects of parallel reward systems? Requirements for tenure, appointments, promotions, targeted project funding, etc., all have their own requirements, which can be at odds with the signals sent by PRFS and can lead to varying responses in different countries, even for basically similar systems (Whitley, 2008, p. 24).

Yet research is seen as one of the major features determining an institution's reputation (European Commission, 2010, p. 56), and an understanding of the impact of any new, and existing, policy instrument is vital for its effective operation.

When the UK Department for Education and Skills (DfES) called for responses to a consultation document on a proposed reform of higher education research assessment and funding in 2006, it found that the question on “possible undesirable behavioural consequences” attracted the most comments, both from answers to the question directly related to the issue, and in more general statements. These implied that perverse behaviours were likely, whatever the system used (metrics or peer review) (DfES, 2006, para. 41).

This chapter consists of five sections. Following this general introduction, the intended consequences of PRFS are briefly surveyed, as this topic is covered in greater detail in Chapter 1 of this volume. The following section draws out existing evidence on the most common impacts generally regarded as unintended outcomes that are attributed to PRFS. Next, some related issues are touched upon, followed by a final brief conclusion.

A number of important distinctions are made throughout the course of this review. It covers both the intended, and the unintended, consequences of the introduction of PRFS across a range of countries. If the research sector responds to PRFS such that the government's underlying objectives are achieved, there can be little argument that this is positive. As Chapter 1 covers in detail the rationales behind the introduction of various systems, the main focus of this chapter is their unintended consequences. Actors in the sector respond in a variety of ways, and many responses are not anticipated by those designing the systems. It is these behavioural changes which are to be examined: What are they? And are they positive or negative?

The chapter also tries to distinguish between reality and perception. Many commentators on the “evils” of a particular system (and it is usually the opponents of a system who are most vocal) are merely postulating what they believe has, or could have, occurred, rather than presenting evidence on what has in fact occurred. McNay's study of the RAE demonstrated this clearly: while the number of staff who reported that they had moved away from interdisciplinary work was relatively small, almost half of those surveyed felt that the RAE hindered interdisciplinary work (McNay, 1998, p. 20). There was no doubt that many academics perceived the UK system to favour discipline-based research, but alternative policy drivers outside the RAE were clearly ensuring that interdisciplinary work continued to thrive.

A distinction is also made between comments that are evidenced-based and those that are purely anecdotal. This raises one of the most difficult aspects of this review. Given the importance of the issue, very little research has sought to identify and study responses to the introduction of PRFS. Gläser and colleagues have been particularly critical of the failure of researchers in science and technology studies (and the sociology of science) to study the effects of these systems (Gläser, 2008, p. 246; Gläser *et al.*, 2002, p. 3). There are relatively few studies to which one can refer for evidence:

- Research Information Network (RIN)/Joint Information Systems Committee survey of 800 UK academics (RIN, 2009).
- A study for the Higher Education Funding Council of England (HEFCE) – focus groups of senior academics and administrators, survey of heads of departments or research units, staff survey, interviews with limited range of users and funders of research (McNay, 1998).
- A second study for HEFCE focusing on the effect of the Research Assessment Exercise on interdisciplinary research, which surveyed 5 505 researchers (37% response rate) and 327 RAE panel members (62% response rate) (Evaluation Associates Ltd, 1999).
- The Roberts' review of the RAE which encompassed 420 consultation responses; 44 consultative meetings, six open public meetings; and nine workshops (Roberts, 2003).
- A Norwegian bibliometric analysis (Sivertsen, 2008).
- An Australian bibliometric analysis done in the Research Evaluation and Policy Project (REPP) at the Australian National University (Butler, 2003).
- A review of the impact of league tables based on a survey of 134 heads of institutions, with a 68% response rate (92) (CHERI & Hobsons Research, 2008).

It is clear that the bulk of evidence is based on the United Kingdom's RAE. This is not surprising given its much longer history, nor is it of great concern, even when, as here, the focus is on a much broader range of PRFS. As Gläser notes, vastly different systems can have quite similar impacts (Gläser, 2008, p. 256). Actors can only respond to the stimuli they receive in certain ways.

Both institutional and individual responses to PRFS will be examined here. Those responses are not necessarily the same, even within a single institution, as messages from university management can be filtered through several intermediaries before they reach individual researchers. Even when they are the same, researchers may well receive contradictory signals as they seek to obtain external funding for their research or face promotions and appointments committees.

Some literature on university rankings is also covered. While these have little direct impact on individual researchers, institutions respond strongly. It is, to a certain extent, a funding response: rankings equate with prestige, which in turn enables universities to attract more students and more interest from end users. Both translate into more funding.

A number of evaluation systems are meant to cover both universities and other types of research institutions, such as government research agencies (Australia), and all other research institutions (France) (European Commission, 2010, p. 96). However, this review shows that the overwhelming majority of the literature, and specifically studies with an evidence base, focus squarely on higher education institutions.

## Intended consequences

Some common themes run through the rhetoric around the introduction of PRFS. All systems, of course, are seen as a means for selectively distributing research funds. But most also seek to use it to drive particular behaviours, most commonly an improvement in the quality of research undertaken; or to increase accountability on the expenditure of taxpayers' money (Frölich, 2008). In all cases, the goal of creating a funding allocation mechanism is achieved. However, the degree to which the funds are effectively distributed, and the extent to which secondary goals are achieved, is not always readily apparent.

The nuances of desired outcomes are apparent in the stated aims of PRFS across countries. The University Grants Committee of Hong Kong, China, lists three main outcomes it hopes to achieve (UGC, 2006, p. 3):

- role differentiation in the UGC sector,
- international competitiveness,
- “deep collaboration”.

The UGC believes it has been effective in informing funding, symbolising public accountability and inducing improvement in research, but provides no evidence to support these assertions (UGC, 2006, p. 4). Indeed, the UGC acknowledges that the earliest research assessment exercises in 1993 and 1999

had “difficulty in distinguishing achievement at the top-end through a broad-brush exercise such as the RAE”, and had taken out of the funding pool a small proportion of the block grant to reward the highest quality research. How this was done was not specified (UGC, 2006, p. 5).

The current Australian system is purely a funding allocation system, with no explicit subsidiary aims along the lines espoused in other countries. This may well arise from its early introduction and the acknowledgment of its deficiencies as a research assessment tool, which has seen alternative systems proposed. The objectives of the most recent initiative, Excellence in Research for Australia (ERA) are to:

- Establish an evaluation framework that gives government, industry, business and the wider community assurance of the excellence of research conducted in Australia’s institutions.
- Provide a national stock-take of discipline-level areas of research strength and areas in which there is opportunity for development in Australia’s higher education institutions.
- Identify excellence across the full spectrum of research performance.
- Identify emerging research areas and opportunities for further development.
- Allow for comparisons of Australia’s research nationally and internationally for all discipline areas (ARC, 2009, p. 6).

As yet ERA has no direct funding implications, but institutions are responding on the assumption that this will change. In the interim the outcomes of the assessments will have significant implications for their ability to attract external funding and students, particularly full-fee-paying international students.

In some cases, existing measures were modified to stimulate the desired changes. Belgium’s Flemish government included bibliometric data in its funding formula with a view to improved research performance (Debackere and Glänzel, 2004, p. 265). The formula will be further refined in 2011 by incorporating data from the social sciences and humanities along the lines of the Norwegian model (Sivertsen, 2010, p. 22).

Many countries are trying to achieve additional outcomes. For example in Sweden, policies have aimed at increasing the share of staff with doctorates. Their new model for distributing general university funds (GUF) includes an element that counts for 5% of the total – the number of staff with doctorates (Elzinga, 2009, p. 6).

## Impact

The impact of a PRFS is multifaceted, and a number of actors influence final responses – institutions, managers and researchers themselves. In addition, many intervening factors can mitigate expected outcomes. Researchers may encounter contrary signals from the rewards system in their discipline, or other government agencies may run programmes that soften the impact of a relatively blunt PRFS.

The following teases out the evidence, or highlights the lack of it, in relation to the major consequences attributed to PRFS.

### *On funding*

To date, the New Zealand government has undertaken the most detailed analysis, with published data, on the funding implications for individual universities of its decision to introduce its Performance Based Research Funding Scheme (PBRF) (Ministry of Education, 2008). In contrast, much of the analysis of the impact of the RAE on funding came not from government, but from the print media, most notably *The Times Higher Education Supplement (THE)*. The funding implications of these two systems are discussed below, followed by a brief discussion of the compounding influence of international university rankings.

#### *New Zealand's PBRF*

The introduction of the PBRF had two main impacts on funding allocations: a shift in research funding away from institutes of technology and polytechnics (ITP) to universities, and a shift of funding to high-cost sciences at the expense of lower-cost fields (unspecified, but presumably social sciences and humanities). This followed a period during which increasing research funding flowed to the ITPs under the previous funding system. The ITPs' share of research funding was estimated to fall from 7.4% under the former scheme, to just 2.1% under the PBRF. Even though the funding pool was larger in the PBRF than the old system, ITPs collectively still received less than they would have under the earlier scheme. It was estimated they lost 65% of their research funding.

At the institutional level, Otago (26% increase in funding) and Lincoln (35% increase) were the big winners, while Auckland University of Technology, a relatively new university, received only half the funds that would have come its way under the old scheme.

In New Zealand's PBRF, fields are given differential weights to reflect the cost of undertaking different types of research. In the two evaluations conducted to date, the government has noted that this appears to have changed the fields that produce the most research income for institutions. Rather than low-cost fields with high student numbers, which previously drove the funding allocations, universities now derive more of their funding from high-cost fields (predominantly in the sciences) with relatively lower student numbers (New Zealand, Tertiary Education Commission, 2010, pp. 3-4).

### *The United Kingdom's RAE*

Successive RAEs increasingly concentrated research funds in a few institutions, particularly the research-intensive Russell Group, until the 2008 exercise. Changing the assessment process from a single "star" rating for a department to a quality profile resulted in some major funding shifts (Corbyn, 2009). David Sweeney of the HEFCE, which is responsible for the oversight of the RAE, calculated a 3% drop in the concentration of research. For example, there were 25 institutions that received no mainstream quality-related (QR) funding in 2008-09 but which received funding in 2009-10. Many other institutions saw dramatic increases in funding at the expense of the research-intensive universities. For example, the University of Lincoln will get GBP 1.6 million in 2010 compared to the GBP 220 000 it received in the previous year. The Russell Group's share dropped from 65% to 60%, the first decline in its share of RAE funds.

However, in the most recent RAE, which introduced a new assessment methodology, there was considerable criticism of the weightings originally given to each star quality level, and the implied rewards for an improvement in performance from one level to the next. Adams and Gurney (2010) have undertaken a detailed analysis of the implications of RAE 2008 for funding selectivity. Under previous RAEs, institutions gained significant benefit from achieving the top ranking, compared to the second highest. This was not the case in 2008. The weights originally given to each level were 0 for 1\*, 1 for 2\*, 3 for 3\*, and 7 for 4\*. In effect, this means that a 3\* was worth 3 times a 2\*, but a 4\* (a significantly harder level to achieve) was only worth 2.3 times a 3\*. The inequity of this was acknowledged and the weightings were changed to 0/1/3/9 for subsequent funding years. This will have the effect of reducing the movement in funds to just 1%, down from the 3% trumpeted by HEFCE immediately after the 2009/10 funding allocations were announced.



### *The importance of reputation*

Funding implications flow not just from the allocations received after PRFS assessments, but also as a result of the reputation achieved as a result of this process. Li *et al.* (2008, p. 6) point out that in a climate in which government funding for public universities is in decline in many countries, the continued viability and prosperity of a university can depend on its ability to attract external funding, which in turn can depend heavily on the institution's reputation. Success or failure validated by PRFS assessments can have significant flow-on effects outside the immediate funding allocations.

Reputation has gained increasing importance for Australian universities with the easing of restrictions on overseas students and the enrolment of full fee-paying students. In fact, international students are so important to Australia that education services are now Australia's third largest export, behind coal and iron ore (Baty, 2010). Many institutions appear to use the revenue they receive from these students to cross-subsidise the cost of research.

Reputation can be gained from sources other than PRFS. International rankings of universities, such as those published regularly by Jiao Tong University and the *THE*, have a strong influence on the regard in which institutions are held, and in turn have strong funding implications such as those mentioned in the preceding paragraphs. This situation exists even in the absence of any PRFS.

### *On human resource issues*

For many years, concerns have been raised about the impact of PRFS on several personnel issues. Most relate to the impact of systems that require universities to select and submit for assessment only the work of "research-active" staff. The implications of this requirement are manifold. It is believed to affect the morale of staff who are not included in their institution's submission; to be biased against women, early career researchers and ethnic groups; to have serious implications for the autonomy of an individual's line of research; and to have led to an active "transfer market" for leading academics.

McNay's study of the RAE also highlighted more general changes in job descriptions, recruitment criteria, creation of posts, appraisal systems, career patterns and rewards:

*“Previously, teaching in the PCFC sector and teaching and research in the UFC sector were conducted by permanent staff and a significant cadre of casual research workers. The greater emphasis on research has prompted some institutions to consider career paths for researchers, to provide bridging funds and to introduce more conscious management and appraisal of research staff. On the other hand, in some cases teaching posts have become casual to provide relief for permanent staff to do research. In general, a greater range of contract terms is being used.”* (HEFCE, 1997, para. 81)

Hong Kong, China, reports similar responses. The focus on publications in the RAE has now been incorporated into the management practices of many universities. Most have established performance standards for academics which dictate the expected number of publications for promotions and appointments (Li *et al.*, 2008, p. 7).

The United Kingdom’s RAE provides the most fertile ground for understanding the impact of PRFS on these issues, and studies undertaken on its impact provide some data to assist in distinguishing between perceptions (or possibilities) and reality. The two issues most often referred to are the designation of research-active staff and the transfer market in academic staff. These are discussed in more detail below.

### *Identifying research-active staff*

A number of PRFS, such as those in the United Kingdom, Hong Kong (China) and New Zealand, require the identification of research-active staff. Only those with the requisite output in the specified time period are submitted, or have their work submitted, for assessment.

McNay found a significant impact on morale after the 1996 RAE when institutional submissions became more selective and only research-active staff were submitted. In the 1996 exercise, a survey of academics found that nearly two-thirds of departments excluded researchers from their submissions, primarily because of the quality of their research (Evaluation Associates Ltd, 1999, p. 23). A study by the Association of University Teachers found the selection of research-active staff biased against women; men were 1.6 times more likely to be entered in the RAE than their female counterparts (Baty, 2004).

Even when handled sensitively by management (and this was not always the case), selectivity became divisive and had a significant impact on staff morale and on the collegiality of institutions more generally (HEFCE, 1997, para. 82-86). It was characterised as the most “traumatic” effect of the RAE. HEFCE postulated that the increased stress on staff could “be mitigated by

greater institutional encouragement and support”, though no studies have reported on whether such support mechanisms have been put in place (HEFCE, 1997, para. 100).

The need to identify research-active staff is also a concern in relation to New Zealand’s PBRF, for similar reasons.

### *The academic transfer market*

In the United Kingdom’s RAE, universities are assessed on the basis of the publications of researchers currently on staff. Recently recruited academics bring with them all the outputs they produced in the preceding six years (the usual census period for assessment). Even if academics have only been on staff for a matter of months, their full track record for the period may be submitted for assessment by their current institution. This has led to what is commonly termed the transfer market in academics leading up to each assessment.

This transfer market was widely regarded as the “most newsworthy aspect of the exercise” in the run-up to the 1996 exercise and its existence was blamed on the RAE (HEFCE, 1997, para. 88-97). It is difficult to quantify the effect. McNay found that only 2% of staff in his survey had moved institutions in the 18 months prior to the 1996 RAE. However senior administrators commented that the 2% were all top-level staff (McNay, 1998, p. 20). In addition, the impact is not solely seen in movement between institutions. The HEFCE report notes that while movement between institutions in the lead-up to the 1996 exercise may not have been remarkable, nearly half the staff surveyed had been promoted in the previous five years. Conceivably, the 1996 RAE had a greater effect on internal movements than on external ones (HEFCE, 1997, para. 90).

This issue was a topic of considerable interest in the press in the lead-up to each subsequent exercise. In defence of its system of selectivity, HEFCE noted that even if the RAE had increased staff movement, this should not necessarily be construed as a negative consequence – it could be a just reward for the most talented researchers (HEFCE, 1997, p. 14).

### ***On productivity***

Increased publication output appears to be a common impact of PRFS, irrespective of the model used, and this has generated a great deal of attention. Much of the discussion is anecdotal, but it is the one impact on which there is considerable bibliometric analysis, accompanied by a belief that it is possible to demonstrate the causal effect of the assessment systems. The United Kingdom, Australia, Spain and Norway have been the focus of detailed studies.

### *The United Kingdom's RAE*

A recent study, which examines annual data across the full period covered by all UK RAEs, looks more directly at the link between RAE cycles and changes in publication behaviour. It finds “pronounced ‘timing effects’ prior to a RAE” (Moed, 2008, p. 157). The data show:

- The United Kingdom’s share of publications in the Web of Science increased significantly in the four years before the 1992 exercise in which assessment was based in part on total publication numbers.
- With the shift to a focus on quality in the 1996 exercise, the United Kingdom’s share of publications in the Web of Science declined, but its share of higher-impact journals increased.
- The initial drop in productivity prior to the 1996 RAE appears to have lessened leading up to 2001 RAE as academics focused more on productivity and collaboration, perhaps in an effort to stimulate the number of staff judged research-active.

One note of caution applies to Moed’s study: it used aggregate data for the United Kingdom and the analysis was not limited to higher education institutions. While universities account for the majority of research publications, a significant minority come from the hospital and non-profit medical research sectors. Ideally, separating the sectors could provide better information for testing the existence and strength of the universities’ response.

Even though the focus of the 1996 RAE was quality, nearly two-thirds of respondents to the Evaluation Associates’ (EA) survey agreed that it had encouraged them to publish more, and that the focus was predominantly on peer-reviewed journals (Evaluation Associates Ltd, 1999, p. 17). These responses are at odds with Moed’s findings, which show that prior to the 1996 RAE productivity actually declined, with a stronger focus on quality than on quantity. It is an example of the mismatch between perception and reality, and raises questions about the extent to which anecdotal evidence, or survey responses, can provide an accurate picture of behavioural responses.

### *Australia*

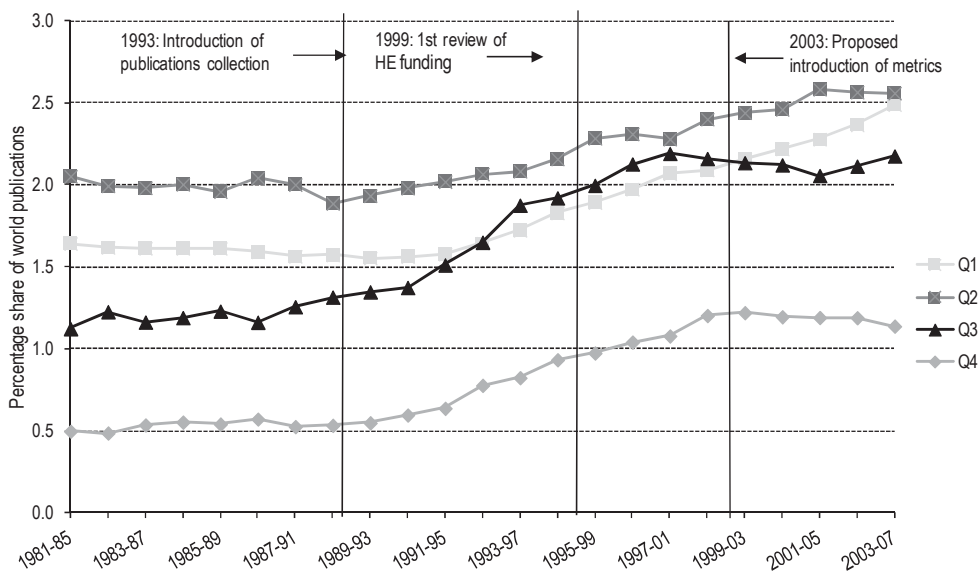
The Australian government has been using quantitative formulas to allocate the research component of the university block grant for nearly 20 years. A detailed analysis of the Web of Science data published in 2003 demonstrated a clear link between the introduction of this PRFS and an increase in productivity (Butler, 2003). By demonstrating that the identified trends applied to all fields of research, but were only present in the

university sector, and not in medical research institutes and other government research agencies, a direct causal relationship could be confidently implied.

Since that publication, the research policy landscape in Australia has shifted significantly. While the research block grants continue to be distributed on the same basis, a number of new initiatives have shifted the focus away from a purely quantity-driven calculation to quality. In 1999, a review of higher education funding first mooted the idea of dropping the element of the formulas based on simple publication counts. This did not happen immediately, but in 2003 a proposal was introduced to distribute funds on the basis of an RAE-style assessment system which would incorporate citation analysis where appropriate. A change of government saw this initiative fail, but a new initiative was developed which, while not yet linked to funding decisions, will assess the quality of research in Australian universities.

Figure 4.1 provides some indication of the response of universities to the various Australian policy initiatives.

**Figure 4.1. Australian university responses to government funding initiatives**



*Note:* Q1-4 indicate the quartile of journals, as measured by citation impact. Q1 is the highest quartile; Q4 is the lowest.

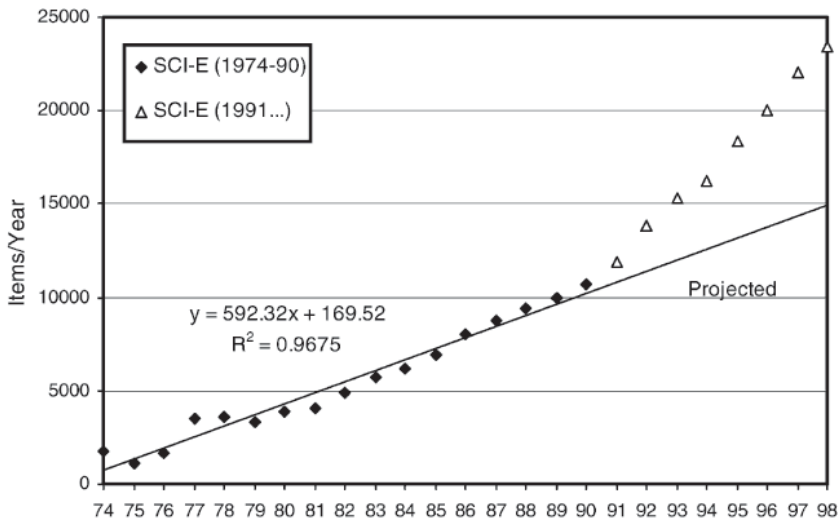
*Source:* Author.

The shaded area in Figure 4.1 reflects the analysis published in 2003, which clearly indicates the sudden increase in publication output from Australian universities after the introduction of the funding formulas, with a particularly sharp increase in the bottom two quartiles of journals (Q3 and Q4), as judged by citation impact. Another seven years of data have been added to that analysis and show that the trends in Q3 and Q4 flattened when the government's focus shifted to quality assessments, particularly after 2003.

### Spain

In 1989, the Spanish government introduced an incentive system for researchers based on publication output. Refinements to the system in 1994 saw increased emphasis on the use of international journals indexed in the Web of Science. A strong link was found between these policy initiatives and a marked increase in output (Jiménez-Contreras *et al.*, 2003). As with the Australian study, many possible alternative explanations for the increase in productivity (such as increased international mobility, funding from non-government sources, etc.) were examined. None of the factors reviewed was able to explain the change in growth rates after 1989. This gave the authors confidence in establishing a causal relationship. Figure 4.2 shows the change in publication output over predicted trends after the PRFS was introduced.

**Figure 4.2. Data-based predicted evolution of Spanish scientific production from 1974 to 1990**



Note: SCI-E is SCI Expanded (a version of the SCI that contains more indexed journals than the CD-ROM version).

Source: Reprinted from Jiménez-Contreras *et al.* (2003), p. 131, with permission from Elsevier.

Spanish output in the Web of Science moved well above the long-term trend line after the introduction of the research incentive system in 1989. The authors believe that while the monetary incentive initially caused a surge in output, the direct monetary rewards were too small to explain fully the continued increased productivity. They believe the increased focus on research, and its associated influence on prospects for promotion, led to the continuing upward trend (Jiménez-Contreras *et al.*, 2003, p. 136).

### *Norway*

In Norway there are concerns that the new PRFS might increase the number, but not necessarily the quality, of publications (Strehl, 2007, p. 51). However, since the introduction of the new system, there has been a substantial increase in publications, and output in both ranks of journals has increased at similar rates (Sivertsen, 2010, p. 26). There is some hesitancy in attributing this improvement to the new model, although Web of Science data show Norway's share of output increasing (2000-08), when that of other Nordic countries is decreasing (Sivertsen, 2009). The increase is in all journal quartiles. A number of qualitative changes have facilitated the increase in productivity – stronger institutional incentives for research; research seen as a communal as well as individual responsibility; increased internal awareness of publication activity; and improvement of research management practices (Sivertsen, 2010, p. 27). The Norwegian Ministry of Education saw the use of a two-level indicator as greatly reducing the risk of an explosion of publication activity at the expense of quality (European Commission, 2010, p. 121).

### *New Zealand*

While the first PBRF did not take place until 2004, the proposal for a performance-based research funding system was first made in 2001. In most universities this triggered a marked increase in publication activity. When a production function approach was used to model the research process in New Zealand universities, the results corroborated the findings on increased productivity (Smart, 2009).

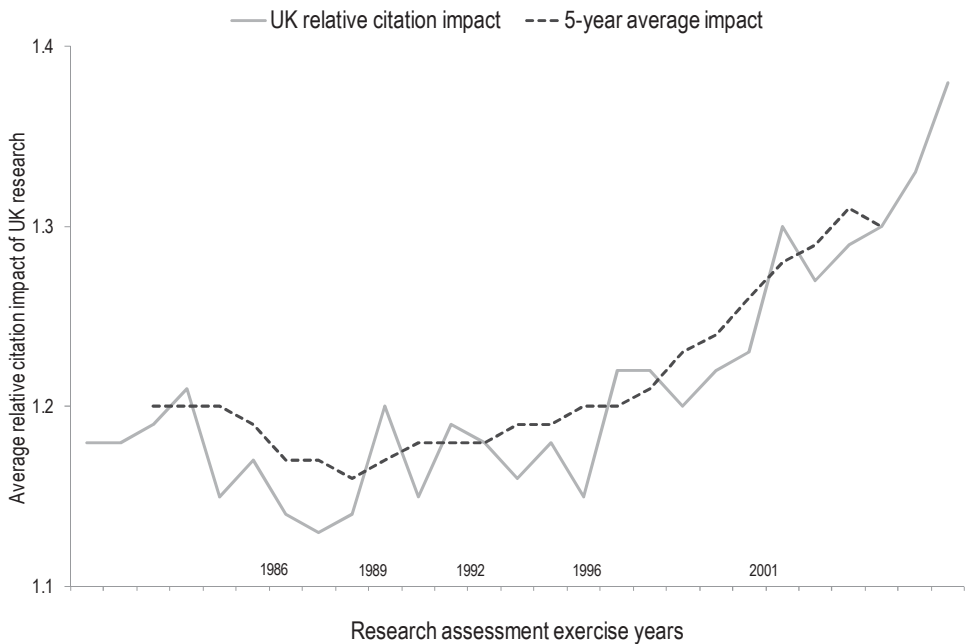
### *Flanders*

The Flemish government explicitly acknowledged the danger of a focus on quantity at the expense of quality if their BOF key continued to rely on publication counts (Debackere and Glänzel, 2004, p. 272). It established a working group to look at refinements to the key that would incorporate citation impact measures. It has also moved to incorporate an indicator similar to the one used in Norway to include output from the social sciences and humanities (Sivertsen, 2010, p. 22).

### *On quality*

A number of the studies described above also address the issue of the impact of PRFS on research quality. The Australian study clearly showed that, with the spotlight solely on productivity, there is a perceived drop-off in quality due to a greater tendency to publish in journals with less impact. With the change in focus to assessing quality, researchers appeared to maximise publication in journals with higher impact. There is no proof that, in itself, this denotes an improvement in the quality of research outputs, but a strong correlation is known to exist between publication in journals with high impact and quality of research, particularly at high levels of aggregation.

**Figure 4.3. The relative citation impact of the UK research base, 1981-2007**



*Note:* The average relative citation impact is citation counts normalised to world average by year of publication and journal category. The five-year average citation impact is the impact in a moving five-year window.

*Source:* Adams and Gurney (2010), p. 10. Reprinted from data and analysis by Thomson Reuters, published by HEPI (Higher Education Policy Institute) as “Funding Selectivity, Concentration and Excellence: How Good Is the UK’s Research?”



As mentioned, in Norway, there were strong initial concerns that the system would increase the number, but not necessarily the quality, of publications (Strehl, 2007, p. 51). However, publication numbers in outlets classified in both the upper and lower ranks have increased at a similar rate (Schneider, 2009, p. 374).

There is a widely held belief that the RAE has been a driver of improved research quality in the United Kingdom (UK Government, 2006, para. 4.6). Bibliometric analyses confirm the improvement in UK science, in terms of relative citation impact. Most bibliometricians caution against asserting that a causal relationship has been proved, but the evidence that this is what the data reveal is mounting (Adams and Gurney, 2010, p. 3; Moed, 2008, p. 159), as Figure 4.3 demonstrates.

Figure 4.3 indicates the timing of the UK RAE cycles. The data show a downward trajectory in citation impact until the introduction of quality-based assessments in the 1992 RAE. The improvement in UK research impact has continued with each subsequent round of the RAE.

### ***On teaching***

A concern expressed in relation to all PRFS was the impact on the quality and standing of the teaching component of academics' work. In the United Kingdom, universities were required to report the total number of staff appointed to teaching-and-research and research-only positions, enabling calculation of the proportion of staff submitted for assessment. One way of improving this proportion was to move less research-active staff to teaching-only appointments.

One of the challenges faced by policy analysts is that while methods for assessing research quality are tried and tested, attempts to assess and reward teaching performance have proved much more problematic. As a result, discussions of the impact on teaching are largely anecdotal, with the exception of a research report from the UK Association of University Teachers that adds some evidential weight to the concerns. A few representative examples of these concerns are given below.

In terms of the effect of the RAE on teaching in the United Kingdom, research undertaken by the Association of University Teachers appears to substantiate claims that, in the lead-up to the 2008 exercise, less research-active staff were being moved to teaching-only contracts (Sanders, 2005). The belief was that this was a reaction to the need to select research-active staff for submission to the RAE and that it implied that teaching-only staff had lower status in the university system than staff with research activities.

When Ireland's new Recurrent Grant Allocation Model (RGAM) was phased in from 2006, there were concerns that it would divert resources from teaching to research, yet only 5% of the core budget is earmarked for research (Strehl, 2007, pp. 39, 50).

In Hong Kong, China, concerns that the RAE places "undue emphasis on research" given the "significant marginal returns" have been raised by the body responsible for distributing the funding (UGC, 2006, p. 4).

In New Zealand, there have always been concerns that any PBRF system will drive a wedge between the teaching and research functions of a university, yet studies that seek to determine whether or not this has occurred do not appear to have been carried out.

### *On discipline mix*

The main concerns in relation to discipline mix centre on the real or perceived advantages or disadvantages that are thought to flow to particular discipline groups. Any quantitative assessment is generally seen to favour science, technology and medicine (STEM), at the expense of the humanities, social sciences and arts (HASS). This is primarily because accepted proxies for assessing research quality exist for STEM subjects in the form of citation analyses, but equivalent measures are not so readily available for the HASS disciplines. A comprehensive review of the issues faced by HASS disciplines is contained in a report published by the Australian Council for the Humanities, Arts and Social Sciences (CHASS, 2006).

In Norway too, where the PBRF does not incorporate any qualitative assessment, there was concern that areas of research in which the production of articles/monographs is not the norm will be discriminated against (Strehl, 2007, p. 51).

In the United Kingdom, concerns have focused more on disciplines with an emerging research culture than on a simple difference between STEM and HASS disciplines. In its response to a consultation on the RAE overseen by the joint UK funding bodies, the Department of Health raises its concerns about disciplines such as nursing and allied health professions:

*"We have had mounting concerns that less well established disciplines such as nursing research or Allied Health Professions' (AHPs) work have been disadvantaged in a process where the quality standards have risen faster than such new subjects could be expected to keep up. This is discouraging at the personal level and financially damaging at the institutional one. This becomes even more pronounced since such new topics are frequently based in newer universities with only limited research resources."* (UK Department of Health, 2001, p. 4)

While the debate on the implications for the discipline mix within universities has not generally pitted STEM against HASS, there remain some concerns about the comparability of results across disciplines, particularly between the sciences and social sciences (Corbyn, 2008). These concerns exist even within these two subject groups and do not appear to have been alleviated by changes to the panel structure for the 2008 RAE that was designed with this issue in mind.

### *On the focus of research*

A general concern of all PRFSs, whether based on qualitative or quantitative assessment, is that they favour “mainstream”, disciplinary-based, basic, “safe” research at the expense of applied, interdisciplinary or speculative research. These concerns are regularly voiced in the academic press, yet little research has been undertaken to prove or disprove these biases. As Gläser notes succinctly:

*“Our knowledge about the development and effects of RES [research evaluation system] has developed rather unevenly. While the science policy processes and organisational dynamics have been investigated for several countries, we have just begun to explore how RES interact with scientific communities and their research.”* (Gläser, 2008, p. 246)

In Spain, concerns about the Evaluation of Research Activity policy span a number of related themes, as it is believed to lead researchers to focus on research that is attractive to Web of Science journal editors, and there is a fear that more applied research, and research with a local/regional focus, may suffer (European Commission, 2010, p. 123). However, studies to determine whether these fears have been realised do not appear to have been undertaken.

Evidence gleaned from the literature on a number of themes relating to the focus on research is presented below.

### *Interdisciplinarity*

A major concern for PRFS, based either on field-based quantitative indicators or peer assessments by panels convened along disciplinary lines, is the impact on interdisciplinary work. It is extremely difficult to find clear evidence on whether such systems are indeed biased against such research.

Quite often perceptions, either from survey responses or discussed on an anecdotal basis in the press, can be quite different from the reality. An example is the regularly discussed concern on the bias of discipline-based systems against interdisciplinary research. The view that the RAE was biased against interdisciplinary research is widespread, and accepted within

government circles, and was one of the reasons quoted for moving to a REF (UK Government, 2006, para. 4.7). It was also a major motivation for the Roberts Review of the RAE after the 2001 exercise (Roberts, 2003, p. 22).

Yet two studies on this issue, undertaken before these policy changes were mooted, found no evidence to support the assertion. McNay's study found that while the number of staff who reported that they had moved away from interdisciplinary work was relatively small, almost half those surveyed felt that the RAE hindered interdisciplinary work (McNay, 1998, p. 20). The Evaluation Associates study found that while nearly a quarter of researchers believed that the RAE was a strong inhibitor of interdisciplinary research, the ratings achieved by departments with a high proportion of interdisciplinary researchers show that there was in fact no such discrimination (Evaluation Associates Ltd, 1999, pp. 13, 28).

McNay's 1997 survey demonstrates that reactions can vary by type of institution and that general statements on an issue may not provide an accurate picture. While a significant minority of researchers, and nearly half of research managers, felt it was not a good strategy to pursue interdisciplinary work, many staff in the post-1992 universities focused on applied/interdisciplinary work, which they saw as their strength (McNay, 1998, p. 20).

#### *“Blue skies” research*

McNay's survey found that nearly half the research managers surveyed felt the RAE hindered the pursuit of new research areas or risky “blue skies” research (McNay, 1998, p. 20). In Norway, the principal fear regarding the new PRFS appears to be that “mainstream” research will be advantaged (Strehl, 2007, p. 51). The Hong Kong, China, RAE was also perceived to have a “narrow focus on traditional research” (UGC, 2006, p. 4).

The RAE time scales compounded these concerns for staff and a significant minority (around one-quarter) reported avoiding new lines of research and speculative topics because they believed quality outputs could not be achieved by RAE submission deadlines (McNay, 1998, p. 20). This was corroborated by Evaluation Associates' study of the 1996 RAE, which found that researchers felt pressure to get results published early to meet cut-off points for the exercise (63%), and a small minority felt that there was little incentive to focus on longer-term research (Evaluation Associates Ltd, 1999, p. 17).

### *Research on the periphery*

A third theme in this area relates to smaller nations and/or those seen to be on the periphery of the large centres of research in the United States and the United Kingdom. University rankings are believed to lead to a neglect of local or regional issues in order to publish in high-impact international journals (Hazelkorn, 2009, p. 2). This is particularly true of PRFS that specifically aim to increase publication in the international journals, as happens in Spain (Jiménez-Contreras *et al.*, 2003). There are legitimate concerns that national journals and research priorities will be neglected in the rush to publish in Web of Science journals. PRFS funding systems that incorporate ranked outlets and/or citation measures have the potential to encourage such a response.

Fifteen years ago Arunachalam raised concerns about the mismatch between the topics on which Indian medical researchers published and the areas of research related to the country's most pressing medical problems (Arunachalam, 1995). While this did not result from PRFS but from academic incentives that rewarded researchers with career advancement for a strong international publication record, it demonstrates that concerns about a reliance on databases with country/language biases are real.

### *Applied research*

A fourth theme focuses on concerns that research targeted at policy, social interventions, professional practice, etc., will be disadvantaged. This is of particular concern in PRFS systems based primarily on quantitative indicators, as much of this research is published in the “grey literature” that falls outside the ambit of standard performance indicators.

### *Researcher autonomy*

One issue that has gained prominence recently relates to concerns over researcher autonomy. With the increased incidence of departmental micro-level research management strategies, there is concern that researchers are being pressured to stay within their own field of expertise and not move down new paths. A quote from one of the respondents to the survey describes these pressures:

*“Unfortunately the RAE categories and decisions are, at least in my institution, permeating decisions about research activity. They’re increasingly looking for RAE publications; in a sense, monitoring people’s research and parts of their careers.”* (RIN, 2009, p. 36)

Nearly one-third of staff surveyed by McNay felt that the direction of their research was increasingly conditioned by the collective priorities of their group or department, rather than their individual preferences (McNay, 1998, p. 20).

In Norway there are also concerns about a loss of autonomy. “Research is now perceived as a common and institutional responsibility, not only as an individual task.” (Sivertsen, 2009, slide 38)

### ***On collaboration***

Concerns about the effect of PRFS on collaboration appear to relate to specific details of the assessment methodology. The choice is generally between whole counts, *i.e.* giving each institution (or author) full credit for an article on which their name appears (which leads to double counting) and fractional counts, *i.e.* an article only counts as a single unit and each institution (or author) receives a fraction of the credit.

Concerns about the effect of fractional counts on collaborative activities led Australian governments to reject this methodology in favour of whole counts. Norwegians appear less concerned and believe their use of fractional counts has not resulted in a decline in collaborative activities (Schneider, 2009, p. 372). Schneider believes that “‘invisible colleges’ and social networks within research specialties have eventually ensured collaboration”. It is believed that the dependence of research on collaboration will counteract any adverse behaviour that might result from the funding model.

In contrast, in response to a survey on the 1996 RAE, researchers felt that the RAE did not encourage them to work more collaboratively, particularly with researchers or users outside the university sector (Evaluation Associates Ltd, 1999, p. 50).

Countries such as Australia and the United Kingdom continue to have dual funding systems and specific schemes aimed at stimulating collaborative activities. Many EU funding programmes require extensive collaborative networks. These work to soften the impact of any system that might otherwise seem to work against this trend.

### ***On institutional management practices***

The first response to the introduction of a PRFS often comes from senior management – it takes time for individual academics to feel its effects. The implications of any PRFS do not appear on an academic’s radar until specific institutional policies put it there, *e.g.* the need to provide information; individual funding incentives for individual researchers to improve performance; changes to requirements for promotion; etc.

### *Information management systems*

A number of countries report significant improvement in the information management systems of universities. With the increased need for data, at ever increasing levels of detail, many universities have redesigned their information management systems. One of the main drivers is the need to bring together data from a range of different internal sources: data on students, staff, competitive grants, external income, research publications, etc.

The rise of PRFS and the increased visibility of university ranking tables have led universities, among other things, to a “renewed emphasis on the accuracy/amount of data gathered and shared with third parties” (Hazelkorn, 2007, p. 13).

Research management also is claimed to have improved in Norway with the aid of complete bibliometric information about research activities (Sivertsen, 2009, slide 38). The same is also regularly claimed for Australia and the United Kingdom.

### *Institutional repositories*

In both Australia and the United Kingdom, moves to introduce a metrics-based system of assessment has led universities to expend considerable effort on improving or establishing institutional repositories for housing publications that are likely to be included in the assessment. This is seen to be a particularly beneficial, unintended consequence of the introduction of a PRFS. Australian librarians noted:

*“The imminent introduction of the RQF [Research Quality Framework] has served to justify, and hence to hasten, the introduction of a repository in some universities, while encouraging better communication between the research office and repository managers.”* (Henty, 2007)

This is a trend that is also gathering pace in the United Kingdom.

### *Research management*

When first introduced, PRFS can have a profound effect on the focus and styles of research management within institutions.

A survey of universities after the 1992 UK RAE found that it had become a driver of institutional research planning, particularly when it became clear that the assessment exercise would be conducted regularly (HEFCE, 1997). However, it was not always clear that these changes could be linked directly to the RAE. A significant minority of respondents (close to one-third) also pointed to the policies of other government agencies, research councils and funders as important drivers of improved, strategic

research management. McNay found a greater focus on research in the post-1992 universities and colleges (McNay, 1998).

The view that these changes could be linked, at least in part, to the rise of assessments/rankings was supported by a more recent survey. Respondents reported that new organisational sections were established, or individuals assigned, “to deal with indicator improvements and monitor rankings” (Hazelkorn, 2007, p. 13).

Another recent study made a related point: many changes may occur anyway, but assessments/rankings made these happen quicker (CHERI and Hobsons Research, 2008, p. 34). The rise of university rankings is also believed to lead to strong reactions from senior management.

### *Departmental restructuring*

One consequence of research assessment that has been directly linked to the RAE has been the internal restructuring of universities after the results of the latest exercise are released.

In some cases, universities respond to poor RAE results by closing departments. After the most recent RAE in 2008, the University of Liverpool announced the likely closure of three departments – statistics, philosophy and politics, and communications studies. All were deemed by the RAE assessors to have no “world-leading” (4\*) activity (Newman, 2009a). The university is also contemplating closing other departments that were in the lowest quartile (based on 3\* + 4\* outputs) – civil engineering, cancer studies, dentistry, American studies and sociology. This, in spite of the fact that the University of Liverpool was an overall “winner” in RAE 2008 (its QR funds increased by 21.3% – GBP 4.4 million). It was also contemplating providing no support for staff assessed as 1\*.

It is feared the poor showing of politics in the 2008 RAE (rated 62 overall out of 67 disciplines, based on the average research grade) will lead to a number of department closures, even though the RAE results were at odds with an Economic and Social Research Council review which assessed the discipline’s overall standing as high (Newman, 2009b). This demonstrates the RAE’s power to drive management strategies in the UK higher education system.

In France, some institutions use the AERES exercise as a “benchmarking exercise” and try to “improve their rating by closing down sectors which show a poor performance” (European Commission, 2010, p. 97).



### *Strategic recruitment*

Some of the implications of the selective submission of research for assessment, particularly for individual researchers, are discussed above. This section looks more specifically at the strategies that institutions have employed in deciding the proportion of staff to submit for assessment.

The issue is perhaps most clearly illustrated in relation to the most recent RAE. A number of 2008 RAE panels (*e.g.* economics and econometrics; accounting and finance; business and management studies) believed that some universities excluded research-active staff in order to gain a higher ranking, while others drafted in “research stars on ‘unusual’ contracts who were not fully integrated members of the research team” (Newman, 2009c). This occurred despite the chair of the social sciences panel having warned, well over a year earlier, that he would be looking at the possibility of “game playing” and urging universities to enter all their research-active staff (Lipsett and Tysome, 2006). However, it was difficult to determine the extent of this practice empirically because there was no information on the proportion of staff submitted owing to ambiguous definitions of eligibility.

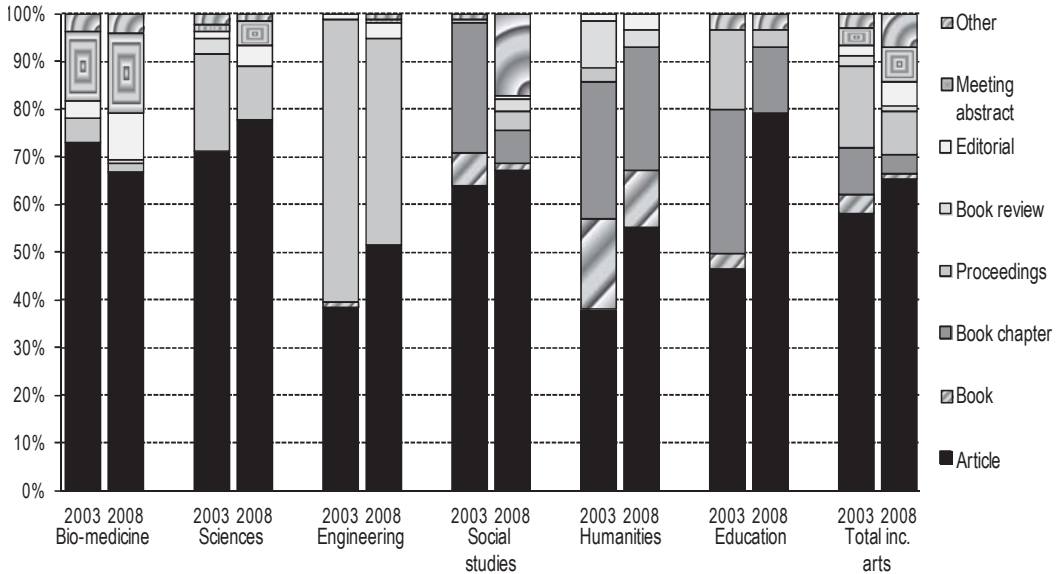
HEFCE had recognised the incentive to omit staff from submissions in order to improve their ranking, and the new system of distributional scoring was introduced in part to overcome such practices. It does not appear to have been successful (MacLeod, 2006). Australia hoped to lessen some of the game playing by including *all* research. However it followed the United Kingdom’s lead in using a census date approach to determining the institutional affiliation of researchers, so it is expected that some form of transfer market will arise (or already has arisen).

### *On where to publish*

One of the major concerns about PRFS relates to the possibility of publication practices being altered to suit what is seen as the best strategy for the assessment methodology. In particular, there is concern about the rise in the use of journals in disciplines that have traditionally favoured book or conference outlets.

A recent online survey of over 800 UK academics provides some very clear evidence on changes that are occurring in academic publishing patterns and on the reasons why. Figure 4.4 provides clear insights into changes that have occurred in the last decade.

Figure 4.4. Change in publication outlets by major field of research, 2003 to 2008



Source: (RIN, 2009, p. 16).

Concerns about these changes are encapsulated in the views of one of the respondents in that study:

*“Sadly, I find myself increasingly moving away from publishing in journals which are important and read by a lot of colleagues, to publishing in high status journals instead. This had led to much longer delays [and] thus adversely affects science, but I feel the pressure to do this in order to advance career-wise.” [Medical and biological sciences] (RIN, 2009).*

McNay’s earlier 1997 study also found that most heads of research units (84%), and a majority of staff (55%), felt that the RAE had led to the targeting of key journals (McNay, 1998, p. 20), and many of those surveyed believed the RAE had led to premature publication (30% of heads of units believe it, 25% of staff admit to it).

As mentioned earlier, in addition to a changing use of publication outlets, the study found that work is being fast-tracked for publication and may be less thoroughly researched than in the past. Some analysts blame the “publish or perish” culture and the pressures this entails as the reason for the increased rate of retractions in scientific journals (Corbyn, 2010). Perhaps these two trends are linked.

### *On author behaviour*

Multiple authorship has become commonplace in research publications, reflecting an increase in collaboration both within and between institutions and countries. However disciplines, and even different groups within disciplines, vary considerably in their standard traditions for author attribution. From an online survey of UK academics, a report conducted for the Research Information Network (RIN) determined the different practices among disciplines from over 800 responses (RIN, 2009). It is worth reproducing the summary of their findings here:

*“The listing of authors in order of contribution (with first author providing the greatest contribution) is the most frequent practice in most disciplines except for the humanities where alphabetical order is the norm. But it is notable that in physical sciences, mathematics and social sciences alphabetical ordering and ordering by contribution are almost equally common. Notable also are the differences of practice within discipline groups. In medical and biological sciences, in physical sciences and maths, and in engineering and computing, ordering by contribution may frequently be modified by placing the senior researcher or grant-holder last. And while the most common practice with papers arising from research undertaken by students is to place the student first in the author list, a significant minority in medical and biological sciences and in social sciences place the supervisor first.”* (RIN 2009, p. 26)

The implications of this for any performance-based funding system that is predicated, even in part, on quantitative measures related to publications is two-fold:

1. Any attempt to fractionate publications across collaborating authors/institutions is fraught with difficulties, particularly if an attempt is made to weight shares according to level of contribution. Practices vary too widely to allow this to be done automatically. Until such time as authors routinely ascribe percentage contributions, it is unlikely that fractionation can be used without seriously disadvantaging some authors and/or disciplines.
2. Undertaking any analysis on the basis of first author only (as in the current Swedish model) is likely to have serious consequences, particularly at lower levels of aggregation. At the national and institutional level, the assessment of performance is likely to be accurate for most countries, but as soon as the methodology is used at lower levels of aggregations, serious distortions can arise.

## Additional comments

### *Overt game playing*

The strategies adopted by universities in response to PRFS are unsurprising. All seek to maximise their returns from a system with direct implications for the funding they receive. It can be difficult to draw the line between acceptable and unacceptable strategies. Much criticism has surrounded the “transfer market” in “star” researchers that occurs before each RAE round. But this can also be seen as a just reward for high achievers, something that was missing prior to the introduction of the RAE.

Yet undoubtedly some strategies are clearly against the interests of the sector. One unambiguous example relates to the new area of assessment based on webmetrics. Some universities have started to use link farms<sup>1</sup> and paid backlinks<sup>2</sup> to improve their positions (European Commission, 2010, p. 132). This clearly undermines the rankings, and the rankings developers have removed these institutions from their tables (Prieto Valverde and Fernández Pérez, n.d.). They have also discovered some universities that are hosting large numbers of academic papers authored by scientists that do not belong to those institutions.

A few instances of game playing have been identified in relation to the Australian system, which is based on publication counts. Quite detailed and restrictive definitions are applied to the four types of publications that can be included in a university’s return – conferences must be “international” and journals must have an editorial board that goes well beyond a single institution. One university overcame the first issue by establishing a conference with “international” in the title and ensuring at least one participant from overseas, but 90% of the presentations were from the institution. An entrepreneurial academic overcame the restrictions on editorial board by establishing one consisting solely of his former PhD students, a number of whom had moved back to their former institutions. Both strategies were well outside the spirit of the process, but fortunately such extreme examples appear to be rare.

### *Perception and reality*

The Research Information Network report clearly shows that researchers submit different kinds of outputs to those they publish – the sciences and engineering submit a much lower proportion of conference papers than in their overall publishing patterns; and the humanities submit a larger proportion of books (RIN, 2009). The survey of academics reveals that a quarter of those who responded believed that the RAE excludes some types of research output which they consider important. This was despite all

panels clearly stating that virtually all types of output are admissible and the guidelines clearly specifying that all should be treated equally. Clearly some academics receive misleading advice, either from institutional research managers or their colleagues.

In relation to peer review, a British Academy report put forward the argument that the criticisms directed at peer review related to “deficiencies of practice rather than the principle of peer review” (British Academy, 2007). The system itself was considered the best, but sometimes less than optimal in practice.

### *Inconsistent outcomes*

Institutions do not always respond in identical ways. A poor assessment may lead one university to close down a department, but another may decide to invest more resources in that department in an attempt to improve outcomes in the next assessment. HEFCE noted that the RAE was said to have stimulated some institutions to recruit new or young staff with research potential, but also to have caused other institutions to adopt a more conservative approach of recruiting proven researchers (HEFCE, 1997, para. 101).

Responses may change over time. In the United Kingdom, institutions appear to have become more selective of the staff/publications/units they submit and more aggressive in their recruitment of “stars” in each successive RAE cycle.

Also, the problem of grade inflation in the 2001 RAE resulted in departments that had maintained a 5\* rating losing funding while, at the same time, some departments that improved their ranking did not receive funding because departments on the lower ratings (1, 2, and 3b) were no longer funded (Tapper and Salter, 2003, p. 19).

### *Other influences*

PRFS do not operate in isolation. Countries such as Australia and the United Kingdom operate dual funding systems, with significant amounts of money distributed through a number of research councils. The methods other funders use to determine the success of grant applications can soften the impact of PRFS, particularly if different assessment processes are used. In Australia, the Research Council uses a traditional peer review approach, in stark contrast to the blunt funding formula used to distribute the research block grant. In the United Kingdom, the systems are less dissimilar, and this is likely to reinforce the impacts of the PRFS.

Unrelated government policies can soften potentially negative unintended consequences. A number of countries have programmes aimed at facilitating cross-institutional collaboration, which work to overcome a potential bias against collaborative work in an assessment system that uses fractional counts (e.g. Sweden). In Australia, the existence of a significant pool of money to fund grants that link researchers with industry helps to counter a potential bias against more applied research.

The Evaluation Associates' study also listed a small sample of other factors that can influence research performance (and in particular interdisciplinary research). Teaching loads; the availability of equipment, software and data; contract funding; institutional "environment"; and the policies of research councils were some of the examples provided (Evaluation Associates Ltd, 1999, p. 17).

### *Responses to specific assessment measures*

Many responses are generic, whatever the methodology employed. But some are quite specific to the particular methodology used. This was touched on in relation to the choice between fractional and whole counts in quantitative-based assessments. Some other measure-specific concerns are briefly covered here.

The recent RIN report reinforces the fears of many academics (22%, or nearly a quarter, of survey respondents) who believed the use of bibliometrics in research assessment would lead to game playing in the form of citation clubs and increased self-citation and citation of collaborators, although few indicated they would cite their rivals' work less often (RIN, 2009, p. 7). What all commentators on the effects of bibliometrics seem to forget is the role that journal editors and reviewers play as gatekeepers of appropriate behaviours.

Assessing only a specified proportion of total output can result in misleading assessments of total research. One panel chair feared that the selectivity used in some institutions could lead to very poor results in disciplines for which the perceived wisdom was that UK research was very good (Corbyn, 2008).

Responses from institutions may depend on details of the process. There are concerns in the United Kingdom that moves to force universities to submit all research-active staff may lead to even more destructive game playing than the existing selective process which encourages the poaching of "stars" (Roberts, 2003, p. 52), but in Australia the inclusion of all staff (and all outputs) does not appear to have led to a major shift in the classification of staff.

### ***Broader impact***

In some instances, the introduction of PRFS can have an impact beyond the higher education or research sectors on which the assessment focuses. Research is not an isolated endeavour. It involves a myriad of inter-linked stakeholders, all with an interest in the processes used in those assessments, and the implications, if any, for them. Some examples of the broader impact of assessment systems are described below.

In Spain there is a perception that several Spanish journals have improved their manuscript assessment and peer review processes in an attempt to be indexed in the Web of Science (European Commission, 2010, p. 123). The scientific publishing industry has a keen interest in the mechanics of these systems as they can have a direct bearing on the attractiveness of their journals as outlets for the work of researchers. Journal editors and publishers took a very keen and active interest in the Australian exercise to classify over 20 000 peer-reviewed journals into four quality bands. The work done by committees attempting a similar exercise for the European Reference Index for the Humanities under the aegis of the European Science Foundation also gained their keen attention.

As with improvements in the standards of Spanish journals, PRFS can have other external benefits. Public pressure can lead to improvements in the quality of data, as in the development of the Berlin Protocol for university rankings. It also occurred with the German CHE rankings, which initially only included professors' publications. There was a strong belief that this distorted the outcomes, and the collection of data was subsequently expanded to include the publication of all staff holding a doctorate (European Commission, 2010, p. 104).

Even the rise of university rankings can be seen as an indirect result of the introduction of PRFS in a limited range of countries – they fill the gap for countries in which no such system exists (European Commission, 2010, p. 111). The developers of the Jiao Tong rankings in Shanghai have often said that the prime motivation of the construction of their index was to provide a method, however crude in its initial form, to answer the question: “Does China have any world class universities?”

## Conclusion

Hicks (see Chapter 1) lists 12 countries with performance-based funding systems, yet there is a relative scarcity of evidence-based analysis of the impact of these systems in the literature. It is possible that some may be buried in the grey literature of government reports. While Norway is known to be very active in this area, and researchers have presented the results of a number of studies at recent conferences, it is difficult to find detailed reports. It is likely that a number of studies will only be written in the language of the relevant country, as the primary audience will be local government agencies. This may explain a lack of studies from Poland and the Slovak Republic.

It is likely that considerable information could be unearthed, and it is hoped that, as is occurring in Italy, scholars will start to scrutinise these systems and publish more evidence-based assessments of their impact. They are certainly needed. Without hard evidence, anecdotes will hold sway.

This literature review shows the mismatch that can occur between perception and reality. It raises questions about the extent to which anecdotal evidence, or survey responses, can provide an accurate picture of behavioural responses. For every anecdote about a particular response to a PRFS, it will nearly always be possible to discover an anecdote that suggests the contrary. Similar contradictions in responses to survey questions have also been highlighted.

Yet the distribution of research funds on the basis of the assessment of performance is here to stay. The architects of PRFS may well anticipate, and even encourage, some of the more obvious impacts of the schemes they introduce. But academics are very creative – they can and will respond to such measures in novel and unforeseen ways. And the institutions they serve will always seek to maximise returns, as is their responsibility. Many impacts occur, no matter what method is employed. Redesigning the system will therefore not change anything – what is needed are alternative policies to overcome unwanted responses from academia.

This leads back to one of the central points made in Whitley and Gläser (2008): What needs to be studied is clear. The challenge is to facilitate the prioritisation of studies on these issues. Some potential studies are easy to identify and scope. They include:



- A broad, multinational assessment of publication trends that compares sectors/countries subject to PRFS with those that are not, thus providing a clearer picture of causality.
- Bibliometric analyses of trends in co-citation and co-authorship patterns to gain insight into the effect of PRFS on interdisciplinarity and collaboration.
- Disaggregating data analyses down to disciplines to investigate differential responses to PRFS.
- Analysing trends in publication strategies (perhaps expanding the RIN study to other countries).
- A detailed study of staff data to determine changes in the classification of staff and movement between institutions to provide some insights into questions such as gender bias, movement to casual and/or teaching only appointments, and the extent of the “transfer market”.

For these studies the data are readily available, but they need to be co-ordinated and have a wider remit than a single country. It is more difficult to delve deeply into issues that require the input of sociologists of science and their qualitative assessment techniques: how the content of research has changed; whether applied research is suffering; what signals researchers are receiving and how they respond; etc.

What is clearly needed is a co-ordinated approach to the study of the impact of PRFS on all aspects of the academic endeavour, incorporating both quantitative and qualitative approaches. Without valid evidence on their impact, it is impossible to assess the efficacy of their use and develop strategies to overcome adverse consequences.

## *Acronyms*

AERES	<i>Agence d’Evaluation de la Recherche et de l’Enseignement Supérieur</i> (France)
AHP	Allied Health Professions (United Kingdom)
ARC	Australian Research Council
BOF	<i>Bijzonder Onderzoeksfonds</i> (Flanders)
CHASS	Council for the Humanities, Arts and Social Sciences (Australia)
CHE	Centre for Higher Education Development (Germany)
CHERI	Centre for Higher Education Research and Information (United Kingdom)
DfES	Department for Education and Skills (United Kingdom)
ERA	Excellence in Research for Australia
HASS	Humanities, Social Sciences and Arts (United Kingdom and Australia)
HEFCE	Higher Education Funding Council of England
ITP	Institutes of Technology and Polytechnics (New Zealand)
PBRF	Performance-Based Research Fund (New Zealand)
PCFC	Polytechnics and Colleges Funding Council (United Kingdom)
PRFS	Performance-based Research Funding System
RAE	Research Assessment Exercise (United Kingdom)
REPP	Research Evaluation and Policy Project (Australia)
RES	Research Evaluation Systems
RIN	Research Information Network (United Kingdom)
RQF	Research Quality Framework (Australia)
STEM	Science, Technology and Medicine (United Kingdom)
THE	Times Higher Education
UGC	University Grants Committee (Hong Kong, China)

## *Notes*

1. According to Wikipedia a link farm is: “any group of web sites that all hyperlink to every other site in the group”.
2. The Wikipedia definition of backlinks is: “incoming links to a website or web page”.

## *References*

- Adams, J. and K. Gurney (2010), “Funding Selectivity, Concentration and Excellence – How Good Is the UK’s Research?”, Higher Education Policy Institute (HEPI), [www.hepi.ac.uk/455-1793/Funding-selectivity,-concentration-and-excellence---how-good-is-the-UK%27s-research.html](http://www.hepi.ac.uk/455-1793/Funding-selectivity,-concentration-and-excellence---how-good-is-the-UK%27s-research.html).
- ARC (Australian Research Council) (2009), *ERA 2010 Submission Guidelines*, [www.arc.gov.au/era/key\\_docs10.htm](http://www.arc.gov.au/era/key_docs10.htm) (accessed 23 March 2010).
- Arunachalam, S. (1995), “Research As If Relevance Mattered: Medical Research in India As Reflected by SCI 1981-85”, in *Proceedings of the Fifth Biennial Conference of the International Society for Scientometrics and Informetrics*, ISSI, River Forest, IL, pp. 33-42.
- Baty, P. (2004), “Sex Bias Limits Women in RAE”, *THES*, 16 July, [www.timeshighereducation.co.uk/story.asp?storyCode=190047&sectioncode=26](http://www.timeshighereducation.co.uk/story.asp?storyCode=190047&sectioncode=26) (accessed 6 March 2010).
- Baty, P. (2010), “Exposed and Vulnerable”, *THES*, 22 April, [www.timeshighereducation.co.uk/story.asp?sectioncode=26&storycode=411285&c=2](http://www.timeshighereducation.co.uk/story.asp?sectioncode=26&storycode=411285&c=2).
- British Academy (2007), *Peer Review: The Challenges for the Humanities and Social Sciences*, [www.britac.ac.uk/policy/peer-review.cfm](http://www.britac.ac.uk/policy/peer-review.cfm).
- Butler, L. (2003), “Explaining Australia’s Increased Share of ISI Publications - The Effects of a Funding Formula based on publication Counts”, *Research Policy*, 32(1), pp. 143-155.
- CHASS (2006), “Measures of Quality and Impact of Publicly Funded Research in the Humanities, Arts and Social Sciences”, [www.chass.org.au/papers/PAP20051101JP.php](http://www.chass.org.au/papers/PAP20051101JP.php).

- CHERI and Hobsons Research (2008), “Counting What is Measured or Measuring What Counts? League Tables and Their Impact on Higher Education Institutions in England”, issues paper, report to HEFCE, April 2008/14.
- Corbyn, Z. (2008), “RAE Results May Not Reflect True Quality of UK Research, Warns Chair”, *THES*, 4 December.
- Corbyn, Z. (2009), “Reversal of Fortunes”, *THES*, 5 March.
- Corbyn, Z. (2010), “Retractions Up Tenfold”, *Times Higher Education*, 22 March, [www.timeshighereducation.co.uk/story.asp?storycode=407838](http://www.timeshighereducation.co.uk/story.asp?storycode=407838).
- Debackere, K. and W. Glänzel (2004), “Using a Bibliometric Approach to Support Research Policy Making: The Case of the Flemish BOF key”, *Scientometrics*, 59(2), pp. 253–276.
- DfES (2006), “Reform of Higher Education Research Assessment and Funding: Summary of Responses”, [www.dcsf.gov.uk/consultations/downloadableDocs/RAE%20response%20summary%20250107.doc](http://www.dcsf.gov.uk/consultations/downloadableDocs/RAE%20response%20summary%20250107.doc).
- Elzinga, A. (2009), “Evidence-based Science Policy and the Systematic Miscounting of Performance in the Humanities”, *The Humaniora/Society Blog*, <http://humaniorasociety.wordpress.com/2009/04/28/evidence-based-science-policy-and-the-systematic-miscounting-of-performance-in-the-humanities/>.
- European Commission (2010), *Assessing Europe’s University-Based Research*, European Commission, Brussels.
- Evaluation Associates Ltd (1999), *Interdisciplinary Research and the Research Assessment Exercise*, HEFCE, [www.rae.ac.uk/2001/Pubs/1\\_99/](http://www.rae.ac.uk/2001/Pubs/1_99/).
- Frölich, N. (2008), *The Politics of Steering by Numbers: Debating Performance-based Funding in Europe*, NIFU STEP, Oslo.
- Gläser, J. (2008), “The Social Orders of Research Evaluation Systems”, in R. Whitley and J. Gläser (eds.), *The Changing Governance of the Sciences*, *Sociology of the Sciences Yearbook*, Springer, Dordrecht, pp. 245–266.
- Gläser, J. *et al.* (2002), “Impact of Evaluation-based Funding on the Production of Scientific Knowledge: What to Worry About, and How to Find Out”, *Expertise for the German Ministry for Education and Research*, Vol. 31.

- Hazelkorn, E. (2007), “The Impact of League Tables and Ranking Systems on higher education decision making”, *Higher Education Management and Policy*, Vol. 19(2), p. 25.
- Hazelkorn, E. (2009), *Pros and Cons of Research Assessment: Lessons from Rankings*, Dublin Institute of Technology, Directorate of Research and Enterprise, Dublin.
- HEFCE (Higher Education Funding Council of England) (1997), *The Impact of the 1992 Research Assessment Exercise on Higher Education Institutions in England*, [www.hefce.ac.uk/pubs/hefce/1997/m6\\_97.htm](http://www.hefce.ac.uk/pubs/hefce/1997/m6_97.htm).
- Henty, M. (2007), “Ten Major Issues in Providing a Repository Service in Australian Universities”, *D-Lib Magazine*, Vol. 13(5/6), [www.dlib.org/dlib/may07/henty/05henty.html](http://www.dlib.org/dlib/may07/henty/05henty.html).
- Jiménez-Contreras, E., F. Anegón and E. López-Cozar (2003), “The Evolution of Research Activity in Spain – The Impact of the National Commission for the Evaluation of Research Activity (CNEAI)”, *Research Policy*, Vol. 32(1), pp. 123-142.
- Li, B., J. Millwater and P. Hudson (2008), “Building Research Capacity: Changing Roles of Universities and Academics”, paper presented at the Australian Association for Research in Education (AARE) 2008 International Education Research Conference, Brisbane.
- Lipsett, A. and T. Tysome (2006), “A Nod to New Blood and a Warning to Game Players”, *THES*, 3 February.
- MacLeod, D. (2006), “Made to Measure”, *The Guardian*, 31 January.
- McNay, I. (1998), “The Research Assessment Exercise (RAE) and After: ‘You Never Know How It Will All Turn Out’”, *Perspectives: Policy and Practice in Higher Education*, Vol. 2(1), p. 19.
- Ministry of Education, New Zealand (2008), “How the PBRF Has Shifted Research Funding”, [www.educationcounts.govt.nz/\\_data/assets/pdf\\_file/0011/18848/Funding\\_factsheet.pdf](http://www.educationcounts.govt.nz/_data/assets/pdf_file/0011/18848/Funding_factsheet.pdf).
- Moed, H.F. (2008), “UK Research Assessment Exercises: Informed Judgments on Research Quality or Quantity?”, *Scientometrics*, Vol. 74(1), pp. 153-161.
- New Zealand, Tertiary Education Commission (2010), Performance-Based Research Fund – Tertiary Education Commission, [www.tec.govt.nz/Funding/Fund-finder/Performance-Based-Research-Fund-PBRF/](http://www.tec.govt.nz/Funding/Fund-finder/Performance-Based-Research-Fund-PBRF/).

- Newman, M. (2009a), “Liverpool May Drop Subjects Ranked Below 4\* in RAE”, *THES*, 12 March.
- Newman, M. (2009b), “RAE Panel Hurt Field by Labelling It ‘Weak’“, *THES*, 26 March.
- Newman, M. (2009c), “Reviewers Raise Concerns about RAE Game Playing”, *THES*, 8 January.
- Prieto Valverde, J.A. and M. Fernández Pérez (2010), “World Universities’ Ranking on the Web 2010”, [www.webometrics.info/](http://www.webometrics.info/).
- RIN (Research Information Network) (2009), *Communicating Knowledge: How and Why Researchers Publish and Disseminate Their Findings*, September, [www.rin.ac.uk/communicating-knowledge](http://www.rin.ac.uk/communicating-knowledge).
- Roberts, G. (2003), *Review of Research Assessment*, Report to the UK funding bodies, [www.ra-review.ac.uk/reports/roberts.asp](http://www.ra-review.ac.uk/reports/roberts.asp).
- Sanders, C. (2005), “20% of Staff Now Just Teach”, *THES*, [www.timeshighereducation.co.uk/story.asp?storyCode=196862&sectioncode=26](http://www.timeshighereducation.co.uk/story.asp?storyCode=196862&sectioncode=26).
- Schneider, J.W. (2009), “An Outline of the Bibliometric Indicator Used for Performance-Based Funding of Research Institutions in Norway”, *European Political Science*, Vol. 8(3), pp.364-378.
- Sivertsen, G. (2008), “Experiences with a Bibliometric Model for Performance Based Funding of Research Institutions”, in Gorraiz J. and E. Schiebel (eds), *Excellence and Emergence: A New Challenge for the Combination of Quantitative and Qualitative Approaches*, Book of Abstracts, 10th International Conference on Science & Technology Indicators, Vienna, Austria.
- Sivertsen, G. (2009), “A Bibliometric Funding Model Based on a National Research Information System”, [www.issi2009.org/agendas/issiprogram/activity.php?lang=en&id=108](http://www.issi2009.org/agendas/issiprogram/activity.php?lang=en&id=108)
- Sivertsen, G. (2010), “A Performance Indicator based on Complete Data for the Scientific Publication Output at Research Institutions”, *ISSI Newsletter*, Vol. 6(1), pp. 22-28.
- Smart, W. (2009), *The Impact of the Performance-Based Research Fund on the Research Productivity of New Zealand Universities*, [www.educationcounts.govt.nz/publications/tertiary\\_education/49897](http://www.educationcounts.govt.nz/publications/tertiary_education/49897).

- Strehl, F. (2007), “Funding Systems and Their Effects on Higher Education Systems – International Report”,  
[www.oecd.org/publicationanddocuments/0,3395,en\\_33873108\\_33873681\\_1\\_1\\_1\\_9\\_1,00.html](http://www.oecd.org/publicationanddocuments/0,3395,en_33873108_33873681_1_1_1_9_1,00.html)
- Tapper, T. and B. Salter (2003), “Interpreting the Process of Change in Higher Education: The Case of the Research Assessment Exercises”, *Higher Education Quarterly*, Vol. 57(1), pp. 4-23.
- UK Department of Health (2001), *Joint Funding Bodies’ Review of Research Assessment: Response from the Department of Health to the Public Consultation Launched in September 2002*,  
[www.hefce.ac.uk/pubs/hefce/2001/01\\_64.htm](http://www.hefce.ac.uk/pubs/hefce/2001/01_64.htm).
- UK Government (2006), *Science and Innovation Investment Framework 2004-2014: Next Steps*, [www.dius.gov.uk/policies/science/science-funding/ten-year-framework/next-steps](http://www.dius.gov.uk/policies/science/science-funding/ten-year-framework/next-steps).
- University Grants Committee (UGC), Hong Kong, China (2006), *Research Assessment Exercise 2006 Guidance Notes*,  
[www.ugc.edu.hk/eng/ugc/publication/prog/rae/rae.htm](http://www.ugc.edu.hk/eng/ugc/publication/prog/rae/rae.htm).
- Whitley, R. (2008), “Changing Governance of the Public Sciences”, in R. Whitley and J. Gläser (eds.), *The Changing Governance of the Sciences*, *Sociology of the Sciences Yearbook*, Springer, Dordrecht.
- Whitley, R. and J. Gläser (eds.) (2008), *The Changing Governance of the Sciences*, *Sociology of the Sciences Yearbook*, Springer, Dordrecht.



**From:**  
**Performance-based Funding for Public Research in  
Tertiary Education Institutions**  
Workshop Proceedings

**Access the complete publication at:**  
<https://doi.org/10.1787/9789264094611-en>

**Please cite this chapter as:**

Butler, Linda (2010), "Impacts of performance-based research funding systems: A review of the concerns and the evidence", in OECD, *Performance-based Funding for Public Research in Tertiary Education Institutions: Workshop Proceedings*, OECD Publishing, Paris.

DOI: <https://doi.org/10.1787/9789264094611-7-en>

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area. Extracts from publications may be subject to additional disclaimers, which are set out in the complete version of the publication, available at the link provided.

The use of this work, whether digital or print, is governed by the Terms and Conditions to be found at <http://www.oecd.org/termsandconditions>.