

NZCPR Research

NEW ZEALAND CENTRE FOR POLITICAL RESEARCH

26 MARCH 2011

PERFORMANCE-BASED RESEARCH FUNDING: WHY IT SHOULD END NOW

By Dr Ron Smith University of Waikato

New Zealand Universities are now well into a third round of Performance-Based Research Funding (PBRF) evaluations which on present trends will be even more protracted and time-wasting than the previous two. Having regard to the manifest defects of the process, both conceptually and practically, and to the likelihood that it has produced no net financial benefit to the universities, this is altogether to be regretted. Indeed, it is a sad testimony to the tolerance or powerlessness of university staffs, that PBRF has persisted this long. It is to be hoped that it gets no further and that the 2012 round will be the last. At very least, the reliability and validity of the judgements made, ought to be subjected to independent scrutiny before any decision is made to continue what is an entirely lamentable practice.

Performance-Based Research Funding is a device to channel some proportion of higher education funding through an all-embracing process to evaluate the worth of academic research and reward ‘outputs’ that are judged to be most worthy. It began in New Zealand in 2003 and was followed by a second round of evaluations in 2006. The process was initially based on the UK Research Assessment Exercise (RAE).

The concept

At the outset, the PBRF process must assume that there is a single model of what constitutes research, or otherwise there would be no valid basis for the comparisons of value that it entails. It will be argued here that there is no such single model. What is

embraced within the term ‘research’ is so diverse in its character, as to make judgments of the kind inherent in PBRF systematically invalid. To add into the performance assessment second order factors such as evidence of ‘peer-esteem’ and ‘graduate completions’ only adds to the weight of arbitrary and subjective factors and makes the consequent ‘scores’ progressively less valid and reliable.¹

Some activities which give rise to new knowledge take place within a framework of assumptions that is virtually unchallenged: this will apply to a great deal of scientific and medical research. Within such a stable paradigm, there is relatively little difficulty in determining the value of a new piece of work. The process of ‘peer-review’ can work well. It can assure readers that appropriate standards of data collection, interpretation and presentation were used and the results produced cohere with other results in the same domain.² That said, it should be noted that even here there may be grounds for doubt where the underlying scientific data has strong ideological or political implications. The revelations of malfeasance at the University of East Anglia, with regard to climate data and publication about climate change, show this very clearly. In this case, the peer-review process was manipulated to favour papers that supported the anthropologically-caused warming hypothesis and to block the publication of those that did not. More generally, there is also an increasing recognition that there are human factors (apart from dishonesty) which can distort results through the incorporation of unconscious bias of one sort or another.³

The situation is different, again, in other areas of academic endeavour, such as the Arts and Social Sciences. Here, there may be no agreed paradigm at all. Indeed, there may be several fiercely competing ways of looking at what is claimed to be valid or true. In this context, the process of peer evaluation is not at all reliable as a determiner of worth. Judgements in this domain tend to reflect the prejudices of the judges as much as they do any objective concept of value. The same phenomenon is noted in the context of Economics by an American writer commenting on the effects of the corresponding

¹ 40% of an individual’s PBRF rating comes from factors such as these. Peer-esteem data is gathered by the individual him or herself and may be based on such things as invitations to give papers at a conference. The number of Masters degrees or PhD completions by students with whom the individual academic is associated also counts, as does the amount of money that may have been collected from official sources to support their intellectual endeavours.

² See <http://www.climategate.com/> for multiple references.

³ Jonah Lehrer, ‘The Truth Wears Off’, *The New Yorker*, 13 December, 2010.

process in the UK. He concludes that RAE has had the effect of almost eliminating all but conventional economics from British universities.⁴

A similar problem in judging the worth of claims to new knowledge, also arises in the sciences when research begins to challenge the paradigm itself. In this case, findings may be dismissed, or even ridiculed by the orthodox.

The history of intellectual activity, both in universities and elsewhere, is full of examples of ‘outputs’ that were dismissed by the ‘experts’ at the time and only seen to be insights of great value at some later point. The German scientist, Wegener, provides a Twentieth Century example, through the response of the scientific community to his notion of continental drift. For some sixty years the theory was derided by the majority of the geophysical community and papers supporting it were declined for publication by leading journals. Here is the comment of the Editor of the *Journal of Geophysical Research* (peer-reviewed, of course) rejecting a paper by Lawrence Morely describing (in 1963) the mechanism of continental drift: “Such speculations make interesting talk at cocktail parties, but it is not the sort of thing that ought to be published under serious scientific aegis.”⁵

It is also essential to note that, for some kinds of research, time horizons may be very long and the product may not easily lend itself to an ordered sequence of academic papers, or even very many papers at all. It may nonetheless ultimately turn out to be of enormous value. An outstanding example of this is provided by the life and work of the philosopher, Ludwig Wittgenstein, who published only one slim volume in his life-time (the central thesis of which was subsequently repudiated by him) and whose enormously influential major work (on how language acquires its meaning) only appeared after his death.⁶

Research of this quite fundamental kind may be intellectually speculative, or practically difficult, or, possibly, both. This latter was the case with Marie Curie who spent many years apparently getting nowhere, as she grappled with the twin problems of devising

⁴ Frederic S Lee, ‘The Research Assessment Exercise, the state and the dominance of mainstream economics in British universities’, *Cambridge Journal of Economics*, 31, 2007.

⁵ Cited in Bill Bryson, *A Short History of Everything*, Black Swan Books, 2003, page 228.

⁶ The books concerned are, *Tractatus Logico-Philosophicus*, published (in English) in 1922 and *Philosophical Investigations*, published in 1953, two years after his death.

methods of chemical separation and evolving a concept of matter which explained the differing activity of what came to be known as radioactive isotopes. Of course, in the end, the worth of her endeavours was recognised by the award of two Nobel Prizes (Physics, 1903; Chemistry, 1911). But how would she have fared under PBRF? Clearly, she would have been highly-rated in mid-career but would she have even got that far?

There are a number of conclusions that might be drawn from this, albeit brief, discussion. There are different *kinds* of research and they are essentially incommensurable. Thus they cannot be made the basis of comparative judgements of intrinsic worth. Similarly, assessments about the intellectual worth of ‘pure’ research cannot be made securely at the time that the work is projected, or even when it is first reported. A process that claims to do this is thus bound to miss some outcomes that time will show to have been of the greatest importance and, conversely, privilege the popular and the ephemeral, such as the contemporary torrent of post-modernist commentary. Insofar as more speculative activity is precisely the kind of research traditionally associated with universities, the acceptance of PBRF is a grievous error. University College, London, academic, David Gillies, talks about this as a ‘systemic failure’ and gives detailed examples, additional to those given here.⁷ More generally, the defects of the system (errors and prejudices) inevitably entail an uncontrolled and uncontrollable injustice, both to scholarship and scholars.

Implementation

The PBRF process is equally flawed at the implementation level. The central tool here is the ideal of the peer-reviewed journal article and its extension, the ‘quality assured’ publication. Again, the range of what might be fitted into the category is so wide as to make the outcome unreliable for any serious purpose. At one end of the spectrum we have the long established international journal, with a formal refereeing process and a record of rejecting a high proportion of the articles offered to it. At the other end (but still accepted as quality assured), is the occasional issue of a periodical of which the editors, and almost all its contributors, are from the same faculty in the same university. More generally, there is the systematic problem of recognition for academic journals and the lack of universally applied standards of judgment. Incidents of the acceptance of hoax articles by prestigious publications also tend to cast doubt on the universal validity

⁷ Donald Gillies, ‘Why Research Assessment Exercises Are a Bad Thing’, *post autistic economics review*, Issue No 37, 28 April 2006.

of the process.⁸ To combine such disparate outputs is to seriously mislead and, to the extent that it is inherent in the process, to seriously undermine the validity of any conclusions that might be drawn. Comparisons between the outputs of individuals are also completely vitiated by the practice of counting *each name* on a publication as a *full credit*.

Consequences

One serious consequence of the process that was noticed and commented on at Waikato University after the first PBRF round, was the humiliation of the substantial proportion of academic staff whose research was evaluated and found wanting. What has been the effect of this (now once repeated) experience on the enthusiasm and dedication of those so treated? Reports on the practice of streaming in secondary schools have regularly shown inferior performance and low self-esteem in pupils categorised as low-achievers. Why would the effect on university staffs be substantially different?⁹

This is the point that Malcolm (former Waikato Vice-Chancellor) and Tarling make in their recent book on the mission and management of New Zealand universities, when they refer (in the context of processes of review) to:

“a feeling of helplessness, of alienation, even at times of fear, that seems to us utterly alien to the proper spirit of a university, and utterly incompatible with its proper aspirations.”¹⁰

⁸ A celebrated example of this is to be found in the 1996 Spring/Summer issue of the journal *Social Text*. The article ‘Transgressing the Boundaries: Towards a Transformative Hermeneutics of Quantum Gravity’, was subsequently revealed by its author, Alan Sokal, to have been a complete spoof. *Social Text* continues to be published by Duke University Press.

⁹ There is some data on this point from a survey conducted by staff of the Labour Studies Department of Waikato University, following the first PBRF round. In this, two-thirds of those responding thought that PBRF was negative for staff morale. (Cochrane, Law and Ryan, ‘The 2003 PBRF Experience: A Survey of Academic Staff at the University of Waikato’, *Centre for Labour and Trade Union Studies*, December 2005.

¹⁰ Wilf Malcolm and Nicolas Tarling, *Crisis of Identity; the Mission and Management of Universities in New Zealand*, Dunmore Publishing, 2007.

This sentiment has also been echoed in a 2011 review of the situation in British universities. In this, a young history academic at a London university describes the ‘bureaucratization of scholarship in the humanities’ as ‘simply spirit crushing’.¹¹

It is not difficult to imagine how staff outside the favoured circles feel, as they increasingly come to understand the flawed nature of the process by which these insulting judgments were made. On the other hand, Professor Cris Shore of Auckland University reports (on the basis of individual interviews) that ‘On balance most staff appeared to support PBRF ... primarily ... because it recognises research as a major element of what universities do’.¹²

In addition, PBRF has corrupted the whole academic institution. We now have a whole apparatus of ‘portfolio managers’, standing ready to advise staff on how to present their efforts to best advantage, how to garner expressions of ‘esteem’, and, above all, how to appeal to the prejudices of those who will sit in judgment. In the political context we would talk of ‘spin-doctors’, or ‘sexing-up’, or simply of deceit.

Since the PBRF process turns particularly on peer-reviewed publication, there has been an explosion of new publication opportunities. Institutions have increasingly been tempted to set up their own journals, with their own ‘peers’ to review them. Indeed, this was the local realist response after the first round. In the PBRF world it is now more important than ever that ambitious individuals attach themselves to influential support groups (‘clusters’, may be the new term) and ‘toe the party line’, to make sure that they are attractive to the cabals that control assessment, or (more proactively) that they contrive to get power themselves.

Then there is the matter of who gets *counted* in the PBRF census. Examples are now accumulating of institutional efforts to hide poor performers and appoint high-flyers for the duration of the assessment.¹³ More generally, PBRF mandates the appointment of persons simply for their ranking and these may be persons who don’t teach or who

¹¹ Simon Head, ‘The Grim Threat to British Universities’, *The New York Review of Books*, January 13, 2011. (<http://www.nybooks.com/articles/archives/2011/jan/13/grim-threat-british-universities/>)

¹² Cris Shore, ‘Playing the tune without the piper? The implications for academic researchers of the changing funding environment’, *Winter Lecture Series*, University of Auckland, 2008,

¹³ Cris Shore (*Ibid*) reports the ‘poaching of star researchers’ and ‘attempts (by some universities) to hide less research-active staff’.

teach with little enthusiasm or interest because they understand that it is an activity that is not valued.

More important, institutional comparisons, based on PBRF scores, will give no indication of the quality of the teaching at the various locations. Indeed, the whole process is likely to encourage the continuation of a long-term trend (overseas as well as here) against valuing university teaching. In order to maximise research 'scores' and money, the teaching is actually done by junior faculty members and senior students. There is another unfortunate side effect of all this and that is to further downgrade academic contributions to what has been seen as the 'public good', or even as 'scholarship' rather than 'research', since none of this activity is significantly valued in the formal process. Indeed, the situation may be much worse than this, with an increasing number of commentators, academic and other, supplying examples where celebrated scholars would have struggled for formal recognition. Here is Roger Brown, Vice-chancellor and professor of higher education policy at Southampton Solent University (UK):

I have it on good authority that Albert Einstein would not even have been considered for entry into the RAE* because his work on relativity would have counted as "scholarship" rather than "research."¹⁴

(* Research Assessment Exercise; the UK equivalent of PBRF)

The final outcome (the PBRF report card) then produces the undignified spectacle of Vice-Chancellors and their deputies bragging about the academic accomplishments of their institutions and their staff. It may be doubted whether the academic reputation of any university (or universities in general) is likely to be advanced by such a process. It also might be argued that the academic standing of a particular institution is actually dependent on the accomplishments of a relatively small number of prominent individuals, who do not need the shrill call of the PR department to establish their reputation. By contrast, no one who knows how the data has been arrived at will be impressed by claims that university X had ten of the top performing departments, as revealed by PBRF scores¹⁵. In addition to all this, it is clear that not only will PBRF fail to

¹⁴ *Times Education Supplement*, 15 June 2007

¹⁵ The systemic potential for distortion was well-illustrated in the 2010 world university rankings compiled by the UK *Times Higher Education*. In this, the University of Alexandria (Egypt) was placed just below the University of Auckland (the only New Zealand university in the top 200) sheerly on the basis of the

achieve its stated objective to raise the level of academic achievement, it will actually inhibit the process by marginalising the sort of creative non-conformity that can lead to great advances in knowledge and, perhaps, institutional reputation. The evidence on this point is continuing to accumulate. It really is about time university leaders (Vice-Chancellors in particular) took their responsibility to scholarship more seriously.

Then there is the extremely important matter of cost. In the case of PBRF, this applies not only to the direct expenditure of operating the scheme both locally and centrally, but also to the second-order costs of staff time devoted to compliance. Nationally, the cost of the first PBRF exercise has been put at up to \$28 million.¹⁶ The same authors also conclude that the value of the overall reallocation of research funds (which is what the exercise is supposedly about) is probably swallowed up in the ‘transaction costs’ involved. In some universities the cost of PBRF is now compounded by the additional imposition of a ‘formative’ exercise. This will surely make PBRF a net loss, not only in terms of financial support for research but also in terms of the time and enthusiasm that staff have for the exercise.

The argument from appeasement

Continuing cooperation with PBRF is sometimes defended on the basis that this is what the Government has decreed and there is nothing that can be done about. The latter claim is simply not true. PBRF would collapse if the universities of New Zealand refused to have anything to do with it. It would be a matter of the individual institutions (and, particularly, their leaders) having the courage of their convictions. A recognition of the conceptual defects of PBRF, together with its corrupting effect on the academy and the injustice that it continues to inflict on staff, ought to take care of the ‘conviction’ part. It is then just a matter of the ‘courage’. If it is not the right way to manage a university, it is not made so by being financially rewarded. It also might be observed that PBRF cannot be justified on the grounds that (as has been claimed) it helps universities who score well in PBRF with their recruiting, both of staff and of students. As with the

apparent performance of a single member of staff who had published in the review period no less than 320 articles in a journal of which he was editor. (<http://www.timeshighereducation.co.uk/>)

¹⁶ Tim Hazeldine and Cliff Kurniawan, ‘Impact and Implications of the 2003 Performance-Based Research Fund Research Quality Assessment Exercise, in Leon Bakker et al (eds), *Evaluating the Performance-Based Research Fund: framing the Debate, Institute of Policy Studies*, 2006, page 270.

corresponding argument that it advantages academic staff who score well, the ends here cannot be justified if the means cannot. On the other hand, it is understood that those (both university administrators and staff) who see themselves as benefitting in one way or another from the system will be more inclined to see its virtues and ignore its defects.¹⁷

Staff to whom this does not apply may feel that, whatever their reservations or resentments, there is nothing they can do about the continuing imposition of PBRF. Here, Shakespeare puts the matter plainly:

Men at some time are masters of their fate: The fault, dear Brutus, is not in our stars but in ourselves that we are underlings. (*Julius Caesar*, Act I Scene 2)

The imposition of a ‘formative exercise’, followed by another round of PBRF cannot proceed without the cooperation of the generality of academic staff. If it is accepted that the process cannot be justified on grounds of academic principle or human rights, then it ought not to be supported and this would apply even if there were grounds to believe that individuals and particular institutions may be advantaged by its continuance.

A different argument for PBRF is sometimes offered. In this, the point of the exercise is simply to make staff work, otherwise they would just slack. Having regard to the nature of the work, it is not evident how data would be obtained on this point, or, indeed, whether the managers who espouse this theory actually have any evidence, but it is clear that a regime that sets out to establish easily measurable targets in the academic domain is likely to significantly influence outcomes, and not necessarily for the better. As argued earlier, the system is encouraging an outcome torrent, the quality of which there is every reason to doubt. This is what Broadhead and Howard say, after acknowledging that in Britain academics ‘appear to be publishing with greater frequency’:

Producing more articles, however, is not the same as doing more research. The regurgitation and multiple-placing of articles is on the increase. This process, although intellectually un-taxing, is time-consuming, reducing time and energy available for both

¹⁷ Strictly speaking, PBRF scores were not to be used for other than collective institutional assessment and not for assessment of individual performance. It is now widely recognised that this is a principle that is now only honoured in the breach.

fresh research and course review. Moreover, as more is being published, recent studies suggest that less is being read.¹⁸

This point was also made by the incoming Director of the (UK) Institute of Historical Research, Professor David Cannadine, in his 1999 inaugural lecture¹⁹. In this he referred to the two thousand books and nearly five thousand articles produced by history academics in the previous year, describing this output as not only ‘prodigious’ but ‘preposterous’, as it inflated the quantity and deflated the quality. Of course Professor Cannadine knew *why* it was being done. It was to satisfy the requirements of the Research Assessment Exercise and probably the only people who read this copious outpouring were the History panel themselves.

There is also evidence of an increasing amount of academic ‘self-plagiarism’ as documented in a study by Australian academics Tracey Bretag and Saadia Carapiet.²⁰ In this, desperate staff endlessly re-write material that has already been published, in order to score a further credit.

An act of faith

Support for research in a university is essentially a matter of faith. Of course, there ought to be an expectation that research activity would underpin good teaching and perhaps some contribution to public discourse on pressing matters of public policy (the critic and conscience obligation). It might also result in some substantial contribution to the advance of human knowledge (or ‘scholarship’, if this is different), although this is less certain and, as has been argued, it may not be at all clear at the time that such an advance has been made. The only question is what proportion of an academic’s time ought to be allowed for this purpose? The present understanding appears to be that it is 40%. In the light of the increasing flow of footling research that the PBRF regime seems to encourage, this may be too high. It may be that we should consider a more

¹⁸ Lee-Anne Broadhead and Sean Howard, “‘The Art of Punishing’: The Research Assessment Exercise and the Ritualisation of Power in Higher Education”, *Education Policy Analysis Archives*, Volume 6, Number 8, 1998.

¹⁹ <http://www.history.ac.uk/ihr/Focus/Whatishistory/cannadine.html>

²⁰ Reported over the by-line Rebecca Attwood in *Times Higher Education*, 3 July 2008.

flexible system, in which local judgments might be made about the contribution that individuals might make, which would not be concerned solely with ‘form’ publication. Overall, though, in the modern university, it may be that less time should be allocated for ‘research’ than the allocation that we have inherited from an earlier time, when a far smaller proportion of the age-group attended and the institutions (and staffs) were correspondingly far smaller.

In such a situation there would be much less of the commissioned, or ‘goal-directed’, research that universities have been encouraged to take on in order to balance their books. Research in universities would then be undertaken because individuals were drawn to unanswered questions that arose from their studies. A formula of this kind might also result in academic staff overall doing more teaching and less ‘research’, which might enable governments (societies) to pay them better because there would be ultimately fewer of them. Clearly governments are reluctant to adequately pay now. This is a discussion that we really need to have.

To suppose that valid and fair judgments can be made about the comparative value of individual research across the academic spectrum is to defy the facts and institutionalise injustice.

The announcement that the University of Waikato was to have an ‘in-house’ Formative Research Exercise as the lead in to the 2012 PBRF round was made in the 13 May 2008 edition of a Waikato internal publication *FASS-E-News Today*. This means that the institution has been in a PBRF frenzy since that time, and there is more than a year to go. Even if there were some virtue to the process (which, in the view of this author, there is not), it is too much. Too much time and effort put into gaining results of such doubtful value.