Critique of research assessments

Jan Currie
Murdoch University
J.Currie@murdoch.edu.au

This article introduces the issue of articles on research assessments and rankings. It focuses on the British Research Assessment Exercises (RAEs) and New Zealand’s Performance-Based Research Fund (PBRF). It contrasts these with systems that are not tied to funding formulas and emphasise quality enhancement rather than quality assurance. It indicates the negative consequences for individuals, institutions and the professional ethos of universities. It does not deny that there are also benefits gained by institutions that win additional funding through these mechanisms. Research managers benefit by being able to concentrate the efforts of researchers and gain more control over the research enterprise. However, scholarship and creativity are the likely losers in these exercises.

[Key words: research assessments, quality, accountability, universities, globalisation]

Portions of this article were reproduced from “Critique of Research Assessment Exercises and Their Consequences for Academics: Meta-analysis of British, Hong Kong and New Zealand Systems” presented at the ANZCIES Conference in Auckland, November 30, 2007.

INTRODUCTION

The word “globalisation” is pregnant with many, often contradictory, meanings. There is an openness to globalisation that keeps us wondering whether it is for the better or worse. In China, for example, globalisation is often interpreted as opening the doors to the West, to new ideas. In contrast, the local is a more constrained space; it is already well constructed by history, by geography, by culture, by demography, by religion and ethnicities. Globalisation creates a space that is not completely constructed; we are not certain where this space will lead us. One of the basic tools of globalisation, the Internet, brings individuals across the world together and should ideally lead to greater collaboration to fight the ills of society, whether these are diseases, global warming or poverty. Yet as we know, it can also lead to constraining individuals by its ability to track and control people.

Another image that comes with globalisation is an open economy, an unregulated economy, a nation without tariffs, a borderless world. This refers to globalisation’s underlying ideology of neoliberalism, let the markets decide. This often pits one country against another and universities around the world against each other. Academics compete to be the first with a discovery, the first to publish it, and then
the first to commercialise it. The spirit of collaboration or working together for the public good becomes trampled in the rush to be first.

Universities have always been competitive by nature, with individuals wanting to excel and push the boundaries of knowledge. And this is praiseworthy. Universities should be creating research and teaching cultures of excellence. Are there ways of developing research cultures of excellence and dynamic teaching cultures in our universities that can also enable them to collaborate and more effectively contribute to the public good?

It appears that many policy makers decide that research assessment exercises are necessary to encourage their universities to compete more effectively in international ranking exercises. Realistically, the universities that can compete in these international rankings are comprehensive research-intensive institutions. “The fact is that essentially all of the measures used to assess quality and construct rankings enhance the stature of the large universities in the major English-speaking centres of science and scholarship and especially in the United States and the United Kingdom” (Altbach, 2006: 42). Marginson (2006: 27) summarised this in his depiction that “the model global university is English speaking and science oriented”.

This limits the number of universities that can truly compete to be world class, another idea that has captured the imagination of policy makers. According to Birnbaum (2006), entering “world class university” in Google produces over 150,000 hits. He noted that even the University of Timbuktu (which apparently was a world-class university in the 12th century) has announced its intention of regaining that status.

Global university rankings (SJTIHE, 2006; Times Higher Education Supplement, 2005) that produce the league tables and the notion of the world-class university also provide a powerful impetus for individual universities to become global actors. The leaders of universities become more fixated on achieving this world-class status and moving up in the league tables than thinking about their local communities. This is an indicator of where the global displaces the local.

QUALITY MEASURES AND NATIONAL RANKINGS

Education ministers and national policy makers are driven to ensure that within their national systems, they have at least one or two universities that are ranked in the top 100 and are considered world class institutions. Sidhu’s article in this issue discusses the pressure created by these national ranking systems. She describes the type of institutional identities that are favoured by quality regimes and the tendency for these regimes to re-conceptualise education as a globally tradeable service.

In surveying the quality assurance mechanisms used in North America, Europe and Oceania, policy makers have veered in at least two different directions to make their universities more competitive: one, an external, national accountability system that is based on metrics, peer assessment of research publications, league tables and
funding formulas and the other, internal based on a rigorous tenure system with evaluation systems focused on improving the culture but without rankings linked to funding formulas. There is in actuality a continuum of measures from quality enhancement to quality assurance. On the quality enhancement side, the focus is on improvement and is based on trust; whereas with the quality assurance side, it is based on proving and is more likely to engender mistrust.

In this introductory essay, I critique the research assessment exercises that have taken place in the United Kingdom and New Zealand with funding attached to them and briefly contrast them with North American and European quality regimes that for the most part are not tied to funding. The former research assessment systems are on the quality assurance side of this dichotomy. They are based on a compliance model where universities have to prove through performance indicators that they have achieved a certain level of excellence and they are rewarded monetarily. These exercises have existed in a variety of forms for at least two decades. The British developed the Research Assessment Exercise (RAE) in the early 1980s and after six assessments decided to abolish it after their 2008 round. New Zealand began its Performance-Based Research Fund (PBRF) assessment process in 2003 and had another one in 2006. Hong Kong’s RAE was modelled on the British one and is in the same compliance mode tied to funding. In contrast, in North America and Europe, there are as yet no national assessments of research that are tied directly to funding formulas. There are national assessments in some European countries but these are mainly used for internal improvement of universities rather than for the distribution of funding and are more aptly described as quality enhancement mechanisms.

This issue examines the underlying rationale for introducing quality regimes and how they become embedded into national systems. It describes how these initiatives were drawn from the field of accounting and the audit society (Singh’s article). Although there is a degree of convergence in this worldwide trend, countries take slightly different paths to maintain high levels of research productivity. Vidovich describes the path taken by Singapore in this issue. She discovered that although it is veering closer towards a compliance model, it has given its universities more autonomy and chosen a low compliance path resembling the American model rather than the British one. Interestingly, those countries that do not have a high compliance model tied to funding seem to do better or equally as well in the national ranking of universities. As reported by Moodie (2008), a new ranking by Taiwan awarded the United States 170 universities in its top 200, Germany 45 and the United Kingdom 39. Canada and Italy had 20 each and France 17. Australia had 13, one or two more than the Netherlands, Sweden and Spain. Out of all these nations, it is only the United Kingdom and Australia that tie research performance indicators to formula funding in national systems. The other European and North American systems do not have a national assessment of research that is tied to funding.
INDIVIDUAL BEHAVIOUR, CONSEQUENCES AND MOTIVATION

A colourful language has developed to describe the behaviour of individuals and universities in their competition for research funding: ‘salami slicing’, ‘shadow publishing’, ‘poaching’ or ‘cherry picking’ academics in a transfer market, ‘cannibalisation’ of junior members of staff to increase their teaching loads for the benefit of active researchers, a swirling non-stop research roundabout that favours short-termism in research projects and products. Newspaper headlines tell a similar story: it’s a “Ratings Game”, the “RAE forces us to fudge”.

‘Salami slicing’ is well described in this quote from a City University of Hong Kong academic: “People try to publish quantity. So sometimes they cut a paper into different bits and submit it to different journals. Thus one paper can generate five papers. In the past, you would submit it as a whole” (Yang, Vidovich & Currie, 2007: 11–12).

Illing (2006) suggests that “Poaching is rife in New Zealand universities following that country’s first national research assessment exercise”. Birkhead (2007) writes that in the United Kingdom “By cherry-picking only those academic staff to submit distorts the process and demoralises excluded staff”. In the latest figures for the 2008 RAE in the United Kingdom, only 46 per cent of research-active staff were submitted. In the most extreme cases, the University of Chichester submitted work by just 18.5 per cent of its academics compared with 60 per cent in 2001 and Keele University was down from 80 per cent to 50 per cent (Times Higher Education Supplement, 5 February 2008).

Sparkes (2007) from the University of Exeter submitted a paper to be judged by the RAE panel in 2008 that criticised the whole process. He used a narrative technique to demonstrate the destructive effects on academics’ health of the audit culture pervading universities. He describes a department head’s meeting with a junior member of staff who was considered research inactive by the RAE process. In a colourful response, the young academic said, “You made me feel like shit. You really did. You made me feel like my research didn’t count. That I didn’t count”.

Policy analysts have warned about the likely consequences of narrow systems of compliance that universities have adopted in countries such as Britain, New Zealand, Australia and Hong Kong. “There is very little top-cover for risk-takers when institutional league tables, bottom lines and reputations have to be guarded” (Nile, 2008). Marginson (2008:33) argues that a ground-breaking approach to innovation would be to foster creativity and this would mean moving beyond the standard policies. “It would mean examining the factors that attract highly creative people to particular places in numbers and diversity, and the freedom they need to produce brilliant ideas and talk about them in engaging manner”.

In a similar vein, academics in the United Kingdom have expressed concerns about the negative consequences of the RAE. Baty (2007) said that it distracts academics from pursuing ground-breaking ideas in favour of low-risk options and Williams
(1998) felt it threatened to crush research creativity, careers, and scientific integrity. In New Zealand the same concerns are raised for new and emerging scholars and for the demoralising effects that the R rating (equivalent to being seen as an inactive researcher) has on new researchers. Of course there are particularly vulnerable groups that seem to lose out in the ratings: in both countries it is female and younger academics and in New Zealand, it is Maori staff and those whose research serves professional, social, cultural and governmental communities (Clarke, 2005; University of Auckland Professors, 2004; Boston, 2004). Harley (2003) noted that men were twice as likely to be entered for the RAE in the United Kingdom as women during the 1990s. Using data from an Australian university, Austen’s article in this issue examines the likely gendered effects of increasing the weight given to research outputs under research assessment schemes. There are not only the important immediate consequences of being seen as research inactive; there is also the longer term potential of not being promoted.

The research assessment exercises have relied on market exchange thinking and extrinsic motivation. Gittins (2007: 21) suggests that one of the problems with focusing on monetary motivation is that individuals are tempted to assume money is the only motivation and it is an external motivation. He asserts that not all motivation is extrinsic, that “motivation is often intrinsic – we do it because we enjoy doing it; it makes us feel good about ourselves”.

A problem with focusing on extrinsic motivation is that in situations where intrinsic motivation is important, introducing monetary rewards can “crowd out” that intrinsic motivation, making things worse rather than better. A quote from a University of Hong Kong academic captures this feeling: “Research must be driven from the heart rather than from external pressures. If you are motivated, then you don’t need these external measures at all” (Currie, 2007: 23).

Academics are professionals and are generally highly intrinsically motivated. Many studies have recorded their love of their work and their identification of their work as a vocation rather than an occupation. Their willingness to give their weekends and nights to research and serving their communities is due to their inherent interest in ideas and sharing their knowledge with the public. Much of the work of academics is based on reciprocity and not being paid for many hours of work that is given in a professional capacity. Relying on payment by results undermines the generosity of academics in assisting each other in a myriad of activities like examining theses, reviewing papers, giving talks to community groups, mentoring younger academics, and participating in national associations and unions.

**PROFESSIONAL AND ETHICAL CONSEQUENCES**

There have been general concerns raised that the focus on these research assessments has led to a lower level of interest in teaching. Quoting Peter Guthrie, professor of engineering for sustainable development at Cambridge University, Attwood (2007) noted that the “level of adventure and excitement in teaching was less than it should be”. This was mainly due to the pressures on academics to
perform against the criteria of the RAE which could lead “inevitably to a lack of attention and focus on new ways of teaching and learning”. Attwood (2007) concluded that the focus on research rankings had left undergraduate engineering courses 20 years out of date and was taking the excitement and innovation out of teaching, according to Britain’s Royal Academy of Engineers.

A more significant ethical concern was raised by Birkhead (2007). He suggested that “the RAE encourages and rewards scientific misconduct as never before. It overburdens journals with too many articles, grant funding bodies with too many grants and is leading to referee fatigue for both of these”. He concluded, “I know of no academic who considers the benefits of the RAE outweigh its costs” (Birkhead, 2007: 33).

Many academics have regretted the divisive nature of these research assessment exercises because departments and individuals are forced to compete against each other with winner take all consequences. Elton (2000) was concerned that many of the consequences might be unintended and it may only be in the longer term that these become apparent. He railed against the kind of culture created—“the competitive, adversarial and punitive spirit evoked by the RAE” (Elton, 2000: 274).

Pitting one department against another is never done on a level playing field. Some areas of study lend themselves to the production of journal articles and the type of research that is more highly prized. Also some areas of study will not produce the same international impact. Sharp clearly delineated the areas that were likely to gain high points and other ones which would be left behind in this RAE competition:

Equally original work in education or law, however, is less likely to make the same international impact, because of the socially and culturally situated nature of the concepts and phenomena with which it deals. This distinction between socially-specific and socially non-specific research may have something in common with the distinction between professional and academic research. Units of Assessment (UoAs) such as Nursing, Communication, Cultural & Media Studies, Art & Design, Education and Agriculture (all of which were amongst the lowest 10 mean ratings in 2001) lie towards the professional end of the higher education curriculum and are also those in which it may be more difficult to make an international impact. (Sharp, 2004: 215–216)

In New Zealand the exercise became more expensive for some universities than for others. In fact some universities spent more on participating than they gained in funding increases (Shewan & Coats, 2006). In the second round, fewer universities/colleges participated because of the potential loss in revenue associated with the exercise. This was even apparent within discipline areas. Boston, Mischewski and Smyth (2005) calculated that the social sciences received a net loss. And when looking at those that performed poorly, it was those that were newer disciplines and more practice-based, for example, journalism, education, nursing and media studies. Should those areas that are important for the development of local knowledge be so penalised in a bid to gain world-class universities?
INSTITUTIONAL BENEFITS AND DISADVANTAGES

Despite all the negative consequences detailed above, there are some benefits and these are mainly for the research managers. The RAE sharpens the focus of research (Gordon, 2005). Paisey and Paisey (2005) remarked that there was a greater focus for research and it stimulated an increase in the number and quality of publications. Sharp (2004) stated that it concentrated their minds and made them more efficient and led to the strategic management of research. Bessant et al. (2003) concurred with this view that it led institutions to consciously focus on developing and managing a research strategy. Hare (2002) found that the RAE improved the volume and quality in terms of journal placement between 1996 and 2001. There was better research management which weeded out poor quality activities and encouraged completion of research. It created bigger research groups that were more internationally competitive. Elton (2000) concluded that research was better managed.

The 2004 PBRF report noted that rankings introduce a “powerful new incentive for universities to concentrate their research around areas of excellence” (Clarke, 2005: 196). The University of Auckland Professors (2004) concluded that it did lead to an increase in the volume and quality of research output. Boston (2004) stated that it should lead to an increase in public funding of research because measured improvement in research performance can provide strong justifications for enhanced public expenditure.

By setting up a competitive fund, the benefit was to get all individuals in these universities competing with each other to get a share of these new funds. The end result, however, has been that a small proportion of the universities get the lion’s share of the funds, such as the Russell group in the British universities and the Group of Eight in Australia. These are the ones referred to earlier by Altbach (2006) and Marginson (2006). They are the older, more comprehensive, research-intensive universities, often with a medical school. In Australia, the selectivity of universities getting the most research funding has led them to form smaller groups to lobby on behalf of their particular type of university, the newer versus the older, the technological versus the comprehensive. Yet every year the Group of Eight (the oldest, sandstone universities) gets the lion’s share of the research funding and it appears that there is ever greater concentration to a smaller number of universities. Sharp (2004) described this struggle for research funding in the United Kingdom clearly: “The RAE is ‘essentially an “old” universities’ exercise designed to give more money to the already well-off and to deny opportunities for newer institutions” (Sharp, 2004: 202). And Ball (1997) quantified this divide by noting how the RAE ranked the new universities and the old with the first 59 in the RAE list almost all old universities and those ranked 60 to 111, the new universities.

In New Zealand, in a study of the social sciences, Boston, Mischewski and Smyth (2005) found that three universities (Auckland, Massey & Otago) accounted for 54 per cent of the total staff assessed in the social sciences. Clarke (2005) commented that there was a strong link between organisational mission and the final rank
ordering of institutions with universities performing the best of all tertiary institutions in New Zealand. She also noted that the roles of different institutions are most likely to be influenced by local economic, social, cultural and historic contexts that will be difficult to change. In Hong Kong, three universities got the bulk of the funding: University of Hong Kong, Chinese University of Hong Kong and the Science and Technology University.

CONCLUSIONS

Globalisation has helped spread ideas like New Public Management, the use of performance indicators and the development of research assessment exercises. The internet, international conferences, and world organisations like the Organisation of Economic Cooperation and Development have been instrumental in connecting policy makers and imbuing them with the importance of being strategic in their allocation of research funding. The image of globalisation is often of homogenisation, marching to the same tune. In this article, another image emerges as equally plausible, one of openness. This image suggests that globalisation is not an inexorable process that sweeps all nations down the same path. There are some nations that are taking alternative paths and this is the case when it comes to research assessment exercises.

Trow (1996) felt that formal accountability was becoming a substitute for trust in countries like the United Kingdom and Australia. New Zealand and Hong Kong could also be included here. In commenting on research that he had been doing with Oliver Fulton in the United Kingdom, Trow stated: “Ironically, the more severe and detailed are accountability obligations, the less can they reveal the underlying realities for which the universities are being held accountable” (Trow, 1996: 5). He described the reports as “more public relations documents which are, shall we say, parsimonious with the truth” (Ibid: 6). Trow much preferred quality assessments based on internal reviews and personal accountability where one is held to account by one’s conscience or by one’s internalised values.

In line with Trow’s (1996) comments, some academics in Taylor’s (2001) Australian study cited self-motivation as their main rationale, as exemplified by one Australian respondent who stressed that “98 per cent of the academics I’ve ever known are driven by the self-motivation of being good academics or becoming known in their field. That drives them” (Taylor, 2001: 387).

Even in the United States where academics complain of micromanagement of their universities (Vidovich & Currie, 1998), there is no national agency that can direct the way performance indicators are applied, and current measures to rank universities are criticised as inadequate despite the many attempts to find comparable benchmarks (Brooks, 2005). According to Jongbloed and Vossensteyn (2001: 140),

Although there are many states that use performance based funding in the US, the amount of funds that is tied (directly or indirectly) to results is relatively small, and usually does not exceed 5% of the institutional
allowance. Most US states adopt performance measures primarily for accountability purposes or for informing students about higher education; they do not (yet) use them for funding decisions.

An interesting phenomenon was revealed in a large American study of 163 colleges and universities (National Center for Postsecondary Improvement, 2000), which assessed whether these institutions were becoming more accountable to their stakeholders and market forces. More public than private institutions reported increasing accountability to a larger number of stakeholders and the least prestigious institutions were subjected to the most accountability reviews. In contrast, most private, name-brand institutions and private liberal arts colleges did not report increased accountability to stakeholders. Thus, those American institutions that were the most sought after, the private prestigious institutions, were not proceeding down the quality assurance route. Yet they are recognised as having the highest market quality in terms of being able to charge the highest tuition fees and garnering the most capable students.

The type of quality programs that Continental European universities have adopted have taken the “lighter touch” of quality enhancement and not taken the “hard” accountability route (Huisman & Currie, 2004; Massy, 1999). In Sweden, Massy (1999: 11) noted that “respondent after respondent said the program’s ‘light touch’ and thoughtful demeanour made it easier to engage professors in a real dialogue about quality work”. If governments want academics to be creative and develop innovative ideas, then the quality enhancement route might be the most effective one.

REFERENCES


