**ACCESS: CONTEMPORARY ISSUES IN EDUCATION**
2008, VOL. 27, NO. 1 & 2, 24–32

ACCESS

# Research assessment exercises and some negative consequences of journal rankings and citation indices

Jan Currie

Murdoch University, Australia

**ABSTRACT**
This article focuses on the limitations of research assessment exercises, journal rankings and citation indices. It uses findings from a case study of Hong Kong's Research Assessment Exercise and secondary sources on the British Research Assessment Exercise and New Zealand's Performance Based Research Funding. It contrasts quality assurance with quality enhancement mechanisms, one based on compliance and the other on trust. It questions whether the short term benefits of gaining more publications and higher rankings can justify the potential negative consequences for the research culture, the development of younger, female and Indigenous researchers and the cost of implementing these exercises.

## Introduction

Education ministers and university policy makers in many parts of the world have decided that research assessment exercises are necessary to force their tertiary institutions 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".

Derived from this leaning toward science oriented research in journals written in the English language was a system to assess the impact of research discoveries and determine the prestige of the journals publishing new findings. Eugene Garfield, founder of The Institute for Scientific Information (ISI), began to publish Journal Citation Reports (JCR) in 1975, including the Science Citation Index (SCI) and the Social Sciences Citation Index (SSCI). The JCR provides quantitative tools for ranking, evaluating, categorizing, and comparing journals. The impact factor measures the frequency with which the "average article" in a journal has been cited in a particular year or period (Garfield, 1994).

Education ministers and national policy makers are also driven to ensure that within their national systems, they have at least one or two universities that are ranked in the top 100 and considered world-class institutions. 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, internally 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 based on trust; whereas with the quality assurance side, it is based on proving one's past performance and is more likely to engender mistrust between universities and governments implementing the system (whether local or national) and within universities between research managers and academics. Although there are positive and negative consequences in using research assessments and some of the positive benefits are mentioned in this article, my intent is to focus on the limitations of using citation metrics, journal impact factors and national systems of research assessments that are tied to funding formulas such as those used in Hong Kong, New Zealand and the United Kingdom (UK).

## The study

The impetus for this article comes from an Australian Research Council study[1] that investigated university responses to globalisation in three Asian regions. Here the findings of the Hong Kong case study on the research assessment exercises that were carried out in the 1990s are used to demonstrate the kind of responses that are felt inside universities about these assessment processes. In addition it describes the benefits and consequences of the UK and New Zealand research exercises and critiques of journal citation indices used to measure the impact of research. These critiques are based on secondary sources drawn from research studies and commentaries in these countries.

In 2003–2004, I gathered the data in Hong Kong in two stages: first, a document analysis at the University Grants Committee (UGC) and individual university levels; and second, semi-structured interviews (39 participants) in two universities, a traditional one and a technological one. From this research, it was evident that there were benefits as well as unintended consequences from Hong Kong's Research Assessment Exercises.

The academic respondents came from a range of discipline areas: sciences, social sciences, law, humanities, education, economics/business and social work. There were seventeen professors, ten associate professors, seven assistant professors (or equivalent) and five administrators (from research and registrar's offices). There were thirty-two men and seven women. It was quite an international group with many either born overseas (fourteen) or worked/educated overseas (sixteen) although more than half were born in Hong Kong or China. The interviews lasted between 30 and 60 minutes and were obtained using a snow-balling technique.

The research assessment systems that are critiqued in this article tend to be on the quality assurance side of the continuum. 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. Table 1 gives an overview of the processes. The British developed the Research Assessment Exercise (RAE) in the early 1980s and after six assessments decided to abolish it in 2008. Hong Kong adopted the British RAE in 1993 and used it for four exercises and decided to increase the number of years between each exercise. New Zealand began its Performance-Based Research Fund assessment process in 2003 and had another one in 2006. These assessment exercises are directly tied to funding formulas. In contrast, in North America and Europe there are as yet no national assessments of research that are tied to funding formulas. There are national assessments in some European countries but these are mainly used for internal improvement of universities rather than the distribution of funding (Jongbloed & Vossensteyn, 2001). So they fall on the quality enhancement side of the continuum.

## Views on the overall research assessments

The following two British quotes are examples of the cynicism with which academics greet these assessment exercises. "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). Ball (1997) demonstrated that most of the RAE funding divides the new universities from the old, with the first 59 almost all old universities, and those listed from 60 to 111 the new universities.

However, there are also those, especially the research managers, who see these as opportunities to improve their university's research productivity. The following comments come from articles describing the RAE in the UK. The RAE sharpens the focus of research (Gordon, 2004). Paisey & 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.

Similar benefits were expressed in New Zealand with its introduction of the PBRF. It has been described as a "powerful new incentive for universities to concentrate their research around areas of excellence" (Clarke, 2005). 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. This was also one of the reasons that the former Australian Minister of Education, Julie Bishop, presented to a colloquium on why the Liberal government wanted to introduce the Research Quality Framework (RQF) in 2007. She stated that the RQF would help convince the Australian Cabinet to provide more research funding and identify where research strengths were in universities so that businesses could link with researchers who were at the cutting edge (Bishop, 2007). The RQF has since been dropped by the incoming Labor government because of the concerns expressed about the impact factor. Nevertheless, a metrics system appears to be the choice of Innovation, Industry, Science and Research Minister Kim Carr. He promised a "new, streamlined, internationally recognised research quality assurance process, using metrics or other agreed quality measures appropriate to each research discipline" (Lane, 2008: 21). The British government has also decided to switch to metrics instead of its current RAE format which relies on peer reviews that are much more expensive. The 2008 RAE will be its last one. However, the switch to metrics will also not be without its detractors. There is still much dispute about how these metrics will be formulated and many complaints about the rankings of journals. Minister Kim Carr acknowledged the debate and stated, "Most of the metrics in the natural sciences is agreed. When it comes to the social sciences, including economics and the humanities and the performing arts, there's a great deal more controversy" (Lane, 2008: 21).

In a study by Coryn, Hattie, Scriven and Hartmann (2007) that evaluated sixteen national models and mechanisms used to evaluate research, the New Zealand PBRF gained the highest rating because of its comprehensiveness, its transparency, its unit of analysis based on individuals rather than institutions and its mixed approach to the evaluation. A more interesting result from this study was that more than two thirds of the models were assigned a quality category of F out of a six letter rating. They were seen has having an absence of merit, clearly inadequate and with fatal deficiencies. They concluded that evaluations that do not work are sometimes worse than none at all because of the substantial costs in time, money and expertise and the sometimes demoralising aspect for the researchers who are obliged to participate.

In every country there are, of course, mixed views as well. Several respondents in the Hong Kong study highlighted both positive and negative views about the RAE. For example,

> I think you get a lot of better than average publications. But you are less likely to get very innovative, very high impact research out because you have to meet the quota. It is a very, very mechanical way of ranking people. (CU109)

In the UK, Baty (2007) remarked that the RAE distracts academics from pursuing groundbreaking ideas in favour of low-risk options. It increases the degree of managerial control through its allocation of research funds but at the same time undermines academic values of autonomy and freedom (Hare, 2002). In New Zealand, Dianne McCarthy, of The University of Auckland said, "It has been good for the quality and quantity of research at the big research universities. But the impact has been more questionable for teaching and community service" (Illing, 2006: 25).

Many academics interviewed for the Hong Kong study focused on the new emphasis of getting published in international journals and gaining a high citation index. This next section begins with some general criticisms of the use of citations indices and some of their unintended consequences for different fields of study and for the research community in general. It ends with several quotes from Hong Kong academics sounding cautious notes about how the emphasis on international journals will be to the detriment of local and regional journals and journals in languages other than English and will not benefit scholarship in the long run.

## Critique of citations indices

A number of writers have questioned the scientific validity of the citation indices and do not feel that they are an adequate substitution for peer reviews of journal articles. Williams (1998) notes that the impact factor is determined by an arbitrary mathematical exercise that is unrelated to the scientific quality of individual papers. He warns that the impact factor should not be used without careful attention to the many phenomena that influence citation rates. He identifies that the inclusion of review articles and letters can manipulate the rankings because review articles are cited more frequently than typical research articles. Journal citation counts do not distinguish between letters, reviews or original research.

There are also differences between fields in what outputs are important. Williams and Van Dyke (2007) note that the research output in the sciences is largely in the form of refereed journal articles; in the humanities, books; and refereed conference papers in engineering and IT. They also note that in science and medicine, the lags between research findings and citations are short so that the time period used for citations can be short. However, in the social sciences, citation counts over periods of less than ten years are of limited usefulness. They cite a World Bank (2006) report suggesting that citation counts might be useful for assessing the long-term impact of an individual scholar's or department's work but are much less useful in the short run (Williams & Dyke, 2007: 3). Oswald (2007) also warned against the excessive use of journal quality as an indicator of paper quality, especially in the long run. He traced through citations to articles in 1981 in six economic journal of varying status. He found that the less highly cited articles in the top journals were easily bettered by good articles in less prestigious journals.

Despite these warnings, citations are increasingly touted as a way of judging the quality of publications. However, Dale and Goldfinch (2005) noted that citations have limitations as a direct measure of research quality. Coryn (2006) identified seventeen practical and technical concerns with citation analysis. One of these is that work that is ahead of its time gets few citations. At the same time, Dale and Goldfinch (2005) as well as other scholars (Phelan, 1999; Smith & Eysenck, 2002) assert that citation indices are highly correlated with other measures of research performance, such as the British RAE, Nobel prizes, awards and fellowships. Geary, Marriott and Rowlinson (2004) analysed the journal rankings in business and management that were revealed as part of the 2001 RAE in the UK and discovered that 80% of the 9,942 publications submitted were journal articles. In addition, 50%

of all citations were to just 126 journals representing a core list of business and management journals. They warned that the editorial policies of certain high count-high rated journals may have exercised considerable influence on the direction of management research.

It is also apparent that members of committees cannot read all the articles they are given for peer assessment. So they tend to resort to secondary criteria like citation indices, journal prestige, the reputation of authors and institutions. Williams (1998) suggests that this makes peer review as much of a lottery as of a rational process.

Bond (2007), writing in *The Australian*, stated that the reputation of the journal in which you publish your work is no measure of research ranking. He cited a number of articles demonstrating that equating research quality with journal of publication was unreliable. He suggested an alternative method of counting the number of publications, citations, citations for each article and the Hirsch index of each researcher. The Hirsch index was devised by Jorge Hirsch in 2005 as a good way to sort the excellent from the very good, and so on. Developing these measures into a metric appears to be the method that is currently favoured to replace the Research Assessment Exercise (RAE) in the UK.

A number of writers indicate that people are manipulating citations, including publishers and academics. Journal editors are running more review articles in order to gain more citations and there are citation clubs of authors who conspire to cross-reference one another (Lane, 2006). It has been noted repeatedly that there are international and North American biases in citation indices. Beyond these, Seglen (1997) found that there were a number of other biases: an English language bias; domination by American publications; research fields with literature that rapidly becomes obsolete are favoured; small research fields tend to lack journals with high impact; and long articles collect many citations and have high impact. For example, he noted that American scientists were prone to cite each other and dominated these databases. A clear case was within clinical medicine where 83% of references in the same year were to other papers by American scientists. There are also differences within fields; for example basic medicine is cited three to five times more than clinical medicine and this is reflected in journal impact factors. Donovan (2006) pointed out the injustices of using science-friendly metrics to evaluate the humanities, arts and social sciences. She found that number-crunching papers had the widest potential for citation. Seglen (1997) asserted that citation impact is primarily a measure of scientific utility rather than of scientific quality. He concluded that there seems to be no alternative to qualified experts reading the publications. Donovan concurred with that view and went further in her work on designing the Australian Research Quality Framework by recommending that different fields need to have different types of peer assessments.

Steele (2007) has demonstrated that there has been an increasing monopoly over citation indices by a limited number of publishers. "In 2006, 20 publishers accounted for 84% of revenues of the $US11 billion publishing market in science, technology and medicine. The top five STM publishers account for 50% of the market." More researchers are chasing high-impact journals for research assessment and league table purposes and this has led to increasing rejection rates. In 2006 the *British Medical Journal* accepted only 7% of 7000 submissions; the *Journal of the American Medical Association* and the *New England Journal of Medicine* published 6%; *Science* accepted 8% of 12,000 submissions while *Nature* in 2005 published 2000 of 25,000 papers received. The *Economic Journal* had a 91.5% rejection rate in 2006.

These concerns are echoed in the Hong Kong data. Hong Kong academics were critical of the issue of importing policy innovations from the West, as shown by the following comments made by three Hong-Kong-born professors of social sciences on its Western bias and undermining of scholarship per se:

> We are concerned that the evaluation criteria may lead to local and regional journals being further neglected. At the end of the day you need to nourish local and regional journals. I think that is really against the spirit of enhancing research scholarship because you want to encourage scholarship that is relevant to the community you are living and working in, particularly for the

> social sciences. Yet this RAE exercise in the eyes of our colleagues is too much of a paper exercise, more for administrative purposes, rather than really enhancing scholarship. (CityU104)
>
> I'm an old-fashioned academic and I don't really believe in this initiative. I believe that academics should be given the freedom to do research and one should be examined very holistically rather than all the time being asked how many pieces have you published this year. I also think having this list of journals is a self-defeating game because everybody is fighting to get their stuff published in that short list of journals and the capacity of absorption is limited. (CityU100)
>
> It seems to have a negative impact on scholarship per se, encouraging people to publish articles that are so-so, leading to a glut of publications, many with little substance or originality. There is no Humanities Index. These measures benefit the scientists more than the social scientists and those in humanities are particularly disadvantaged by this system. Humanities subjects are often culture bound and area specific whereas the sciences have no boundaries. (CityU105)

The focus on particular journals, the emphasis on quantity of publications, and a tendency to benefit scientists more than others were all offered as criticism by a range of academics from the two Hong Kong universities, as shown below:

> You basically exclude most of the Chinese language journals because most of them are not listed in the citation index. You will also probably avoid contributing book chapters because book chapters usually aren't counted that much. (CityU115)
>
> The way that research is assessed is very science-orientated and journal rather than monograph-orientated. Therefore the arts faculty feels that the numerical grade may not reflect their true strength. (HKU109)
>
> The rule of the game is that one will get more recognition with more publications, which is quantity before quality. Another problem is that arts research is more often than not individual research that takes longer to do. Medical and science professors work in teams and their research publications bear a long list of authors. Arts colleagues end up with fewer publications and have been ridiculed for not being as productive and for not doing collaborative research. (HKU107)

## Discussion

Research assessment exercises create winners and losers within a research community. There is considerable evidence to indicate which areas of research are rewarded and which lose out on this additional funding. Sharp identified quite clearly those that miss out in the UK:

> This differential distribution of funding was evident in New Zealand as well:
>
> Within the social sciences, newer disciplinary areas (communications, journalism, media studies) and those with a higher concentration of practice-based researchers (education, nursing) performed less well than long-established disciplines (anthropology and psychology) and those with proportionately fewer practitioners. (Boston, Mischewski & Smyth, 2005: 64)

A similar disparity is created for individual academics within a university; there are those who get the A grades and those who are considered research inactive. In New Zealand, individual academics were rated as opposed to rating the Units of Assessment (UoAs) or departments. This can be quite devastating for early career researchers who often received R (0) ratings, which were given to 40 per cent of PBRF-eligible staff in the 2003 round.

> A stigma gets attached to the R rating which impacts on new and emerging researchers where it was difficult to produce enough evidence for the research environment and peer esteem scores with potentially demoralizing effects. (Clarke, 2005: 187–188)

Despite clarifying that the R label does not mean that the academic was research inactive, "the evaluation evidence is that it became almost universally understood as meaning 'research inactive'" (WEB Research, 2004: 29).

A group of twenty-seven University of Auckland professors expressed concern about the plight of those new to the academy after the first research assessment exercise in New Zealand. They also identified the difficulties faced by particular academics and those located in certain professional areas.

> There is demoralization of new researchers. There is a struggle to identify research paths, especially for Maori staff and those whose research serves professional, social, cultural and governmental communities in NZ. (University of Auckland Professors, 2004: 5–6).

In New Zealand's PBRF, there was greater emphasis on sole authorship which can have the long term consequence of discouraging collaboration (Boston, 2004). It is only after each exercise is completed that some of these patterns emerge. As a result, new guidelines are usually written for the next exercise in an attempt to correct biases discovered in previous exercises.

However, it is difficult to get the process right and free from bias against particular groups within the academy. Writing in *The Australian* about the British experience of research funding reform, Birkhead (2007: 33) concluded, "I know of no academic who considers that the benefits of the RAE outweigh its costs". 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" (Birkhead, 2007: 33).

## Conclusion

Until recently, research assessments were carried out largely by the academic community for the academic community as a self-regulating mechanism in most parts of the world. Orr observed that with the current evolution of research assessment practices, they are now instigated by the state with other ultimate goals in mind (2004: 346). Is the aim towards "wealth creation" at the expense of "knowledge creation" (University of Auckland Professors, 2004)?

As governments have stepped into the management of research outcomes, has the emphasis on short-term products become one of their goals because of the nature of government terms? There has certainly been a greater focus on productivity and generating income with research products. In the end, government ministers and policy makers may need to explore to what degree their current forms of corporate accountability are a threat to research cultures that could nourish creativity and be freer to explore ideas. In their rush to get their universities on to the global league tables, governments may damage the prospects of younger, female and marginal groups within their universities to contribute their talents and ideas on how to create a better society.

As pointed out at the beginning of this article, research assessments seem to fall on a continuum from quality enhancement to quality assurance, the one based on trust and the other on compliance. The three research assessment exercises reviewed in this article fall on the quality assurance side. The British model is the one that has influenced Hong Kong's and New Zealand's. In contrast to these models where funding is attached to results, the European and North American forms of quality assessments have not tended to tie funds to the results. This may be the crucial element which tilts an assessment towards mistrust. It is fitting to end with this quote from a professor of social sciences because it sums up a number of concerns that surfaced in the Hong Kong study and the UK and New Zealand commentaries. It also reflects the underlying differences between the types of research assessments in different countries:

> There's no trust, there's no discretion and there's no willingness to think outside the box, so you get a very defensive structure. Whereas, in the global academic system, the successful places have people that think outside the box. They take risks. And you trust people's professional ability, you trust their dedication and basically if someone says, "I'll take care of that." You trust them to do it properly. If you don't have that trust, you end up with a system that is policed from top to bottom, academically and administratively. We have become far less collegial. With cutbacks you start to have people looking over their shoulders and you get this kind of policing mentality. (HKU100)

## Note

1. Australian Research Council Project, 2003–2007, University Responses to Globalisation in Singapore, Hong Kong, and the People's Republic of China: Focus on accountability and autonomy, by J. Currie, L. Vidovich and R. Yang.

## References


Altbach, P. (2006). The Dilemmas of Ranking. *International Higher Education*, 42. Retrieved January 15, 2008, from http://www.bc.edu/bc_org/avp/soe/cihe/newsletter/Number42/p2_Altbach.htm

Ball, D. F. (1997). Quality Measurement as a Basis for Resource Allocation: Research assessment exercises in United Kingdom universities. *R&D Management 27*(3), 281–289.

Baty, P. (2007, Friday, June 1). Research Wrecked by 'Bean Counters'. *The Times Higher Education Supplement*. Retrieved January 21, 2008, from http://timeshighereducation.co.uk/story

Bessant, J., Birley, S., Cooper, C., Dawson, S., & Gennard, J. (2003). The State of the Field in UK Management Research: Reflections of the Research Assessment Exercise (RAE) panel. *British Journal of Management, 14*, 51–68.

Birkhead, T. (2007, Wednesday, June 6). Beware the Grant-Getting Lottery. *The Australian*, p. 33.

Bishop, J. (2007). Speech addressing the Symposium held at the Institute of Advanced Study on Rankings and Research Assessments: Their Impact on Asian and Australian Universities, University of Western Australia, February 15.

Bond, N. (2007, Wednesday, January 10). Not All Views are Equal: Peer judgments need relativity to be valid. *The Australian*, Letter to the Editor, p. 30.

Boston, J. (2004). The Future of the Performance-Based Research Fund: Issues and options. Paper prepared for Royal Society Forum: *Evaluating the Assessment Framework*, 21 May.

Boston, J., Mischewski, B., & Smyth, R. (2005). Performance-Based Research Fund: Implications for research in the social sciences and social policy. *Social Policy Journal of New Zealand, 24*, 55–84.

Coryn, C. L. S. (2006). The Use and Abuse of Citations as Indicators of Research Quality. *Journal of Multi Disciplinary Evaluation, 3*(4), 115–120.

Coryn, C. L. S., Hattie, J. W., Scriven, M., & Hartmann, D. J. (2007). Models and Mechanisms for +Evaluating Government-Funded Research: An international comparison. *American Journal of Evaluation, 28*(4), 437–457.

Clarke, M. (2005). Quality Assessment Lessons from Australia and New Zealand. *Higher Education in Europe, 30*(2), 183–197.

Dale, T., & Goldfinch, S. (2005). Article Citation Rates and Productivity of Australasian Political Science Units 1995–2002. *Australian Journal of Political Science, 40*(3), 425–434.

Donovan, C. (2006, August 11). An Instrument too Blunt to Judge Sharp Minds. *The Times Higher Education Supplement*. Retrieved January 21, 2008, from http://timeshighereducation.co.uk/story

Elton, L. (2000). The UK Research Assessment Exercise: Unintended consequences. *Higher Education Quarterly, 54*(3), 274–283.

Garfield, E. (1994). The ISI Impact Factor, *Current Contents* (June 20).

Geary, J., Marriott, L., & Rowlinson, M. (2004). Journal Rankings in Business and Management and the 2001 Research Assessment Exercise in the UK. *British Journal of Management, 15*, 95–141.

Gordon, D. (2005). UK's Research Assessment Exercise (correspondence). *The Lancet*, *365*(9464), 1025.

Hare, P. G. (2002). The UK's Research Assessment Exercise: Impact on institutions, departments, individuals. Paper prepared for OECD Conference on Higher Education, September, Paris, France.

Hirsch, J. E. (2005). An Index to Quantify an Individual's Scientific Research Output. Retrieved January 28, 2008, from http://www.arxiv.org/abs/physics/0508025

Illing, D. (2006, Wednesday, November 15). Stars of Research Lured by Top Dollar. *The Australian*, p. 25.

Jongbloed, B., & Vossensteyn, H. (2001). Keeping up Performances: An international survey of performance-based funding in higher education. *Journal of Higher Education Policy and Management, 23*(2), 127–145.

Lane, B. (2006, Wednesday, August 9). Over-citation puts integrity under cloud. *The Australian*, p. 33.

Lane, B. (2008, Wednesday, January 23). Research review heats up. *The Australian*, p. 21.

Marginson, S. (2006, Wednesday, December 6). Rankings ripe for misleading. *The Australian*, pp. 26–27.

Orr, D. (2004). Research Assessment as an Instrument for Steering Higher Education: A comparative study. *Journal of Higher Education Policy and Management*, *26*(3), 345–362.

Oswald, A. J. (2007). An Examination of the Reliability of Prestigious Scholarly Journals: Evidence and implications for decision-makers. *Economica, 74*(292), 21–31.

Paisey, C., & Paisey, N. J. (2005). The Research Assessment Exercise 2001: Insights and implications for accounting education research in the UK. *Accounting Education: An international journal*, *14*(4), 411–426.

Phelan, T. J. (1999). A Compendium of Issues for Citation Analysis. *Scientometrics 45*(1), 117–136.

Seglen, P. O. (1997). Why the Impact Factor of Journals Should Not be Used for Evaluating Research. *British Medical Journal (BMJ), 314*(15), 497–508.

Sharp, S. (2004). The Research Assessment Exercises 1992–2001: Patterns across time and subjects. *Studies in Higher Education 29*(2), 201–218.

Smith, A., & Eysenck, M. (2002). The Correlation Between RAE Ratings and Citation Counts in Psychology. Retrieved October 1, 2007, from http://psyserver.pc.rhbnc.ac.uk/citations.pdf

Steele, C. (2007, Wednesday, October 31). Alternatives to Bonuses (Letter to the Editor). *The Australian*, p. 33.

University of Auckland Professors. (2004). *Submission on the University of Auckland Draft Strategic Plan 2005–2012*. Auckland, New Zealand: University of Auckland (Group of 27 professors).

WEB Research. (2004). *Evaluation of the Implementation of the PBRF and the Conduct of the 2003 Quality Evaluation*. Wellington, New Zealand: Centre for Research on Work, Education and Business Limited.

Williams, G. (1998). Misleading, Unscientific, and Unjust: The United Kingdom's research assessment exercise. *British Medical Journal, 316*(4), 1079–1082.

Williams, R., & Van Dyke, N. (2007). Measuring University Performance at the Discipline/Departmental Level. Paper presented at *Symposium on International Trends in University Rankings and Classifications,* 12 February, Griffith University.

World Bank. (2006). *An Evaluation of World Bank Research 1998–2005*. Retrieved January 20, 2008, from www.tinyurl.com/yck7wa