A ‘perverse incentive’ from bibliometrics: could National Research Assessment Exercises (NRAEs) restrict literature availability for nature conservation?

Michael C. Calver[1], Maggie Lilith[1], Christopher R. Dickman[2]

1.School of Biological Sciences and Biotechnology, Murdoch University, Murdoch, Australia

2.Institute of Wildlife Research, School of Biological Sciences, University of Sydney, Sydney, Australia

Abstract

National Research Assessment Exercises (NRAEs) aim to improve returns from public funding of research. Critics argue that they undervalue publications influencing practice, not citations, implying that journals valued least by NRAEs are disproportionately useful to practitioners. Conservation biology can evaluate this criticism because it uses species recovery plans, which are practitioner-authored blueprints for recovering threatened species. The literature cited in them indicates what is important to practitioners’ work. We profiled journals cited in 50 randomly selected recovery plans from each of the USA, Australia and New Zealand, using ranking criteria from the Australian Research Council and the SCImago Institute. Citations showed no consistent pattern. Sometimes higher ranked publications were represented more frequently, sometimes lower ranked publications. Recovery plans in all countries also contained 37 % or more citations to ‘grey literature’, discounted in NRAEs. If NRAEs discourage peer-reviewed publication at any level they could exacerbate the trend not to publish information useful for applied conservation, possibly harming conservation

efforts. While indicating the potential for an impact does not establish that it occurs, it does suggest preventive steps. NRAEs considering the *proportion* of papers in top journals may discourage publication in lower-ranked journals, because one way to increase the *proportion* of outputs in top journals is by not publishing in lower ones. Instead, perhaps only a user-nominated subset of publications could be evaluated, a department's or an individual's share of the top publications in a field could be noted, or innovative new multivariate assessments of research productivity applied, including social impact.

Keywords: Australia; New Zealand; USA; UK; Nature conservation; Threatening process; ERA; PBRF; VTR; REF; NRAE; RAE

Introduction

In many parts of the world governments are implementing policies to evaluate and improve the quality of research that they support, especially in the tertiary sector. Northcott and Linacre (2010) proposed the acronym National Research Assessment Exercises (NRAEs) for such government-driven assessments of university research (as distinct from research evaluations in non-university government research laboratories, which we do not consider). OECD (2010) recognises a sub-category of NRAE, the performance-based research funding schemes (PBRF), where research outputs are evaluated retrospectively and funding is allocated based on the evaluations.

Many assessment schemes now exist. In 2010, for example, the Australian Commonwealth implemented the first iteration of its 'Excellence in Research for Australia' (ERA) initiative, a major evaluation of the research conducted by Australia's universities (Cooper and Poletti 2011). Other examples include the Valutazione Triennale Della Ricerca (VTR) in Italy (Corsi et al. 2010), the Performance Based Review Fund (PBRF) in New Zealand (Shewan and Coats 2006; Edgar and Geare 2010; Sampson and Comer 2010), and the Research Assessment Exercise (RAE) and its successor the Research Excellence Framework (REF) in the UK (Oppenheim 2008; Hicks 2009). In addition to these countries, Northcott and Linacre (2010) claim that France, Hong Kong, Spain and Sweden either

have or will soon implement NRAEs. The Netherlands (Broadbent 2010), South Africa (Visser 2009) and Norway (Gihus and Sivertsen 2009) can be added to this list. The USA does not have a government NRAE, but Hicks (2009) argued that 'free-lance' (i.e., non-government) assessments such as those of the National Research Council (NRC) are influential and approximately equivalent.

National Research Assessment Exercises generate intense controversy. Arguments in favor are that they ensure that universities spend large amounts of public funding wisely (Oswald 2010), raise professional standards in the sector (Butler and McAllister 2009), increase research outputs such as attendance at major international conferences and publications in top journals (Butler and McAllister 2009; Box 2010; Hodder and Hodder 2010), and lead to 'financial and reputational gains' for good institutions and good researchers (Broadbent 2010, p.15). Criticisms are that NRAEs lead to an emphasis on good scores rather than good science (Butler 2007), encourage short-term research on 'hot' or 'bandwagon' topics that result in a narrowing of research diversity (Box 2010; Corsi et al. 2010; Northcott and Linacre 2010), divert significant funds from supporting research directly (Hicks 2009; Broadbent 2010), change editorial policies to favor review papers, theme issues and 'hot' topics (Steele et al. 2006; Falagas and Alexiou 2008), are retrospective whereas research is forward-looking (Elton 2000), discourage interdisciplinary research or the development of new disciplines (Elton 2000; Marsh et al. 2012) and fail to give proper recognition to lower-ranked journals publishing clinical or applied research that influence professional practice rather than academic citations (Shewan and Coats 2006). In sum, these changes could reduce publications on applied topics of local significance, although Butler (2010) pointed out that the critics have amassed little rigorous data to support their position. There is therefore a clear need to attempt data-based assessments of the criticisms.

Conservation biology provides a unique opportunity to evaluate one contentious area—the possible impact of changes in academic publishing practices on the work of practitioners. This is because conservation biology uses species recovery plans, which are practitioner-authored '… information and management documents that set out actions to assist threatened taxa and ecological communities within a planned and logical framework' (Horwitz and Wardell-Johnson 2009, p. 574). Although

neither perfect nor without their critics (Clark et al. 2002), these plans are often mandated in legislation and are regarded as important factors in recovering endangered species and reducing their risk of extinction (Ortega-Argueta et al. 2011). They can be used to determine the literature used by conservation biology practitioners and, on this basis, predict possible consequences of any changes in academic publishing practices on practitioners' work. We hypothesized that if lower-status academic publications are of disproportionate importance to the practitioners writing and implementing recovery plans, as predicted by some critics of NRAEs, then they should be cited disproportionately highly in the reference lists of those plans.

We tested this hypothesis by profiling the literature cited in random samples of 50 species recovery plans from each of Australia, New Zealand and the USA as examples of countries where recovery plans are used widely. The profiles reveal the relative extent to which species recovery plans, as indicators of the research applied in practical conservation actions, use the types of research publication likely to be regarded highly by NRAEs. They also indicate whether different categories of literature are used in different proportions in the recovery plans of the different countries. The likely impact on conservation planning of any diminution in the availability of different types of literature can then be assessed.

Methods

Species recovery plans for Australia, New Zealand and the USA are available online at http://www.environment.gov.au/biodiversity/threatened/recovery-list-common.html, http://www.doc.govt.nz/publications/science-and-technical/products/series/threatened-species-recovery-plans/ and http://www.fws.gov/endangered/species/recovery-plans.html, respectively. In July and August 2010, we selected at random 50 recovery plans from each country. These represented 8.1 % of all Australian recovery plans available, 80.6 % of all New Zealand recovery plans and 1.7 % of all US recovery plans. We first classified the references cited in the plans into journal articles, conference proceedings, books, book chapters and 'grey literature' (reports,

theses, websites, and the like that are ephemeral or unavailable or difficult to access through conventional library or publisher sources—see Debachere 1995; Calver and King 2000). We then classified the journal literature specifically according to the ranking system used by the Australian Research Council (ARC) in its 2010 'Excellence in Research for Australia' (ERA) evaluation (Cooper and Poletti 2011) and 2010 data from SCImago Lab (SCImago 2007).

The ARC system classified journals as A+ (top 5 % in their field), A (the next 15 %), B (the next 30 %) and C (the remaining 50 %). The full database can be downloaded from http://www.arc.gov.au/era/era_2010/archive/era_journal_list.htm#1. Where a journal title was not listed in the ARC database, we searched for it online to check if the journal had changed its name. If so, we assigned the ranking for the new name. Where a journal was unlisted and we found no evidence of a name change, we used a ranking of 'C?' on the assumption that an unlisted journal was unlikely to be scored at a higher level than C if evaluated. Although the ARC ranking system will be replaced in future years by a journal indicator profile (ARC 2012), we used it as an easily understood measure that reflects the ranking aspects of NRAEs.

To our knowledge the SCImago Lab ranking is not affiliated with any NRAE, but its classification of journals according to citation impact in different subject areas accords with the general philosophy of quality assessment in NRAEs. It places journals in specified subject areas within a quartile range, where Q1 represents the top quartile and Q4 the lowest. Rankings are based on the SCImago Journal Rank (SJR) (Colledge et al. 2010; Jacsó 2010). SCImago Journal Rank uses an algorithm similar to the 'page rank' employed by the Google search engine to prioritize internet sites corresponding to a search term. SJR assesses the quality of the citing sources (citations from highly cited sources carry more weight than those from infrequently cited sources), partially excludes self-citations and uses a 3-year time frame to evaluate a journal in a given year. Jacsó (2010) considered these points, plus the basing of SJR on the broad Scopus database, to be significant advantages. Where journals were ranked in more than one subject area, we used the median quartile in our evaluation. Scoring the four quartiles as 1, 2, 3 and 4 (with values of 1.5, 2.5 and 3.5 in some cases where a median quartile was calculated) led to an eight point ranking scale, with smaller numbers more highly ranked: 1, 1.5, 2,

2.5, 3, 3.5, 4 and `4?' (journal not listed in SCImago, with no evidence found that it had changed its name to a listed title).

The ideal test of the hypothesis that species recovery plans cite lower ranked journals more than might be expected would be to compare the incidence of citations of journals in recovery plans to the proportions predicted by the ARC and SCImago journal categories. For the ARC, the expected values would be 5, 15, 30 and 50 % in each of the categories A+, A, B and C and C? combined, respectively, while for SCImago it would be 25 % in each of the four quartiles (combining 1 and 1.5 together, 2 and 2.5 together and so on). However, this will only hold true if journals in the different categories publish approximately similar numbers of papers. We tested this for the ARC and SCImago classifications separately. First, we selected randomly five journals from each country in each of the four ARC categories and then five journals from each country in each of the four SCImago categories. We then determined the number of papers published by each journal in 2010, using the Scopus database. For SCImago quartile 4 for New Zealand data, there were only two journals listed in Scopus so they were the only ones used. To determine if the number of papers published per journal was similar across the journal rankings, we used the number of papers published by each journal in 2010 as a dependent variable in two-way ANOVAs with the factors of 'country' and 'ranking category', separately for the ARC and SCImago classifications. Data were log-transformed to correct for unequal variances. Where significant differences were found, they were explored further using Tukey's HSD test for equal sample sizes (Australian and US data) and Tukey's HSD test for unequal sample sizes (New Zealand data).

We decided that if ANOVAs revealed that the number of papers published per journal was similar across the rankings, we could legitimately compare the frequency of citations in the recovery plans against the expected values given above. If not, we would focus on assessing associations between citations to journals in different categories and country, and on a general description of the patterns of citations.

Results

The numbers of papers published per journal in the four ARC categories were unequal ($F_{3,48} = 15.6$, $p < 0.001$), but there were no differences between countries ($F_{2,48} = 0.6$, $p = 0.56$) and there was no interaction between countries and journal categories ($F_{6,48} = 0.9$, $p = 0.48$). The A+ and A journals published significantly more papers than the B and C journals (Table 1). There was a similar result for the SCImago classifications. The numbers of papers published per journal across the four quartiles were unequal ($F_{3,45} = 13.2$, $p < 0.001$), but there were no differences between countries ($F_{2,45} = 0.8$, $p = 0.46$) and there was no interaction between countries and journal categories ($F_{6,45} = 0.8$, $p = 0.59$). Quartile 1 journals published significantly more papers than the journals in the other three quartiles (Table 1). Based on these results, we decided against comparing the frequency of citations for the different journal categories against the notional proportions of journals in each category because of the bias created by the unequal number of papers published by journals in different categories. Instead, we describe the patterns of citations.

There were striking qualitative similarities in the data for all countries (Table 2). The grey literature cited was substantial (between 37.3 and 42.6 % for each country), and similar to that for journal citations (33.7–40.3 %). Citations to A + and A journals under the ARC ranking system were between 11.1 and 24.6 % of all journals and 4.5 and 9.2 % of all references. Citations of SCImago journal categories 1 and 1.5 were higher (27.3–42.0 % of all journals), but representing between 10.2 and 15.7 % of all references.

The three countries differed in the proportions of references that were cited in plans in the categories journals, books/book chapters, conference proceedings and grey literature, with USA plans using more grey literature and less journal literature than those from Australia or New Zealand ($\chi 26\chi 62 = 41.8$, $p < 0.001$). With regard to journals specifically, ARC journal rankings were associated strongly with country ($\chi 28\chi 82 = 230.7$, $p < 0.001$); recovery plans in the USA used proportionately more references cited as ‘A and C?’ than did those from Australia or New Zealand. SCImago journal rankings also were associated strongly with country ($\chi 214\chi 142 = 269.8$, $p < 0.001$),

with Australian recovery plans using proportionately more references in the first two categories than the other countries (Table 2).

Discussion

Characteristics of sources cited in species recovery plans

The profiles of references cited in species recovery plans suggest that research in all classes of journals and in books, conference proceedings and the grey literature is important for nature conservation. Recovery plans from Australia, New Zealand and the USA all made extensive use of lower ranked journals and grey literature in their quest for important background information, a phenomenon also noted by Stergiou and Tsikliras (2006) in a survey of the literature cited in four reviews of fish biology in the Mediterranean. These qualitative similarities overshadow the statistically significant differences in the distribution of references across publication categories in different countries and support the conclusion that all categories of literature are important in recovery planning.

While Stinchcombe and Moyle (2002) called for greater involvement of academic researchers in preparing species recovery plans to increase the uptake of conservation research in recovery planning, this may be a lesser problem than encouraging the formal publication of information in the grey literature that otherwise risks being overlooked or lost altogether Calver and King (2000). There is also a need for more basic ecological research to inform recovery plans (Ortega-Argueta et al. 2011) and for integrating the knowledge and experience of local community groups into research and publication (e.g., Fairfull and Williams 2003).

Formal publication of local information is unlikely in the top-ranked journals, which often emphasize studies of broad international appeal (Calver et al. 2010). For example, Meffe (2006) (then editor of *Conservation Biology*) acknowledged that some papers he described as ‘critical studies’, ‘important’ and ‘breakthrough’ lacked novelty or broad appeal and hence were unsuitable

for *Conservation Biology* (an ARC A+ and a SCImago quartile 1 journal). Similarly, when referring to *Biological Conservation* (an ARC A and a SCImago quartile 1 journal), Primack (2009) (the journal's editor-in-chief) observed '…now we reject many papers that are worthy of publication due to lack of space'. Nevertheless, if published in regional journals these studies would be a valuable conservation resource. This point is illustrated by important recent studies of long-term population trends in different groups of European insects Schuch et al. (2011, 2012a,b) that were only possible because of detailed baseline studies 40–50 years old that were published in local journals.

Implications for NRAEs

Overall, the references cited in species recovery plans show that much specific, practical information needed for regional conservation application is not published in the major international journals that attract the greatest credit under NRAEs. If, in response to the requirements of NRAEs, researchers reduce their rate of collecting and publishing of such regional, applied data, then less valuable information would be available for local conservation efforts and constitute a threat to nature conservation.

Butler (2010) cautioned against jumping to such a conclusion and accepting anecdotal or perception-based criticism of negative outcomes of NRAEs. She argued that NRAEs are but one of a wide range of influences on researchers' behaviour. Requirements for academic promotion, personal commitment and motivation, sources of funding and the like may all influence what research projects are undertaken and what papers are published. For example, McNay (1998) found that many academics in the UK believed that interdisciplinary research would suffer under the RAE. However, few staff involved in interdisciplinary research shifted their focus so the general belief was not reflected in practice. Furthermore, NRAEs are attempting to evaluate research quality, not research impact, so criticising them for not giving a strong weighting to impact is to misrepresent their key role.

In contrast, McNay (1998) observed that NRAEs change managerial policy. Therefore, the view that alternative forces to NRAEs balance pressures to concentrate publications in top outlets may be challenged at the institutional or departmental level if managers seek to gain advantage from NRAEs

by focusing on high scores rather than quality science (Elton 2000; Butler 2007; Hicks 2010; Luwel 2010; Cooper and Poletti 2011. Lawrence (2007) p. R583) complained that: 'Consequently, over the last 20 years a scientist's primary aim has been downgraded from doing science to producing papers and contriving to get them into the "best" journals'. Research topics or academic activities unlikely to lead to publications in top journals may be disadvantaged, ultimately leading to their neglect (Roa et al. 2009; Sampson and Comer 2010; Marsh et al. 2012). The result is that: 'Rather than genuinely fostering relevant knowledge, the emphasis on ranking seems to be driven by a desire to identify winners and losers in a game of academic prestige' (Adler and Harzing 2009, p. 74).

While statements such as these may be taken as examples of the perception-based or anecdotal arguments Butler (2010) cautioned against, she did acknowledge that NRAEs can change management practices within higher education institutions, that institutions do seek to take advantage of NRAEs and departmental research heads encourage publication in leading journals. This reality was noted by Australia's then Minister for Innovation, Industry, Science and Research when he announced that the ARC journal ranking system was to be replaced with a journal indicator profile: 'There is clear and consistent evidence that the rankings were being deployed inappropriately within some quarters of the sector, in ways that could produce harmful outcomes, and based on a poor understanding of the actual role of the rankings. One common example was the setting of targets for publication in A and A+ journals by institutional research managers' (Carr 2011). With regard to the new profile, ARC (2012) notes: 'The change enables journal quality to remain an indicator for ERA 2012, while discouraging the use of assessments of journal quality beyond their role as an ERA indicator. It will ensure that the indicator is kept in proper perspective, while maintaining ERA's rigour and focus on quality.' The problem was not the measure itself, but the way managers applied it.

This is an example of what Marsh et al. (2012) call a 'perverse incentive'—an undesirable consequence of individuals or organizations seeking to optimize scores under the rules. It is not alone. Here are three similar examples of abuse of ranking data quoted by Adler et al. (2008, p. 10), which they claim are only a few of many reported to them:

'***Example 1***

My university has recently introduced a new classification of journals using the science citation index core journals. The journals are divided into three groups based only on the impact factor. There are 30 journals in the top list, containing no mathematics journal. The second list contains 667, which includes 21 mathematics journals. Publication in the first list causes university support of research to triple; publication in the second list, to double. Publication in the core list awards 15 points; publication in any *Thomson Scientific* covered journal awards 10. Promotion requires a fixed minimum number of points.

Example 2

In my country, university faculty with permanent positions are evaluated every 6 years. Sequential successful evaluations are the key to all academic success. In addition to a curriculum vitae, the largest factor in evaluation concerns ranking five published papers. In recent years, these are given 3 points if they appear in journals in the top third of the *Thomson Scientific* list, 2 points if in the second third, and 1 point in the bottom third. (The three lists are created using the impact factor.)

Example 3

In our department, each faculty member is evaluated by a formula involving the number of single-author-equivalent papers, × the impact factor of the journals in which they appear. Promotions and hiring are based partly on this formula.'

Possible consequences of obsession with journal status are the demise of some local journals because of lack of copy or institutional support (Gowrishankar and Divakar 1999; Steele et al. 2006), or changes in editorial policies to aspire to international status that reduce opportunities to publish local work (Bryant and Calver 2012). Furthermore, quantitative studies demonstrating how ranking criteria can bias against particular fields of study are emerging. In comparing Innovation Studies Units against Business Management Schools in the UK using a range of journal-based and citation-based metrics, Rafols et al. (2012) concluded that the top business journals span a narrower disciplinary range than

lower-ranked journals, leading to a more favourable assessment of the of the Business Management Schools than the Innovation Studies Units, using journal ranking criteria.

Possible reforms for NRAEs

While accepting Butler's (2010) point that there is a substantial gap between demonstrating the potential for a perverse incentive and documenting conclusively that it is occurring, if a modest modification in policy prevents any chance of a problem such a change is, we believe, worthwhile. This approach is in keeping with the well-known precautionary principle, which argues that where there are plausible risks of serious environmental damage (in this case changes in publication practice) lack of certainty about the impacts should not preclude precautionary action (e.g., Deville and Harding 1997; UNESCO 2005). Simultaneously, research can be undertaken to reduce uncertainty Calver et al. 2011). In this case, we believe that the importance of a breadth of literature to conservation professionals that we have demonstrated, plus the quantitative study by Rafols et al. (2012) that shows that journal ranking approaches can bias against certain disciplines, establishes a need for precautionary action.

One important reform would be for NRAEs to reduce any emphasis on the *proportion or percentage* of an individual's, a department's or an institution's publications in top-ranked journals. For example, SCImago's Q1 indicator calculates the percentage of an institution's scientific publications in the top 25 % of journals according to SCImago's SJR indicator (SCImago 2012). One easy way for a well-published individual, department or institution to achieve a higher proportion of papers in top journals is to reduce publications in less prestigious local journals while keeping the contributions to higher ranked journals at least constant. The unfortunate result would be that research of local interest is not done or, if it is, that it remains unpublished.

Marsh et al. (2012) suggest the simple and transparent solution of assessing only the top 80 % of papers for any unit of evaluation. This would allow publishing of important regional information without penalty if it was balanced with publications in top journals. Another approach is illustrated by the SCImago (2012) Excellence Rate, indicating '… which percentage of an institution's scientific

output is included into the set formed by the 10 % of the most cited papers in their respective scientific fields.' This is uninfluenced by any lower-ranked publications and so removes any disincentive to publish in them. While such a measure might be volatile, averaging it over a period of 3–5 years could give a more stable and potentially useful indicator.

A more challenging suggestion is to tackle the acknowledged difficult area of research impact (Hicks 2010). This would allow individuals, departments or institutions undergoing evaluation to present evidence of the impact of their work beyond the standard bibliometric measures of citations and journal rankings, as recommended by Roa et al. (2009) to improve recognition for ethnically focused research. It is planned for the UK's REF (Smith et al. 2011). Their arguments are equally applicable to local or regional research aimed at conservation management or policy development. Examples include uptake of research by professional bodies or government instrumentalities (Witten and Hammond 2010), download-statistics indicating wide readership (Bollen et al. 2009), or publication in languages other than English to reach local communities or local professionals (Adler and Harzing 2009). Editors of top-ranked journals, who often reject scientifically sound papers with a regional or local focus for reasons of space, could help by suggesting that the papers be sent to specific, named regional or local journals.

Such suggestions are easy to dismiss as 'too hard', and Martin (2011) is fearful about the cost of implementing them. However, Lane (2010) and Lane and Bertuzzi (2011) argue that accurate, multivariate measures of impact are essential for a true understanding of the effects of scientific research. Lane (2010) highlighted three main projects that she believes show the way: STAR METRICS (Science and Technology for America's Reinvestment: measuring the effects of research on innovation, competitiveness, and science), which aims to document diverse achievements from US researchers receiving Federal funds; the Lattes Platform in Brazil, which combines multivariate data on Brazilian researchers from many different institutions; and MESUR (Metrics from Scholarly Usage of Resources), which notes the frequency and duration of online access to articles. In addition to these, European agencies from The Netherlands, Spain, France and the UK are collaborating in the SIAMPI (Social Impact Assessment Methods through Productive Interactions) project, which aims to

assess the social impact of research, including consideration of the effectiveness of granting bodies (Molas-Gallart and Tang 2011; Spaapen and van Drooge 2011). Such approaches would have a far lower chance of creating perverse incentives by giving credit for a very broad profile of researcher performance.

Concluding remarks

Box (2010) noted that there are few data-based evaluations of NRAEs, with the result that many claims of benefits or problems are unsupported by strong evidence (Butler 2010). We have taken a preliminary step toward using a data-based approach in evaluation, which we believe shows the potential (as distinct from establishing that it is actually happening) that NRAEs run a real risk of threatening nature conservation by discouraging applied research and publication on specific regional and local problems. Or, as Lane (2010, p. 488) expressed it, 'If we do not press harder for better metrics, we risk making poor funding decisions or sidelining good scientists.' We also suggest that some modifications to the approaches in NRAEs could reduce the risk by removing disincentives to publish work only suited to regional or local journals.

This is not an argument against publication in top journals—rather, researchers or their institutions should not be penalized for publishing material of regional interest that, as our data show, is used by practitioners in preparing relevant local conservation responses. Action on removing the emphasis on the *proportion* of work in the top journals could be done immediately, removing the perverse incentive not to publish in local journals. Although measures to assess the social impact of research will require more careful implementation, writers such as Hicks (2010) are calling for it and Lane (2010), Lane and Bertuzzi (2011) and Spaapen and van Drooge (2011) show that it can, and is, being done.

Acknowledgments

We thank, without implication, H. Recher, B. Dell, D. Saunders and an anonymous reviewer for detailed and constructive feedback on earlier versions.

References

Adler, R., Ewing, J. & Taylor, P. (2008). Citation statistics. A report from the International Mathematical Union (IMU) in cooperation with the International Council of Industrial and Applied Mathematics (ICIAM) and the Instititute of Mathematical Statistics (IMS). http://www.mathunion.org/fileadmin/IMU/Report/CitationStatistics.pdf Accessed 28 August 2012.

Adler, N. J., & Harzing, A.-W. (2009). When knowledge wins: transcending the sense and nonsense of academic rankings. *Academy of Management Learning and Education, 8*, 72–95.

ARC (2012). ERA 2012 frequently asked questions. Available from http://www.arc.gov.au/era/faq.htm. Accessed 12 August 2011.

Bollen, J., van de Sompel, H., Hagberg, A., & Chute, R. (2009). A principal component analysis of 39 scientific impact measures. *PLoS ONE, 4*(6), e6022. doi:10.1371/journal.pone.0006022.

Box, S. (2010). Performance-based funding for public research in tertiary education institutions: Country experiences. In OECD Performance-based Funding for Public Research in Tertiary Education Institutions: Workshop Proceedings. Paris: OECD Publishing. doi:10.1787/9789264094611-en.

Broadbent, J. (2010). The UK research assessment exercise: Performance measurement and resource allocation. *Australian Accounting Review, 20*, 14–23.

Bryant, K. & Calver, M. (2012). Adaptive radiation in Australian journals in the Arbustocene ERA: an empty niche for JANCO? In P.B.Banks, D. Lunney & C.R. Dickman (Eds.), *Science under siege* (in press). Sydney: Royal Zoological Society of New South Wales.

Butler, L. (2007). Assessing university research: a plea for a balanced approach. *Science and Public Policy, 34*, 565–574.

Butler, L. (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. Paris: OECD Publishing. doi:10.1787/9789264094611-en.

Butler, L., & McAllister, I. (2009). Authors' response to reviews. *Political Studies Review, 7*, 84–87.

Calver, M. C., & King, D. R. (2000). Why publication matters in conservation biology. *Pacific Conservation Biology, 6*, 2–8.

Calver, M. C., Grayson, J., Lilith, M., & Dickman, C. R. (2011). Applying the precautionary principle to the issue of impacts by pet cats on urban wildlife. *Biological Conservation, 144*, 1895–1901.

Calver, M., Wardell-Johnson, G., Bradley, S., & Taplin, R. (2010). What makes a journal international? A case study using conservation biology journals. *Scientometrics, 85*, 387–400.

Carr, K. (2011). Improvements to Excellence in Research for Australia. Canberra: Australian Government. Available from

http://archive.innovation.gov.au/ministersarchive2011/carr/MediaReleases/Pages/IMPROVEMENTSTOEXCELLENCEINRESEARCHFORAUSTRALIA.html. Accessed 9 April 2011.

Clark, J. A., Hoekstra, J. M., Boersma, P. D., & Kareiva, P. (2002). Improving US Endangered Species Act recovery plans: key findings and recommendations of the SCB recovery plan project. *Conservation Biology, 16*, 1510–1519.

Colledge, L., De Moya-Anegón, F., Guerrero-Bote, V., López-Illescas, C., El Aisati, M., & Moed, H. F. (2010). SJR and SNIP: Two new journal metrics in Elsevier's Scopus. *Serials, 23*, 215–221.

Cooper, S., & Poletti, A. (2011). The new ERA of journal ranking: the consequences of Australia's fraught encounter with 'quality'. *Australian Universities' Review, 53*, 57–65.

Corsi, M., D'Ippoliti, C., & Lucidi, F. (2010). Pluralism at risk? Heterodox economic approaches and the evaluation of economic research in Italy. *American Journal of Economics and Sociology, 69*, 1495–1529.

Debachere, M.-C. (1995). Problems in obtaining grey literature. *IFLA Journal, 21*, 94–98

Deville, A., & Harding, R. (1997). *Applying the precautionary principle*. Sydney: The Federation Press.

Edgar, F., & Geare, A. (2010). Characteristics of high- and low-performing university departments as assessed by the New Zealand performance based research funding (PBRF) exercise. *Australian Accounting Review, 20*, 55–63.

Elton, L. (2000). The UK research assessment exercise: unintended consequences. *Higher Education Quarterly, 54*, 274–283.

Fairfull, S. J., & Williams, R. J. (2003). Community involvement in natural resource management: Lessons for future water management in catchments of New South Wales. In P. Hutchings & D. Lunney (Eds.), *Conserving marine environments: Out of sight, out of mind* (pp. 55–61). Sydney: Royal Zoological Society of New South Wales.

Falagas, M. E., & Alexiou, V. G. (2008). The top-ten in journal impact factor manipulation. *Archivum Immunologiae Et Therapiae Experimentalis, 56*, 223–226.

Gihus, N. E., & Sivertsen, G. (2009). Publishing affects funding in neurology. *European Journal of Neurology, 17*, 147–151.

Gowrishankar, J., & Divakar, P. (1999). Sprucing up one's impact factor (multiple letters). *Nature,401*(6751), 321–322.

Hicks, D. (2009). Evolving regimes of multi-university research evaluation. *Higher Education, 57*, 541–552.

Hicks, D. (2010). Overviews of performance-based research funding systems. In OECD Performance-based Funding for Public Research in Tertiary Education Institutions: Workshop Proceedings. Paris: OECD Publishing. doi:10.1787/9789264094611-en.

Hodder, A. P. W., & Hodder, C. (2010). Research culture and New Zealand's performance-based research fund: Some insights from bibliographic compilations of research outputs. *Scientometrics, 84*, 1–15.

Horwitz, P. & Wardell-Johnson, G. (2009). Cultural conservation biology. In M. Calver, A. Lymbery, J. McComb & M. Bamford, M. (Eds.) *Environmental biology* (pp. 559–578). Melbourne: Cambridge University Press.

Jacsó, P. (2010). Comparison of journal impact rankings in the SCImago journal and country rank and the journal citation reports databases. *Online Information Review, 34*, 642–657.

Lane, J. (2010). Let's make science metrics more scientific. *Nature, 464*, 488–489.

Lane, J., & Bertuzzi, S. (2011). Measuring the results of science investments. *Science, 331*, 678–680.

Lawrence, P. A. (2007). The mismeasurement of science. *Current Biology, 17*, R583–R585.

Luwel, M. (2010). Highlights and reflections: rapporteur's report. In OECD Performance-based Funding for Public Research in Tertiary Education Institutions: Workshop Proceedings. Paris: OECD Publishing. doi:10.1787/9789264094611-en.

Marsh, H., Smith, B., King, M., & Evans, T. (2012). A new era for research education in Australia? *Australian Universities' Review, 54*, 83–93.

Martin, B. R. (2011). The research excellence framework and the 'impact agenda'. Are we creating a Frankenstein monster? *Research Evaluation, 20*, 247–254.

McNay, I. (1998). The research assessment exercise (RAE) and after: 'you never know how it will turn out'. *Perspectives: Policy and Practice in Higher Education, 2*, 19–22.

Meffe, G. (2006). The success—and challenges—of conservation biology. *Conservation Biology,20*, 931–933.

Molas-Gallart, J., & Tang, P. (2011). Tracing 'productive interactions' to identify social impacts: an example from the social sciences. *Research Evaluation, 20*, 219–226.

Northcott, D., & Linacre, S. (2010). Producing spaces for academic discourse: The impact of research assessment exercises and journal quality rankings. *Australian Accounting Review, 52*, 38–54.

OECD (2010). Performance-based Funding for Public Research in Tertiary Education Institutions: Workshop Proceedings. Paris: OECD Publishing. doi:10.1787/9789264094611-en.

Oppenheim, C. (2008). Out with the old and in with the new: the RAE, bibliometrics and the new REF. *Journal of Librarianship and Information Science, 40*, 147–149.

Ortega-Argueta, A., Baxter, G., & Hockings, M. (2011). Compliance of Australian threatened species recovery plans with legislative requirements. *Journal of Environmental Management, 92*, 2054–2060.

Oswald, A. J. (2010). A suggested method for the measurement of world-leading research (illustrated with data on economics). *Scientometrics, 84*, 99–113.

Primack, R. (2009). Why did we reject your paper? *Biological Conservation, 142*, 1559.

Rafols, I., Leydesdorff, L., O'Hare, A., Nightingale, P., & Stirling, A. (2012). How journal rankings can suppress interdisciplinary research: A comparison between innovation studies and business and management. *Research Policy, 41*, 1262–1282.

Roa, T., Beggs, J. R., Williams, J., & Mohler, H. (2009). New Zealand's performance based research funding (PBRF) model undermines Maori research. *Journal of the Royal Society of New Zealand, 39*, 233–238.

Sampson, K. A., & Comer, K. (2010). When the governmental tail wags the disciplinary dog: some consequences of national funding policy on doctoral research in New Zealand. *Higher Education Research and Development., 29*, 275–289.

Schuch, S., Bock, J., Leuschner, C., Schaefer, M., & Wesche, K. (2011). Minor changes in orthopteran assemblages of Central European protected dry grasslands during the last 40 years. *Journal of Insect Conservation, 15*, 811–822.

Schuch, S., Bock, J., Krause, B., Wesche, K., & Schaefer, M. (2012a). Long-term population trends in three grassland insect groups: A comparative analysis of 1951 and 2009. *Journal of Applied Entomology, 136*, 321–331.

Schuch, S., Wesche, K., & Schaefer, M. (2012b). Long-term decline in the abundance of leafhoppers and planthoppers (Auchenorrhyncha) in Central European protected dry grasslands. *Biological Conservation, 149*, 75–83.

SCImago (2007). SJR—SCImago Journal & Country Rank. Consejo Superior de Investigaciones Científicas (CSIC), University of Granada, Extremadura, Carlos III (Madrid) & Alcalá de Henares, Spain. Available from http://www.scimagojr.com. Accessed July–August 2010.

SCImago (2012). SCImago institutions rankings. SIR World Rankings 2011: Global ranking. Available from http://www.scimagoir.com/ Accessed 9 April 2012.

Shewan, L. G., & Coats, A. J. S. (2006). The research quality framework and its implications for health and medical research: time to take stock? *Medical Journal of Australia, 184*, 463–466.

Smith, S., Ward, V., & House, A. (2011). 'Impact' in the proposals for the UK's Research Excellence Framework: shifting the boundaries of academic autonomy. *Research Policy, 40*, 1369–1379.

Spaapen, J., & van Drooge, L. (2011). Productive interactions as a tool for social impact assessment of research. *Research Evaluation, 20*, 211–218.

Steele, C., Butler, L., & Kingsley, D. (2006). The publishing imperative: the pervasive influence of publication metrics. *Learned Publishing, 19*, 277–290.

Stergiou, K. I. S., & Tsikliras, A. C. (2006). Underrepresentation of regional ecological research output by bibliometric indices. *Ethics in Science and Environmental Politics, 6*, 15–17.

Stinchcombe, J., & Moyle, L.C. (2002). The influence of the academic conservation biology literature on endangered species recovery planning. *Conservation Ecology* 6(2), 15. http://www.consecol.org/vol16/iss12/art15/.

UNESCO (2005). The precautionary principle. World Commission on the Ethics of Scientific Knowledge and Technology (COMEST). United Nations Educational, Scientific and Cultural Organization, Paris.

Visser, G. (2009). Tourism geographies and the South African National Research Foundation's Researcher Rating System: international connections and local disjunctures. *Tourism Geographies, 11*, 43–72.

Witten, K., & Hammond, K. (2010). What becomes of social science knowledge: New Zealand researchers' experiences of knowledge transfer modes and audiences. *Kotuitui, 5*, 3–12.

Table 1 Mean number of papers (with standard errors) published per journal in 2010 for a random selection of journals in each of four categories based on **a** ARC and **b** SCImago criteria

Category	Mean	Standard error	*N*
a			
A+[a]	241.5	37	15
A[a]	506.1	200.1	15
B[b]	76.2	24.7	15
C[b]	41.5	12.6	15
b			
Quartile 1[a]	387.3	140.1	15
Quartile 2[b]	59.7	8.8	15
Quartile 3[b]	63.7	13.2	15
Quartile 4[b]	26.1	5.5	12

Data were log-transformed before analysis. *Superscripts* adjacent to means indicate significant differences between categories based on Tukey's HSD tests

Table 2 Classification of references used in 50 randomly chosen species recovery plans from each of Australia, New Zealand and the United States

Reference type	Australian plans	New Zealand plans	United States of America plans
Journals—ARC[a]			
A+	38	24	99
A	100	56	90
B	191	239	105
C	138	274	221
C? (journal unlisted)	93	128	275
A+ and A (% journals)	24.6	11.1	24
A+ and A (% all references)	9.2	4.5	8.1
Total (% all journals)	560 (37.5)	721 (40.3)	790 (33.7)
Journals—SCImago[a]			
1 (top quartile)	228	195	220
1.5	7	2	18
2 (second quartile)	108	192	137
2.5	0	22	4
3 (third quartile)	66	147	66
3.5	0	0	25
4 (fourth quartile)	40	10	11
4? (journal unlisted)	111	153	309
1–1.5 (% journals)	42	27.3	30.1
1–1.5 (% all references)	15.7	11	10.2
Total (% all references)	560 (37.5)	721 (40.3)	790 (33.7)
Book			
Authored	130	211	256
Edited	30	14	12
Chapters	151	95	198
Total (% all references)	311 (20.8)	320 (17.9)	466 (19.9)
Conference proceedings	65 (4.4)	32 (1.8)	89 (3.8)
Grey literature			
Reports	477	624	877
Theses	34	75	80
Websites	22	4	21
Pers comms	12	6	11
Statute, maps etc.	12	4	10
Total (% all references)	557 (37.3)	713 (40.0)	999 (42.6)
Grand total	1,493	1,786	2,344

[a]Journal articles are further classified by the ranking system of the Australian Research Council (ARC) in the Excellence for Research in Australia 2010 program (Cooper and Poletti 2011) and SCImago Lab (SCImago 2007)