Jointly published by Akadémiai Kiadó, Budapest and Springer, Dordrecht

Scientometrics, Vol. 74, No. 1 (2008) 153–161
DOI: 10.1007/s11192-008-0108-1

UK Research Assessment Exercises: Informed judgments on research quality or quantity?

HENK F. MOED

Centre for Science and Technology Studies (CWTS), Leiden University, Leiden (The Netherlands)

A longitudinal analysis of UK science covering almost 20 years revealed in the years prior to a Research Assessment Exercise (RAE 1992, 1996 and 2001) three distinct bibliometric patterns, that can be interpreted in terms of scientists' responses to the principal evaluation criteria applied in a RAE. When in the RAE 1992 total publications counts were requested, UK scientists substantially increased their article production. When a shift in evaluation criteria in the RAE 1996 was announced from 'quantity' to 'quality', UK authors gradually increased their number of papers in journals with a relatively high citation impact. And during 1997–2000, institutions raised their number of active research staff by stimulating their staff members to collaborate more intensively, or at least to co-author more intensively, although their joint paper productivity did not. This finding suggests that, along the way towards the RAE 2001, evaluated units in a sense shifted back from 'quality' to 'quantity'. The analysis also observed a slight upward trend in overall UK citation impact, corroborating conclusions from an earlier study. The implications of the findings for the use of citation analysis in the RAE are briefly discussed.

Introduction

As from 1986, higher education institutes receiving funding from UK Higher Education Funding Councils have been subjected to Research Assessment Exercises (RAE) on a regular basis, with the primary purpose to produce ratings of research

Received February 12, 2007

Address for correspondence:
HENK F. MOED
Centre for Science and Technology Studies (CWTS), Leiden University
P. O. Box 9555, 2300 RB Leiden, The Netherlands
E-mail: moed@cwts.leidenuniv.nl

0138–9130/US $ 20.00
Copyright © 2007 Akadémiai Kiadó, Budapest
All rights reserved

quality, to be used by higher education funding bodies in determining the main grant for research to the institutions they fund. RAE procedures and criteria were continuously debated, and the system underwent several important changes [BENCE & OPPENHEIM, 2004]. Table 1 summarises basic characteristics of the four most recent RAEs.

Table 1. Basic characteristics of RAE 1989–2001

RAE Year	Type of institution	Time period of assessment	Staff included	Publication data requested	Assessment criteria
RAE 1989	Universities	1984–1988	All academic staff	Total publication counts; up to 2 'very best' publications	Quality and quantity equally important
RAE 1992	All higher education institutes	1989–1992			
RAE 1996		Jan. 1992–Mar 1996	Research active staff only	Up to 4 'best' publications per research active staff member	Emphasis on quality
RAE 2001		1996–2000			

Many claims as regards the effects of the RAE seem to be based on anecdotal evidence or personal impressions. But bibliometric methods provide valuable tools to assess possible effects of RAE upon domestic authors' publication practices and the effects of funding procedures upon the performance of a national research system. This paper presents the first results of a detailed bibliometric analysis of the UK publication output during the time period 1985–2004. The bibliometric patterns that were found are interpreted in terms of the responses of UK scientists to changes in the principal evaluation criteria applied in the various RAEs.

Data and methods

Data on published UK articles and received citations were extracted from the CD-ROM version of the Science Citation Index (SCI) published by Thomson Scientific, covering the time period 1985–2003. Only research articles, letters and reviews were taken into account. A UK article was defined as an article in which the list of the authors' corporate addresses contains at least one UK institution. Eight bibliometric indicators were calculated on a yearly basis, listed in Table 2.

Citation impact measures of UK articles published in a particular year were based on citations received during the first three years after publication date. Author self citations were not included. The citation counts also include citations from the year 2004. The citation impact indicator applied in this study is a normalised indicator, defined as the ratio of the average citation rate of UK articles and the world citation average in the subfields covered by these papers. Similarly, the journal prestige is defined as the average citation rate ('impact factor') of the journals in which UK articles were published, compared to the world citation average in the subfields covered by

these journals. For more technical details and applications of these indicators, the reader is referred to [MOED & AL., 1995] and [VAN RAAN, 2004].

Table 2. Mean annual growth rates in 8 indicators for the UK

Aspect indicated	Definition	MAGR (± STD) 1985–2003 (%)
UK publication output	Number of papers in the SCI with at least one UK author or institution	+ 1.5 (± 2.2)
World share of UK papers	% UK papers, relative to database total	– 0.2 (± 1.4)
Number of UK authors	Number of unique, publishing UK authors	+ 2.1 (± 1.7)
World share of UK authors	% UK unique, publishing authors, relative to database total	– 0.8 (± 1.3)
UK publication productivity	% UK papers / % unique publishing UK authors	+ 0.6 (± 0.8)
Intra-UK co-authorship	Average number of UK (co-)authors in the byline of a UK paper (intra-UK co-authorship)	+ 0.7 (± 0.9)
Citation impact of UK papers	Citation rate of UK papers compared to the world average in the subfields in which UK authors were active	+ 0.6 (± 2.7)
Prestige of journals in which UK authors published	Average citation rate ('impact factor') of the journals in which UK authors published, compared to the world citation average in the subfields covered by these journals.	+ 0.3 (± 1.6)

MAGR: Mean Annual Growth Rate. STD: Standard deviation.

A substantial fraction of the UK articles resulted from international collaboration and were co-authored by scientists from institutions outside the UK. Although the database does not for any particular paper provide tags between its authors and their corporate addresses, a statistical approach was developed, aimed at recreating these tags as accurately as possible. This approach identified for each UK article the authors that were – most probably – working at UK institutions, applying a number of plausible assumptions and decision rules.

For instance, it was assumed that the first author of an article is linked to the first address in the list of corporate addresses. Many authors, particularly senior authors, do not appear as a first author. In view of this, a second assumption was that if the corporate address lists in *all* papers authored by a particular scientist have one country in common (e.g., a person published 8 papers and *all* these contain at least one UK address in their lists of corporate addresses), he or she is most probably affiliated to an institution in that country. This methodology provided estimates of the number of UK authors publishing in a year, and allowed for the calculation of the world share of UK authors, UK publication productivity and intra-UK co-authorship. For more details the reader is referred to [MOED, 2005].

For each indicator and any pair of subsequent years, annual growth rates were calculated. In order to account for changes in the database's journal coverage over the years, a 'variable fixed journal set' was applied in the following way: the growth rate in year J compared to J-1 was based on articles in journals that were processed for the SCI during *both* years.

It must be noted that the data presented below relate to the UK as whole, i.e., to research activities in all sectors (higher education, public research institutes, hospitals, private sector). In this paper no attempt was made to collect data for the higher education sector only, the sector that constitutes the object of the RAE. A secondary analysis showed that higher education institutes (mainly universities) account for about 75 per cent of the total UK publication output. Therefore, bibliometric patterns in articles published from higher education institutes will undoubtedly be reflected in the total UK publication output, but possibly somewhat less pronounced.

Results

Figure 1 presents annual growth rates during 1985–2003 in the world share of UK papers and publishing authors. Figure 2 gives such rates for the degree of intra-UK co-authorship and UK publication productivity. Finally, Figure 3 presents annual growth rates in the citation impact of UK papers and the prestige of the journals in which they were published. Table 2 gives definitions of the various indicators plotted in these figures.

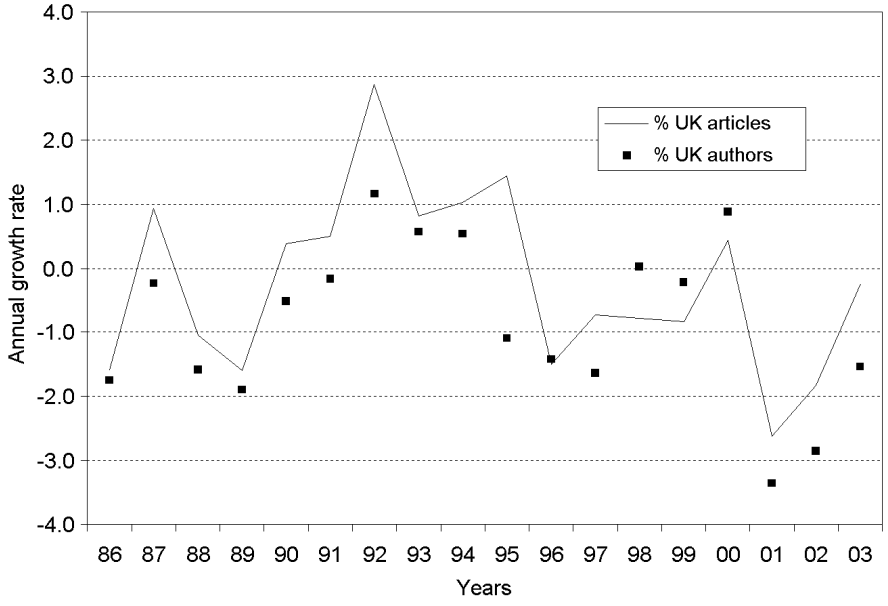

Figure 1. Annual growth rates in world share of UK papers and authors

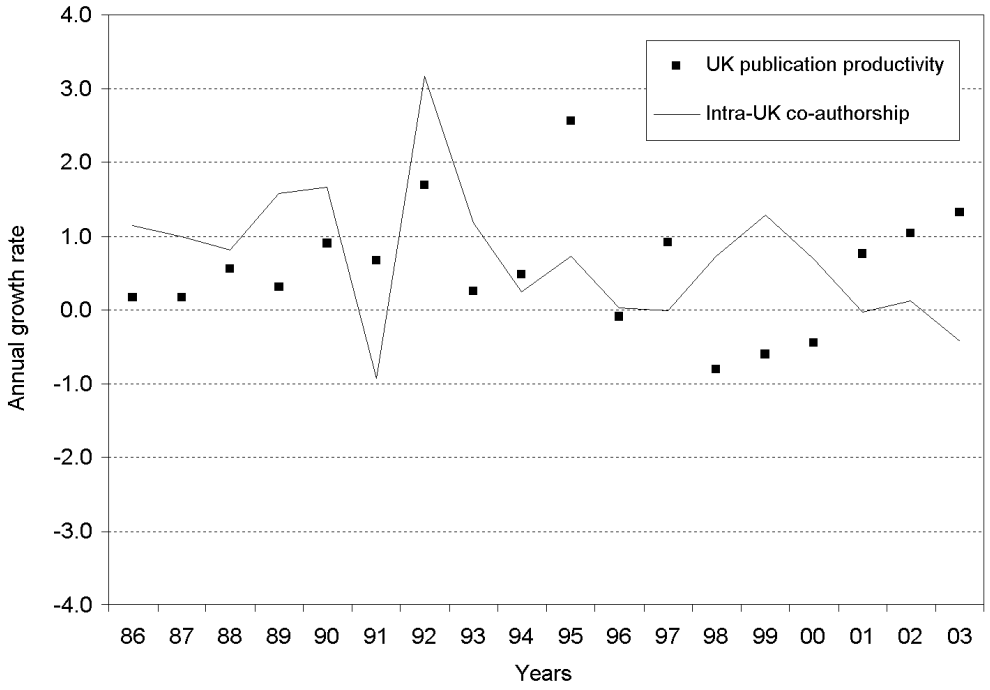

Figure 2. Annual growth rates in UK publication productivity and intra-UK co-authorship

Table 2 also presents mean annual growth rates during 1985–2003 in eight bibliometric indicators calculated for UK science. It shows an increase in the absolute number of UK articles and unique, publishing UK authors, but a slight decline in the world share of UK articles and UK authors. UK publication productivity, intra-UK co-authorship, citation impact of UK papers and the prestige of the journals in which these were published all show a slightly positive mean annual growth rate. But Table 2 also shows that the standard deviations in the annual growth rates in all indicators are rather large. This variability can at least partly be explained by analysing annual growth rates on a year-by-year basis, and taking into account the characteristics of the various RAEs carried out during the time period 1985–2003.

Figures 1–3 reveal for most indicators pronounced 'timing effects' prior to a RAE. Peaks tend to occur in the last or semi-last year of a time period covered by a RAE. Figure 1 shows that, during the four years prior to the RAE 1992, in which total publication counts were collected and assumingly used as principal indicators, the world share of UK publication output and UK publishing authors increased substantially. The figure also shows an upward peak in 1987 that can be attributed to the RAE 1998, although it is uncertain why it occurred one year earlier.

During the four years prior to the 1996 RAE, that announced a shift from 'quantity' to 'quality' and collected data on 'top' publications only, scientists published less, and also the world share of UK authors decreased substantially. At the same time, Figure 3

reveals that during this time period UK scientists gradually published relatively more papers in journals with a high citation impact. Also in the year 2000 there is a substantial increase in the prestige of the journals in which UK papers were published. The same upward peaks can be observed in 1990 and 1992.

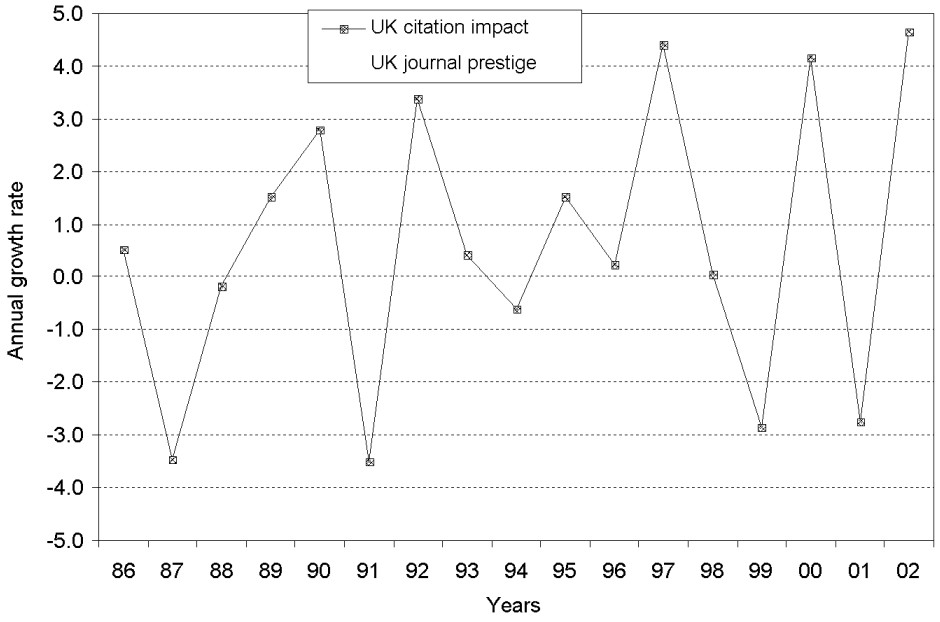

Figure 3. Annual growth rates in UK papers' citation impact in the prestige of the journals in which they were published

During 1996–2000, following a substantial drop in 1996, the annual growth rates in the world share of UK articles and authors increased, but the latter increased faster than the former. In the year 2000 particularly the share of UK authors reached a positive value, though lower than in 1992, another year in which the growth rate of the share of UK authors fairly exceeded that of the share of UK papers. Figure 2 clarifies the underlying mechanism. It shows that during 1998–2000 intra-UK co-authorship, – measured through the average number of UK authors per paper –, increased, whereas UK publication productivity – the number of UK papers per publishing UK author – declined. This phenomenon also occurred in 1992.

Figure 3 reveals that citation impact and journal prestige show similar patterns. An increase in the annual grow rate of one indicator tends to be accompanied by an increase in that of the second. This figure relates to the UK as a whole. In order to examine the degree of correlation between these two measures at lower levels of aggregation, a secondary analysis calculated these indicators for all 2,150 UK authors publishing at least 10 papers per year during the time period 1996–2005. Such authors tend to be senior researchers, and their publication oeuvre tends to represent the output of a group

of researchers rather than that of an individual. In the case that a department's director co-authors all the papers published by its members, his oeuvre represents the output of an entire department.

Table 3 shows that indicator 'number of published articles' explains none of the variance in the indicator 'citation impact'. In other words, the Pearson correlation coefficient between these to variables is 0. The indicator 'journal impact factor' is the un-normalised or 'raw' impact factor of the journals in which an author had published. It explains only 11 per cent, and the normalised journal impact indicator 23 per cent of the variance in citation impact.

Table 3. Variance in citation impact of 2,150 senior UK authors

Indicator	Explained variance in citation impact
No published articles	0 %
Journal impact factor	11 %
Normalised journal impact / prestige	23 %

Discussion and conclusions

It is questionable whether the decline in the world share of UK articles and papers observed in Table 2 can be interpreted as a genuine decline of UK research activity or output. Most bigger western countries show during 1998–2002 a similar trend in the share of domestic authors, including the USA, Canada, Germany and France (Moed, 2005). This trend may at least partly reveal the effect of displacement of papers and authors from 'established' countries by those from scientifically developing ones, such as China and Korea. Table 2 also reveals a slight upward trend in relative citation impact per UK paper and prestige of journals in which these were published. This outcome is consistent with King's conclusions from an earlier bibliometric analysis covering the time period 1993–2001 [KING, 2004].

The observed changes in annual growth rates in most cases do not exceed 2 per cent, and for some indicators, particularly those measuring journal prestige and citation impact, the behaviour is rather peaky and standard deviations (numbers between brackets) are high. Follow up studies should examine the extent to which such behaviour reflects a 'natural' variability, by analysing the research output and impact of other countries of which the academic research system is not assessed periodically and so rigorously as that in the UK. Hence, trend figures presented above should not be 'over-interpreted'. Nevertheless, the main pattern could be interpreted in terms of scientists' responses to the principal evaluation criteria applied in an RAE, or at least to their perceptions or expectations of which criteria would be the most decisive in making a final RAE judgement.

The bibliometric pattern observed in the years prior to the RAE 2001 shows a much stronger resemblance to that observed during the years prior to the 1992 RAE, than it does to the pattern obtained for the years preceding the 1996 RAE, when a shift from quantity to quality was announced. This finding suggests that, along the way towards the RAE 2001, evaluated units in a sense shifted back from 'quality' to 'quantity', particularly by stimulating their members to collaborate or at least to co-author more intensively, and thus increase the number of research active staff. This outcome substantiates evidences of a more anecdotal nature [WATKINS, 2005].

Analyses of changes in publication and citation practices are illuminating [BUTLER, 2004], but the principal question is not whether or not scientists' practices change under the influence of the use of performance indicators, but rather whether or not the application of such measures as a research evaluation tool enhances research performance and scientific progress in general. This can be assessed in longer-term studies only. The analyses presented above constitute a valuable contribution to such longer-term studies, but further research is needed, collecting non-bibliometric evidence as well.

Interestingly, at the aggregate level the annual growth rate in the citation impact of UK articles published in a year shows a strong correlation with that of the prestige of the journals in which the papers were published. Figure 3 illustrates this clearly. It is unclear which causal relationships are responsible for this correlation. On the one hand, papers in journals with high impact factors may tend to generate more citation impact because of the higher prestige and visibility of such journals, but on the other hand they have a high importance or potential influence in order to be accepted for publication in such journals. At the level of senior UK authors, representing research groups or even departments, it was found that these two indicators show a weak correlation. This finding is consistent with the conclusion from other studies that 'expected', journal citation impact is not a good predictor and not a surrogate of actual citation impact, particularly at lower levels of aggregation (e.g., [SEGLEN, 1994; GARFIELD, 1996]).

It is indeed worthwhile considering the application in the RAE of measures of actual citation impact. Proper use of sophisticated citation based indicators as quality markers can shift the assessment's focus from publication counts and prestige journals (as argued, for instance, by LIPSETT & FAZACKERLEY [2005] towards research quality.

But this paper illustrates that it should not be overlooked that the use of statistics in evaluation procedures of any kind of human achievement may influence the behaviour of the subjects under evaluation, and hence that the application of bibliometric measures in research performance assessments may and actually does influence scientists' publication practices. The use of a sophisticated citation analysis may reduce unintended effects of earlier – either formal or 'anticipated' – RAE assessment criteria upon UK scientists' publication and authoring practices, and this provides an additional reason for the application of citation analysis in the RAE 2008.

On the other hand, the notion that any evaluation system should take into account 'strategic' behaviour also applies to the construction of valid citation impact indicators and the definition of appropriate units of assessment. In addition, the use of citation analysis should be founded on the idea that citation impact, though a most useful and valuable aspect in its own right, does not fully coincide with notions as intellectual influence, contribution to scientific progress or research quality.

References

BENCE, V., OPPENHEIM, C. (2004). The role of academic journal publications in the UK Research Assessment Exercise. *Learned Publishing*, 17 : 53–68.

BUTLER, L. (2002). A list of published papers is no measure of value. *Nature*, 419 : 877.

GARFIELD, E. (1996). How can impact factors be improved? *British Medical Journal*, 313 : 411–413.

KING, D. A. (2004). The scientific impact of nations. *Nature,* 430 : 311–316.

LIPSETT, A., FAZACKERLEY, A. (2005). RAE shifts focus from prestige journals. *Times Higher Education Supplement*, 22 July.

MOED, H. F., DE BRUIN, R. E., VAN LEEUWEN, TH. N. (1995). New bibliometric tools for the assessment of national research performance: database description, overview of indicators and first applications. *Scientometrics,* 33 : 381–442.

MOED, H. F. (2005). *Citation Analysis in Research Evaluation*. Springer, Dordrecht, the Netherlands.

SEGLEN, P. O. (1994). Causal relationship between article citedness and journal impact. *Journal of the American Society for Information Science,* 45 : 1–11.

VAN RAAN, A. F. J. (2004). Measuring Science. In: MOED, H.F., GLÄNZEL, W., SCHMOCH, U. (2004) (EDS). *Handbook of Quantitative Science and Technology Research*. The use of publication and patent statistics in studies of S&T systems. Dordrecht (the Netherlands): Kluwer Academic Publishers, 19–50.

WATKINS, D. (2005). Authors per paper. *Sigmetrics Digest*, 19–20 May 2005 – Special Issue (#2005-77).