Estimating the Quality of Business and Management Journals from a Peer Review Exercise: The 2008 RAE in the UK

John Mingers (corresponding author)

Director of Research, Kent Business School, University of Kent, Canterbury CT2 7PE

E: j.mingers@kent.ac.uk

T: 01227 824008

Kate Watson

Kent Business School, University of Kent, Canterbury CT2 7PE

E: kate@waterdancer.co.uk

Maria Paola Scaparra

Kent Business School, University of Kent, Canterbury CT2 7PE

E: m.p.scaparra@kent.ac.uk

Estimating Business and Management Journal Quality from the 2008 Research Assessment Exercise in the UK

Abstract

The 2008 Research Assessment Exercise in the UK involved the peer review of over 12,500 research outputs in Business and Management, of which 92% were journal articles. Each output was graded on a 4-point scale from "world leading" to "national" with a fifth point being unclassified. These grades were accumulated for each department to provide an overall quality profile in terms of the proportions of its outputs in each category. The assessments of individual papers were not made public but the papers submitted by each department were. This data provides a major opportunity for addressing issues of concern about the evaluation of research and the effects of journal rankings, as well as the possibility of reconstructing the judgements made by the Panel about journal quality. Given the submission details and the resulting grade profile for each department, we have used linear programming to produce the best estimate of the grades awarded to papers from each journal that had more than three entries. This provides both a grade profile for each journal and a single quality estimate. The results are shown to have good validity in comparison with other journal rankings. Apart from providing a ranking of 700 journals based on the RAE results, the paper is also able to shed light on issues such as the accuracy and coverage of the ABS ranking; the degree of selectivity of submissions; the dispersion of grades for a journal; and differences between different subject areas.

Key Words: journal rankings, linear programming (LP), Research Assessment Exercise (RAE), research evaluation

1. **Introduction**

There is an increasing drive towards measuring the research quality of academics whether it is for the purposes of promotions, jobs, or assessing the performance of departments, research institutes or even whole universities. Quality can only be judged through the activities and publications produced, especially journal papers as that, in Business and Management, is the primary currency. However, assessing the quality of individual papers by peer review is itself time consuming, requires expert(s) in the area, and is open to disagreement. Partly for these reasons, the quality of the journal that the paper is published in is often taken as a proxy for the quality of the paper itself. This then displaces the problem to judging the quality of journals (and assuming that all papers within are of equal quality) – hence the proliferation of journal rankings.

There are two main ways of generating rankings – stated preference (peer review), where some group of experts determines a ranking, and revealed preference where actual publishing behaviour is measured usually in terms of the citation impact factor (IF) or the h-index. A third approach is to statistically combine a set of already existing lists (Mingers & Harzing, 2007). There are many ranking lists available on the Harzing website (http://www.harzing.com/) but for Business and Management the list produced by the Association of Business Schools (ABS) (http://www.the-abs.org.uk/?id=257) has become extremely important in the UK. There is evidence from, for example, the Committee of Professors in Operational Research (COPIOR) that it is being used in decisions about appointments, promotions and even probation/tenure as well as to influence academics'

publication decisions. There are two major problems, however – the lists do not agree with each other, and in reality there will be a range of quality within any one journal.

The use of journal rankings as a proxy for quality is actually extremely contentious. For example, Paul (2007, 2008), who was a member of the 2008 RAE Panel, states that "One major conclusion appears to be that journal rankings are not a good indicator of the quality of any paper published in that journal, nor necessarily the combined quality of all the papers" (Paul, 2008, p. 324). Macdonald and Kam (2007), in a bitter critique, suggest that the whole world of academic publishing in management is one of gamesmanship and game playing with the so called quality journals simply reproducing standard, consensual research within a small elite community. Clark and Wright (Clark & Wright, 2007), then editors of the *J. of Management Studies*, disagreed and argued that journals do develop and change in response to their communities, and that the reviewing processes of high quality journals do in fact lead to high quality papers. Adler and Harzing (2009) provide another strong critique of the dysfunctional effects of academic ranking systems and journal rankings in particular. The main complaint is that they lead to a narrowing of the discipline, concentrating research into the narrow confines of established journals and discouraging innovation and interdisciplinary work.

In the light of these debates, the latest (2008) UK Research Assessment Exercise (RAE) provides a major opportunity to investigate the extent to which journal rankings are concordant with direct judgements of the quality of individual papers. A quality profile was generated in terms of the proportion of the department's research that was judged to be on a 4-point scale plus unclassified. Full details of the RAE can be found in various reports issued by HEFCE (RAE, 2004, 2005, 2006) prior to it, and the results were announced in 2008 (RAE, 2008). Also available online are the subject overview report for the Business and Management Panel (RAE, 2009a); the complete submissions (RAE, 2009b); and the quality profiles (RAE, 2009c).

Quality itself was defined in terms of three characteristics – originality, significance and rigour – and the levels were:

- 4*: Quality that is world-leading, that has become, or is likely to become, a primary point of reference in the field or sub-field.
- 3*; Quality that is internationally excellent, that has become, or is likely to become, a major point of reference in the field or sub-field.
- 2*: Quality that is recognised internationally, that has made, or will make, a contribution to knowledge, theory, policy or practice.
- 1*: Quality that is recognised nationally, that has made or will make a limited contribution.
- Unclassified (0*): Quality that falls below the standard of nationally recognised work or which does not meet the definition of research.

The requirement of the Panel was to assess a department's quality in terms of three dimensions: their submitted research outputs (publications of all types, although

predominately journal papers); the research environment; and the esteem of the staff members. A profile was generated for each dimension and these were combined (70%, 20%, 10%) to produce the overall quality profile for the department. Thus a department might have ended up with a profile of 25% 4*, 40% 3*, 25% 2* and 10% 1*. From the profile a single-valued grade point average (GPA) is calculated (in this case 2.8) and this is used for producing league tables of departments and then universities. As we are only concerned with the outputs, i.e., papers, the environment and esteem profiles will not be considered.

The Panel was therefore required to produce a quality grading for every single piece of work submitted, in this case 12,575 papers. This was clearly a huge task and initially the B&M Panel stated that it would look at 25% in detail although the actual number was never published. The results that were made public consisted of the grade profile (i.e., % of papers in each quality level) for each of the 90 institutions that submitted together with details of all the publications. However, the grades for individual outputs are not available.

Prior to the exercise, the Panel Chair (Professor Mike Pidd) made it clear on several occasions that the Panel did not intend to use journal ranking lists in making their judgements. He also stated that they expected to find a range of qualities within a single journal. What was not clear was how they would in fact grade outputs if they were only actually going to read a proportion of them. In the event, the Panel claimed that "most outputs were read in considerable detail" (RAE, 2009a, p. 5). It would seem to be a formidable task: 12,600 outputs to be read by 18 academics (700 each) in a few weeks, but nevertheless it does represent a major exercise in directly assessing the quality of research outputs. However, little is said in the review reports about precisely how the quality judgements were made, how the grade boundaries were determined, or the extent of consensus or dissensus among Panel members.

The purpose of this paper is to try to use the peer review quality judgements made by the Panel to evaluate journal quality and journal ranking lists such as the Association of Business School's one. Geary et al (2004) performed a similar task after the 2001 RAE although their approach was somewhat indirect. They assumed that staff in higher quality departments will tend to publish in higher quality journals and that therefore one could assess a journal's quality by the RAE grade of the departments that submitted it. Frequency counts were calculated for each journal and it was then awarded points on a 7-point scale corresponding to the RAE grades for each department that submitted it. An average score for the journal could then be calculated using the mean, mode or median. This method has obvious drawbacks: it does not discriminate that well between journals; the score for a department depends on things other than the research outputs; and it ignores the fact that there may be "islands" of excellence in otherwise weak departments.

A similar method could be used for 2008 based on the mean or GPA of the department's profile, but we are proposing a more sophisticated approach that relates the outputs submitted and the quality profile awarded to them for each of the 90 departments. This is done using linear programming (LP). In brief, we create a set of decision variables for each journal that represent the 5 possible quality levels (including unclassified as zero). We then use LP to find

the values of those variables that minimise the difference between the estimated quality profile (calculated from the variables) and the actual quality profile awarded to each department. The approach is analogous to least squares regression but with several thousand variables to be determined. The result is an estimate of the proportion of papers from a journal that were awarded the various levels of quality. In undertaking this analysis we are not suggesting that the Panel came to its decisions by using journal rankings. We are exploring the extent to which the Panel's actual results, given only in aggregate form, can be replicated by an analytical method. We are not trying to replicate the process used in arriving at the results.

In the first section we compare the outputs submitted to the 2008 RAE with those from previous ones. We then describe the methodology including the mathematical model(s) developed, and the data cleaning and manipulation. Finally, we explore the results obtained and comment on their validity.

2. Comparison of the 2008 and 2001 submissions

In this section we will just present the basic facts of the RAE submissions in comparison with previous ones without considering the quality levels. Note that the journal data has been the subject of a cleaning process which is described in the next section. In particular, papers that were submitted as internet journal publications, i.e., they had been published electronically but not in print, were allocated to the appropriate print journal.

Table 1 shows that since 1996 there has been a contraction of the number of submissions, and presumably departments, but an increase in the number of staff and publications submitted. Staff have increased by 43% but outputs by nearly 60%. It is also noticeable that journal papers have come to dominate the submissions reaching 90% in 2008. Other forms of research such as authored books, edited books and research reports are certainly not being submitted to the RAE, whether or not they are actually being produced. The number of different journal titles is also rising inexorably although it is not necessarily the case that they are all highly regarded as later results will show. The mean number of entries per journal and entries per institution has also risen significantly.

Table 1 Submission statistics for the last three RAEs

The dominance of journals can be seen more clearly in Table 2. Authored books and especially book chapters have fallen dramatically. This may well be attributable to comments by Panel members before the submission that one had to be careful with outputs that had not been peer reviewed, i.e., that were not in refereed journals. It will be interesting to see how

the REF (the next RAE) tackles this problem with its increased emphasis on the external impact of research not merely its scholarly impact.

Table 2 Number of publications by output type

Looking within the journals, it has always been the case that submissions follow the Pareto rule – a small number of journals account for a large proportion of the submitted outputs and vice versa.

Table 3 Dispersion of submitted journal outputs

We can see from Table 3 that around 50% of the journals only have a single entry in the RAE although this proportion reduced in 2008. Over 70% of journal titles have 4 or less entries. On the other hand, a relatively small number of journals account for a high proportion of total entries. The 105 most common journals between them account for 50% of the journal outputs submitted and as Figure 1 shows, the top 20% of journal titles account for almost 80% of the submitted outputs.

Figure 1 Pareto curve for the number of entries per journal in the 2008 RAE

Concentrating on these, we can compare the most popular journals with those from 2001. The 20 most frequently submitted journals from 2008 are listed in Table 4 along with their relative positions on the Geary equivalent for 2001. These journals represented 22 % of the 2008 outputs, compared to 20% of those in 2001. 15 journals have retained their place in the top twenty, while five have slipped out - *Human Resource Management Journal, Industrial*

Relations Journal, Personnel Review, Applied Economics and *Long Range Planning*. The top twenty journals cover most of the spectrum of business and management and it is interesting that two top journals classified as social science have entered – *Regional Studies* and *Research Policy*. All but *Service Industries Journal* are 3* or 4* in the ABS rankings.

In the light of the discussion in the introduction about rankings lists leading to standardisation, it is interesting to see what proportion of the journals submitted are actually included in the ABS list (note that this research used version 9 of the list, the latest one available at the time) . Figure 2 shows that there were 825 journals in the RAE that are not in ABS; 224 journals in ABS that were not entered in the RAE; and 814 that were in both. The first figure shows that 50% of the RAE journals are not actually included in ABS which is the most comprehensive listing of B&M journals there is. There would seem to be two possible reasons: genuine business and management journals that ABS has not yet included, and journals that are not business ones. These would typically either be applications journals, e.g., health services or construction, or other relevant disciplines e.g., philosophy or social science. The latter examples could be seen as healthy interdisciplinary and applicability, or they could be seen as business schools being somewhat of a "dumping ground" for academics who do not fit well in other, more focussed, departments.

Figure 2 Numbers of journals in the RAE and the ABS list

The issue of fragmentation within B&M submissions is important as Bence and Oppenheim discuss (2004). First, it is very difficult for Panel members to genuinely have expertise across such a wide range of subjects although they are able to cross-refer papers to other panels. The 2001 Panel expressed quite strong concerns about this problem (Bessant, et al., 2003). This time the Panel overview was more sanguine but still concluded that "Some submissions … seemed to be of little or no relevance to business and management studies and … some submissions were an over-eclectic mix of outputs" (RAE, 2009a, p. 5). The second, related concern is that even if Panel members consider themselves competent to judge a paper there may be an unconscious bias towards papers published in core business journals rather than more peripheral ones. Some evidence relevant to this will be presented in the results section.

On the other side, we can see that 22% of the ABS journals were not actually submitted in the RAE. At first sight this seems quite high, especially given the large range of journals that were submitted. The main explanation is likely to be that they are the lower ranked journals which departments chose not to submit for fear of getting low quality gradings. This is confirmed when we see that of the 224 unsubmitted ABS journals over 80% are graded at 1* or 2*. In the results section it is shown that there is a lower proportion of 1* journals in the submitted ABS journals than in the ABS list as a whole.

Table 4Twenty journals most frequently submitted in the 2008Business and Management RAE

3. Modelling the aggregate RAE quality evaluations

3.1 The LP Model

Linear programming (LP) is a mathematical method which determines the values of a set of decision variables so as to maximise or minimise a linear function of those variables (the objective function) subject to a set of linear constraints. In our situation we know the quality profiles for each institution and we also know how many entries for each journal the institution submitted. We can then create two sets of variables – the grade profile for each journal and the estimated grade profile for each institution. The grade profile for a journal consists of five variables each of which represents the proportion of the journal's entries judged to be in a particular category (0* – 4*). The sum of the 5 variables for a journal must sum to 1.

The estimated grade profiles for institutions are formed from the journal grades, weighted by the number of articles an institution submitted from each journal. The estimated profile is, for each institution at each grade, the sum of each journal grade at grade (0*- 4*) multiplied by the number of articles from the institution submitted from that journal divided by the total number of articles in that department's submission. The objective (function) is then to minimise the difference between the estimated profile and the actual profile for each institution by finding the best values for the journal grades.

Initial model (QP1)

Let:
j index the journals (j = 1 .. no. of journals)
g index the grades 0* - 4* (g = 0 .. 4)
i index the universities (i = 1 .. no. of institutions)
u_{ig} be the actual proportion of research at grade g for university i
n_{ij} be the number of entries of journal j submitted by university i
e_{ig} be the estimated proportion of research at grade g for university i
p_{jg} be the estimated proportion of the outputs of journal j graded at grade g

$$Min. \sum_i \sum_g \left(u_{ig} - e_{ig}\right)^2 \qquad (1)$$

s.t.

$$e_{ig} = \frac{1}{\sum_j n_{ij}} \sum_j n_{ij} p_{jg} \qquad \text{for each institution } (i) \text{ and grade } (g) \qquad (2)$$

$$\sum_g p_{jg} = 1 \qquad \text{for each journal } (j) \qquad (3)$$

$$p_{jg} \geq 0$$

$$e_{ig} \geq 0$$

The objective function (1) minimises the squared differences between the actual and the estimated proportion of research outputs at each grade level for each department. Constraint (2) defines the estimated proportion in terms of the number of entries of a journal multiplied by the proportion of the journal at a particular grade and divided by the total number of entries for that department. Constraint (2) ensures that the grade proportions for each journal sum to 1. It is possible to formulate this model without explicitly using an estimated proportion variable, but we have done it this way for clarity.

As formulated, this is actually a quadratic program as the objective function is quadratic. Since solving large quadratic programs is generally computationally more expensive than linear ones an alternative model was produced with a linear objective minimising the absolute difference rather than the squared difference.

Alternative model (AbsVal1)

$$Min. \sum_i \sum_g \left| \left(u_{ig} - e_{ig} \right) \right|$$

Although the absolute value function is itself non-linear it can be easily linearised by generating a new variable (err_{ig}) and two new constraints:

$$Min. \sum_i \sum_g err_{ig} \qquad (4)$$

s.t.

$$err_{ig} \geq \left(u_{ig} - e_{ig} \right) \qquad (5)$$

$$err_{ig} \geq \left(e_{ig} - u_{ig} \right) \qquad (6)$$

$$e_{ig} = \frac{1}{\sum_j n_{ij}} \sum_j n_{ij} p_{jg} \qquad \text{for each institution } (i) \text{ and grade } (g) \qquad (2)$$

$$\sum_g p_{jg} = 1 \qquad \text{for each journal } (j) \qquad (3)$$

$$p_{jg} \geq 0$$

$$e_{ig} \geq 0$$

$$err_{ig} \geq 0$$

Here constraints (4) and (5) between them ensure that *err* takes on only the positive difference between actual and expected.

A third model was also developed with the idea of determining a single integer quality grade for each journal rather than a grade profile. This was easy to achieve in the formulation by simply restricting the journal grade variables (p_{jg}) to being 0-1 integers. The constraint that they must sum to 1 for each journal ensures that only one of the five possibilities will actually be 1 and so each journal will have only one grade level. This model (**MIP1**) proved very difficult to solve computationally using either the quadratic or the absolute value objective function.

We should discuss one possible concern with the model as formulated. The model is trying to estimate the proportions of 4*, 3* etc. within a particular journal across all the departments that submitted papers from that journal. It is possible to show that, even if you know the actual proportion of different grades in the journal for a particular department (which in practice you never do), the model does not yield the correct estimate of the department's overall quality. In other words, it would appear as though there is some built-in error in the model preventing it getting the right answer.

The reason for this is that the proportion of 4*, 3* etc estimated (or even known) for the journal as a whole comes from all the departments that submitted it, not just one. In other words, the overall proportions of grades within a journal are actually a kind of average across all the departments. It cannot therefore be expected to reflect the actual values for any one department unless, of course, the distribution of grades for the journal was identical in all departments, which is extremely unlikely.

In general, the greater the disparity of grade distribution for a particular journal across different departments, the greater will be the inaccuracy for any one department. We cannot actually know the real figures because the RAE did not release them but we do not think that there is any reason to assume it should be particularly extreme. We cannot see that, in practice, papers from a particular journal submitted by one department would systematically receive very different gradings to those in the same journal submitted by another department.

Even if it did occur, perhaps by chance in particular examples, we do not think that it would produce a systematic bias – merely a degree of over-or under- estimation as one would expect with an averaging process. We believe that the overall validation of the results, as presented later in the paper, demonstrate the underlying soundness of the model.

3.2 Cleansing the data

Each submitting institution used a pro-forma to enter details of their outputs. After the publication of the RAE 2008 outcomes, the details of the individual outputs were released. The spreadsheet <RA2> was downloaded from the business and management sub panel section of the RAE website. There were 12575 records in the data set. Each output had to be placed within one of twenty categories (summarised in Table 2).

The RA2 data was used to derive a list of all journals, along with the number of outputs from each journal, cross-referenced with the institutions submitting those outputs. As journal nomenclature can be imprecise, journal titles were checked to ensure that no journal is listed more than once ("The Journal of Example" and "Journal of Example" must be resolved, is there an error in one of the titles? Are there two distinct journals?). We also used the journal ISSNs which were part of the RAE data but again there was a good deal noise here: some were entered as text and some numeric; some were incorrect; and some journals actually have more than one ISSN. The online papers were treated as a different category in the RAE data (type H) but we amalgamated those with their printed equivalents. The journal titles adopted in the ABS list of journals were used in preference to any other variants found in the RA2 data.

The next issue was what to do with the outputs that were not journal papers, in the main books and book chapters. We could not include each item individually as if it were a journal because the model only works to the extent that the same journal occurs in a number of submissions. We could simply leave them out which would increase the residual variation in the results but lose information. So, what we did was to include each type (authored book, book chapter etc) as if it were a journal. Thus all 285 authored books were included as if they were a single journal. This increased the accuracy of the model and also allowed us to see how these output types were treated by the RAE Panel. Were books rated highly or lowly? These categories (Book, Book Chapter, Edited Book, External Report and Other) represented 950 outputs (7.6% of the whole dataset).

We also had to decide what to do with all the journals that had only a small number of entries. The problem is that if the journal only occurs a small number of times it becomes essentially unconstrained and the model can use it simply to fill in unexplained variation. After some experimentation we decided to only include in the model those journals that had at least three entries. This meant excluding around 57% of the journal titles (see Table 2). We recognize that this may introduce some bias into the model but it was not apparent what this might be other than they were generally of a low quality.

The final output of this process was a 2-dimensional array indexed by journal name and institution. The full list of cleaned data is available from the authors. As noted by Geary et al (2004), the process of cleansing the RA2 data was the most intensive part of the project, and the results produced may not be identical to others attempting the same task. Using the ABS list as a standard and automating the search process errors are kept to a minimum.

3.3 Solving the models

With the data arrays prepared and the linear programme constructed, the programme was coded using the OPL Studio 4.1 modelling language and solved by the CPLEX 11.0 optimizer. Several runs of different versions and sizes of the model were conducted. The final version of the AbsVal1 included 701 journals and 89 institutions. It included roughly 4,400 variables and 2,050 constraints. It solved in about a minute and gave an objective function value of 23.6, i.e., the sum of all the 445 errors. With a model like this where the variables are relatively under-constrained, there may be many solutions which differ marginally and

give broadly similar results. The sensitivity was explored and although there were many reduced costs with low values there were none with zero. In terms of the validity of the solution this is best evaluated in terms of concordance with other evidence, a task that is carried out in Section 4.1

The integer version proved to be computationally very expensive. After running continuously for 35 days it had still not reached an optimal, fully integer solution. This is not unusual with models that have a large number of integer variables (3500 in our case). It had in fact converged to a near optimal which did not change significantly over 21 days but could not be shown to be the actual optimal.

This model gave a grade profile for each journal but for the purpose of constructing a journal ranking and comparing it with existing ones it is more appropriate for each journal to have a single grade. There are several ways of achieving this: take the modal grade, i.e., the one with the largest proportion; calculate the mean grade (i.e., the GPA) and then round this to the nearest integer; or get the LP to calculate the best value with the integer version of the model (MIP1).

After inspecting and comparing the results for the three different methods – mode, rounded mean, and MIP1 it was decided that the mode gave the fairest and most consistent results and so this has been used in the ranking comparisons, but the final table of results (Table 5) includes the grade profile for each journal and the MIP1 results.

4. The Results

The full results are too large to be printed in the paper but are referred to as Table 5. This presents results for all the journals included in our model, i.e., those with at least three entries. We show the grade awarded based on the mode of the journal profile; the profile itself in terms of the proportions judged to be in each rank; the number of items submitted, the ABS rank where available and the subject classification. The Table is available from the following website: http://www.kent.ac.uk/kbs/community/staff/profiles/mingers_john_bio.html.

Table 5Ranking of journals reconstructed from the RAE (available from http://www.kent.ac.uk/kbs/community/staff/profiles/mingers_john_bio.html)

4.1 Assessing the validity of the reconstruction

Before presenting the results in detail, it is important to evaluate their degree of validity. The philosophy of the model is that, given the aggregate results from the Panel and knowing the papers that were submitted, it should be possible to reconstruct to some extent the grades that were awarded at the journal level. Clearly, if the results we obtain are wholly at odds with our preconceptions of journal quality we might conclude that they were not capturing

anything meaningful. But, we would not expect them to be identical with the existing rankings, partly because of noise in the data resulting from the non-journal outputs being removed, and partly because the Panel were clear that their results did not mirror the existing lists (RAE, 2009a, p. 1). So validity is a matter of degrees of concordance.

We first consider the extent of concordance with existing journal rankings.

Table 6 Correlations between reconstructed RAE grade and journal rankings

Table 6 shows the correlations between the reconstructed RAE grades and the ABS, Kent (Mingers & Harzing, 2007) and the Geary et al (Geary, et al., 2004) rankings. Given the large numbers of observations (shown in parenthesis) all the correlations are highly significant. It is noticeable, however, that they are not as high as the correlations between the rankings themselves tend to be, as shown in the ABS documentation. We can see for example that the correlation between the Kent ranking and the ABS one is significantly higher. Nevertheless, it is clear that the RAE model is broadly in line with these rankings. We also looked specifically at the extremes – the 1* and 4* journals. These had a higher correlation (0.609) perhaps showing that there is greater agreement about the best and worst journal and less about the boundary between 2* and 3*.

Some further evidence is shown by the treatment of non-journal outputs. As explained in the previous section, rather than totally ignore outputs such as books, book chapters and reports we included them as if they were a single journal. This generated a score for each of these categories so that we could see how the category was treated in comparison with the journals. The results are shown in Table 7.

Table 7 Reconstructed grades for non-journal outputs

In the third column we can see the mean grade awarded to each output type. From a validation perspective the order of these types is what we would have expected, i.e., authored books were graded most highly, going down through book chapters to edited books. Other types (e.g., software) and external reports were least valued. This again gives us a degree of confidence in the overall method. In terms of the actual numbers, there was a concern before the RAE that books would be downgraded because they were not refereed. This does seem to

have happened in that one might expect that a quality book would be regarded more highly than a single paper and so books should have achieved a high grading – at least 3* or more. We can also see that external reports scored poorly which does not bode well for the REF trying to encourage the submission of work that has external impact.

Finally we look at the journals that come out top from our reconstruction in Table 8. In terms of our estimation, the best journals are those that have the highest proportion of 4*. In the Table, we have selected all journals that have at least 50% 4* and we have restricted it to those with at least 12 submissions. This results in 30 journals which are ordered in terms of the % 4*. Firstly, all 30 journals are in ABS and only three were less than 3* or 4*. They were also generally ranked highly in other lists including Geary's analysis of the 2001 RAE and the citation impact (CI) factor. Interestingly, 14 of them are also included in the FT top-40 list of journals which is used to rank business schools worldwide. Of the rest of the FT-40 list, all but four were graded 3*, those being *Human Resource Management (USA), International J. of HRM, J. of Business Ethics,* and *J. International Business Studies* which were only graded 2*. Table 8 also includes a sprinkling of the very top American journals such as *AMR, Management Science, Organization Science, HBR* and the *American Economic Review*. Given that these results have been generated purely by the model it does give us confidence that the results do reflect judgements about journal quality.

Table 8 Top 30 journals from RAE (>=50% 4* and >= 12 entries)

4.2 Comparing the RAE grades with the ABS ranking

As the ABS list has become the *de facto* standard for Business and Management in the UK, and is used extensively, for better or worse, in making decisions about appointments, promotions and submissions, it is important to see how it compares with the reconstructed RAE grades.

Table 9 Proportions of journals in particular ranks comparing ABS with RAE grades

Table 9 shows the proportions of journals awarded different grades from the ABS ranking and our RAE reconstruction. The first column shows the proportions in the total ABS list, whether or not they were submitted in the RAE, with a GPA of 2.17. Column 4 shows the proportions for all those journals in our RAE list (remembering that it excludes journals with less than 3 entries) with a GPA of 2.34. The proportions are significantly higher (X^2_3=31.9)

than the ABS list but that is to be expected because of the selectivity exercised in submitting to the RAE (this will be discussed later). The proportion of 4* and 3* is higher, and 2* and 1* lower. A fairer comparison is to consider only those journals that are in common between ABS and our RAE list – columns 3 and 6. The two proportions are in fact very similar with GPAs of 2.43 and 2.42 although there are significant differences (X^2_3=15.6) within grades with the RAE giving more 4* and 1*, and less 2*. In other words, the RAE gave more of the extreme grades.

We can also look at the distribution of differences between the RAE and ABS. Table 10 shows, for each RAE grade, the numbers of journals that were 0, 1, 2 or 3 grades away in ABS. So, of those journals graded 4* in the RAE, 31 were also 4* in ABS, 38 were 3* in ABS, 32 2* and 4 1*. At the other end, 18 ABS 3* were graded 1* in the RAE and 2 4* were graded 1*. These were: *Accounting Review* and *Journal of Rural Studies.* although they had relatively small numbers of submissions.

Table10 Differences between RAE 2008 and ABS

4.3 Selectivity of journal submission

We now move to the issue of selectivity of journal submission. On the one hand, as we saw in Section 2, there were an increased number of journals entered into RAE 2008 and a significant number of these are not even in the ABS list. This suggests a wide range of material. However, at the same time there is continual pressure on institutions to submit only the best work and this pressure will grow. There is currently concern that increasingly the top business schools will limit their academics to publishing only in the top A-rated journals. Indeed, the Parliamentary Select Committee on Science and Technology raised these very concerns in a report in 2004:

"The perception that the RAE rewards publication in journals with high impact factors is affecting decisions made by authors about where to publish. We urge HEFCE to remind RAE panels that they are obliged to assess the quality of the content of individual articles, not the reputation of the journal in which they are published." (Select Committee on Science and Technology, 2004).

Guidelines recently issued concerning the 2013 RAE (the REF) (HEFCE, 2009) say that they aim to support quality rather than quantity and the number of academics and papers is likely to reduce. This will lead to institutions focussing even more on those believed to be high quality journals.

The degree of selectivity can be seen from column 2 of Table 9 which shows the grade proportions in those ABS journals that were not submitted in the RAE. These are significantly different to the profile of ABS journals that were submitted (X^2=60.1). We can

see that 45% of those not submitted were 1* while only 4% were 4*. This clearly shows the extent to which attention was focussed on those ABS journals that are at least 2*.

The possible results of this effect can also be seen in Figure 3 which is a scattergram of the proportion of an institution's submission in ABS journals on the x-axis and the GPA gained by institution on the y-axis. The correlation coefficient is highly significant (0.6) and it explains 36% of the variation in GPA by itself. Taken at face value, this shows that the greater the concentration on ABS journals the better an institution did in its GPA. This might suggest that the RAE Panel grades papers from ABS (or at least mainstream business and management if not ABS *per se*) journals more highly than others.

Figure 3 Scattergram showing association between GPA and proportion of an institution's submitted journals that are in ABS

There are other possible interpretations of this association. One might suppose that high quality institutions produce more papers that are in the mainstream of B&M anyway, and that there will be more papers available to be selected, so that the institution can choose mainly ABS ones. Whereas poor quality institutions have to make do with what papers they have, and may include more academics from the fringe areas. On this interpretation, the association would be indirect rather than causal – the high GPA and the high proportion of ABS both reflect underlying high quality rather than one causing the other. Alternatively, one could interpret it as reverse causality and as evidence for the selectivity effect mentioned above – the better quality institutions are more rigorous in limiting their staff to ABS-only journals.

We can get some more evidence directly from columns 5 and 6 of Table 9. This shows the distribution of reconstructed RAE grades for ABS and non-ABS journals. Did the RAE Panel actually grade ABS journals higher than non-ABS ones? They are significantly different (X^2_4=39.9) but although there are fewer 4* and 3* than would be expected in the non-ABS journals, the biggest difference is that 13% of the non-ABS journals were allocated 0* as opposed to only 2% of the ABS ones. In other words, according to our estimates a significant proportion of the non-ABS papers were considered to be of no research merit. This could be a legitimate response of the Panel to submissions that were not relevant to business and management. They state,

"In a very limited number of cases, such left-field outputs were given low grades because of their lack of relevance" (RAE, 2009a, p. 5)

But it could also reflect a conscious or unconscious bias towards recognised journals regardless of paper quality.

Overall, we feel that there is evidence both of extensive selectivity in submissions and possible bias in judgements against non-mainstream B&M journals.

4.5 Dispersion of grades for a journal

Another issue in connection with journal rankings is the extent to which the RAE Panel would award all papers in a particular journal the same grade which would indicate that they simply went by the ranking of the journal. The Panel stated both before and after that they did not intend to do that, and the results do back them up to some extent.

Given that we are choosing the grade of a journal by its modal grade, i.e., the grade with the greatest proportion, we can measure the degree of dispersion by the percentage that is *not* in the modal grade. Journals with 100% in one grade will thus have zero dispersion. The greatest dispersion a journal could have is 66% with 34% being in the modal grade.

Table 11 shows the frequency distribution of dispersion. In fact the majority (62%) have been found to have 100% at a particular grade. Clearly this is only the estimate from our model and we do not know if this is the actual case but there is no reason for our model to choose 100% particularly and one would expect that it would do the best it could to match the grade profiles so as to minimise the squared deviations in the objective function. So this evidence would suggest that many journals were seen as having only one quality level (although not necessarily the same as their ABS grading of course).

Table 11 Distribution of dispersion of journal grading

Table 12 shows those journals that are ranked as 4* in ABS but which have a high degree of dispersion in the RAE results. As can be seen, there are some well known journals here and most have a large number of entries so the results should be reliable. In many cases the split is just between two adjacent grades, e.g., *Organization Studies* or *J. of Marketing*, but in some cases it is split much more widely, e.g., *British J. of Social Psychology* or *Organizational Research Methods*. The top journals with little dispersion were shown in Table 8.

Table 12 ABS 4* journals with a high degree of dispersion

4.6 Differences between subject areas.

It is of interest to look at the relative grading between subject areas. Reports from both the 2001 RAE (Bessant, et al., 2003) and the 2008 RAE (RAE, 2009a) make it clear that the subjects were seen to have different levels of quality. So, to what extent is that borne out by the ratings? Figure 4 shows the mean journal grading by the ABS sector for the journal where it was in ABS. Those not in ABS have been given the title "#N/A".

Figure 4 Mean RAE score across various subject areas

The data show a significant difference from the highest sector, Psychology, with a weighted average score of 2.8 down to Tourism and Hospitality with a score of 1.3. The non-ABS journals have an average of 2.0. The ABS sectors are somewhat different to the subject groups that the RAE Panel report discusses. We can see that many of the long-established disciplines (e.g., Psychology) and management areas (e.g., Accounting and Finance, Operational Research and Organisational Studies) scored highly while newer and perhaps more applied areas (e.g., ethics, management development, innovation and tourism) did less well. Some surprises perhaps are the poor score for Strategy and perhaps the relatively high scores for General Management (which is a bit of a catch-all category) and Public Sector.

We want to look in more detail at specific subject areas and have chosen Operational Research as that is where we have expertise. Note that some, more mathematical, OR groups were submitted to the Statistics and OR Panel so their contributions are not included here. Table 13 shows all OR journals ranked in terms of the reconstructed RAE grade and then the number of entries. Those with a "#N/A" in the ABS Grade column were not classified in the ABS list but we have added them in as they all would be considered as OR journals. There are seven 4* journals (all comments in this section refer to grades reconstructed by the model) although all but *Management Science* have small numbers and three do not appear in ABS. Some of these are likely to be due to the small sample, but *Decision Sciences, J. of Heuristics* and the *SIAM* journal are generally considered to be strong. In the 3* journals comes *EJOR*, with the second largest entry, and a wide range of other journals, many with small entries. It is interesting that *Operations Research*, the other top US journal, has virtually all its entries graded as 3* rather than 4*.

The next point of interest is the ranking of *J. of the Operational Research Society (JORS*) as 2*. This has the largest entry (third highest in the whole RAE) and, together with *EJOR*, is the main publication outlet for UK academics who find it hard to publish in the US journals. *JORS* figured highly in the Geary analysis of the 2001 RAE because of their methodology which rated journals in terms of the departments which published in them. It happens that the largest groups of OR academics are at Lancaster and Warwick which were top rated in 2001

and so JORS secured a high grade. However, this was unrealistic in terms of the journal's world rating (Mingers & Harzing, 2007) as the current result shows. It is a surprise, however, that it has gained no 4* work at all in our reconstruction. One other anomoly is Naval Research Logistics which is one of the original OR journals and did have a strong reputation. However, in recent years its impact factor has fallen considerably, and in the recent COPIOR OR journal list it was only graded as 2*.

Table 13 Journals in OR/Management Science ranked by RAE grade and number of entries

5. Discussion

In this section we wish to reflect on the extent to which our analyses can shed light on some of the theoretical debates surrounding journal rankings, the effects of the RAE on research, and peer review.

Journal rankings are a cause of great debate and controversy. Several issues emerge: the extent to which the quality of the journal can stand as a proxy for the quality of a paper published within it - "don't judge a paper by its journal"; the extent to which the quality of a journal can in any case be captured by a single ranking; and the extent to which departments and schools are using journal rankings inappropriately to make promotion and employment decisions, and to limit the types of journals in which their faculty can publish.

Many academic decisions, at all levels, depend on a judgement of the quality of a paper or other research output but this is far from straightforward. In principle, one might say that it should be done by competent experts in the subject, i.e., peer review, but even this approach is very fallible. There have been many studies of the process of reviewing papers for journals (Starbuck (2003; 2005) provides a good overview) and the general conclusions are that there is a high level of disagreement between reviewers and many biases are prevalent. In a classic experiment, Peters and Ceci resubmitted 12 papers already published in high quality journals to the journals they were published in under assumed author names. Only three of the twelve were spotted, and eight of the nine were subsequently rejected.

Problems with obtaining valid peer judgements leads decision-makers to fall back on the quality of the journal as a proxy for the quality of the paper. This relies on two assumptions: that journals can be effectively ranked and that all papers within a journal are of the same quality. Considering the latter issue, it is commonly felt that there is not a uniform quality within journals. Certainly, as mentioned above, the RAE Panel took the view that there was not, and claimed strongly not to have based their judgements on journal rankings. The question has been approached theoretically by Starbuck (2005) who developed a statistical model of the journal review process which showed that, based on divergences of opinion in referees' reports, significant proportions of high quality work would be rejected from top journals and trickle-down into lower tier ones, and equally, less good work would be

accepted by top tier ones. The results deduced in this paper also show a degree of dispersion of grades across a journal as reported in section 4.5 but the results are equivocal. 62% of the journals were allocated to only one grade, thus 38% were given a range of quality levels. Moreover, as Table 12 shows, there were some ABS 4* journals that had a range of quality levels. Moreover, the dispersion on some journals was very wide encompassing both 4* and 1*.

The second issue to be considered is the extent to which the RAE, and and disciplinary processes such as journal ranking lists more generally, are having pernicious effects on the quality and direction of academic research (Nkomo, 2009; Suchan, 2008). Such effects include: pressure on academics to publish only in a narrow range of top-ranked journals (Truex III, Cuellar, & Takeda, 2008); pressure to concentrate on papers at the expense of other forms of output and a corresponding need to produce bite-sized pieces of research rather than more major and significant contributions (de Rond & Miller, 2005); pressure to avoid the more marginal areas of research or interdisciplinary research which are likely to be rated less highly. In turn, these pressures make it very difficult for new subjects (Stewart, 2005) or areas of research (Ozbilgin, 2009) and new journals to become established; lead to much game-playing in the publication process (Dulek, 2008; Macdonald & Kam, 2007); reduce the level of external engagement and impact – rigour at the expense of relevance (Syed, Mingers, & Murray, 2009); and generate a self-perpetuating hegemony of theories, methodologies, journals and institutions (Adler & Harzing, 2009).

There is a degree of empirical support for some of these arguments. Moed (2008) analysed publications patterns in UK science over 20 years, covering three RAEs (1992, 1996 and 2001), which showed changes in response to changing RAE requirements – increased quantity for 1992, increased quality for 1996, and a greater degree of collaboration in 2001. In terms of this RAE we can see that journal papers have come to dominate the submissions (Tables 1,2), from 69% in 1996 to 92% in 2008 with authored books and book chapters falling dramatically. We can also see (Table 7) that non-journal outputs (and particularly external reports) were not ranked highly by the Panel despite the general RAE guidelines which stated that all forms of output should be treated equally.

In terms of the pressures for journal selectivity there is mixed evidence. On the one hand, a greater range of journals were submitted (Table 1) and half of them were not included in the ABS list (Figure 2), but we cannot tell from the data whether this was a matter of necessity, because they were the only publications that were available, or genuine choice. I have looked in detail at the submissions of four top-30 Schools and they all had very low proportions (< 10%) of non-ABS journals. On the other hand, there was clear evidence of selectivity in submission with a low proportion of ABS 1* journals being included (Table 9), and there is possible evidence that journals included in ABS were rated more highly by the Panel. Figure 3 shows an association between the proportion of a School's submission that was in ABS and their overall GPA, and Table 9 shows that non-ABS journals were graded lower with a significant proportion being judged as 0*.

The final area to be discussed is the question of the peer review process itself. Again, this is the subject of strongly held opinions with some arguing that its inherent subjectivity and openness to bias means that more objective, bibliometric measures should be used (Doyle & Arthurs, 1995; Doyle, Arthurs, McAulay, & Osbourne, 1996) whilst others claim it is the only way of getting a fair and rounded evaluation (Jones, Brinn, & Pendlebury, 1996a, 1996b). As mentioned above, there have been studies of peer review in journal refereeing processes, and also in the awarding of grants (Marsh, Jayasinghe, & Bond, 2008) but nothing specifically on the RAE peer review process. Thomas and Watkins (1998) list the main problems with peer review in this context as: the influence of the author's reputation on rating; rating journals that one publishes in oneself more highly; having only a limited range of expertise and down-grading work in other areas; and being overly influenced by input factors such as grants. To this could be added being affected (positively or negatively) by: the institution (the Russell Group effect), the journal, personal knowledge of the author, or the theoretical or methodological nature of the research.

We are not suggesting that such biases are conscious or intentional and we have to recognise that the context necessarily involves conflicts of interest (Hackett, 1997; Scott, 2007), in this case between the beliefs of the individual Panel member, the RAE processes and the Panel as a whole, and loyalties to institutions and disciplines. Research suggests that experts often believe that they are behaving rationally and fairly even within a situation of conflicting interests when in fact they are not. Moore, Tetlock et al (Bazerman, Moore, Tetlock, & Tanlu, 2006; Moore, Tetlock, Tanlu, & Bazerman, 2006) have investigated how the auditors in major scandals such as Enron appeared unaware of the extent to which they had become compromised. "Psychological research on the impact of motivated reasoning and self-serving biases questions the validity of this assumption {that an auditor is either independent or in collusion}. This evidence suggests that intentional corruption is probably the exception and that unconscious bias is far more pervasive." (Moore, et al., 2006, p. 16). They suggests mechanisms such as selective perception, escalation of commitment, inaccuracies of self-perception and the effects of accountability that underlie this motivated reasoning.

Research into questions of conscious or unconscious bias could be of two types –direct observation and analysis of the RAE processes themsleves, which seems infeasible given the secrecy imposed by HEFCE, or analysis of the results in comparison with the Panel membership. The reconstruction of journal grades here published could be used as part of such an analysis.

6. Conclusions

The 2008 RAE has been a huge exercise in peer review and the judgements that were made would have been extremely valuable in addressing some of the issues that surround the whole idea of journal rankings. Unfortunately, the gradings of individual outputs have been kept secret which was, in our view, both unnecessary and undesirable. What we have attempted to do in this paper is to reconstruct the judgements made by the RAE Panel at least at the level of individual journals although not at the level of papers. We have done this by developing a

mathematical programming model that determines the best grade profiles to match the overall institutional profiles for all journals submitted that had at least three entries.

We have shown, both in terms of internal reliability and in terms of correspondence with existing ranking lists such as the ABS list, that the results we have generated have a high degree of plausibility. It is extremely unlikely that they do not represent to a reasonable degree the actual judgements made by the Panel although clearly we can never actually assess the extent of the residual error.

With these results, we have been able to comment on several issues that have arisen concerning the conduct and effects of the RAE, as well as produce an RAE-based ranking for around 700 journals in Business and Management and related areas. Many of these journals are not included in the ABS list. Care should be taken in interpreting the results, especially for journals that had few entries.

- Comparing the grades given by the RAE with those in ABS, on those journals that are in common the overall results are very similar in terms of the average grade awarded. However, there are differences in the proportions of each grade with the reconstreucted RAE giving more 4* and 1*, and there are differences for particular journals with some being two or even three grades apart.

- In terms of the RAE leading to selectivity, there is evidence in both directions. There were a very wide range of journals submitted, many of them not in ABS, but many of these non-management journals were given a low rank. It is clear that there was selectivity in the submissions with relatively few ABS 1* journals being submitted. There is also a clear association between the GPA awarded to an institution and the proportion of its submission that was in ABS journals although the direction and nature of the causality is unclear.

- The RAE Panel was clear that it was not grading papers on the basis of the journal they were published in. There is evidence that supports that since many journals, even top ones, had a degree of dispersion in their gradings. However, our results also produced 62% of journals with 100% in a single grade suggesting a considerable degree of uniformity in judgement.

- As expected, there were significant differences in the gradings given to different sectors with Psychology, Accounting and Finance, Management, and OR doing well and Management Development, Innovation, and Tourism faring worst.

- For non-journal outputs, our results show that books (GPA 2.4) and book chapters (GPA 2.2) gained grades that were commensurate with journals, but reports (GPA 1.4) and other forms of output (GPA 1.3) were seen as poor. This does not bode well for the REF and its focus on external research impact.

In terms of wider significance, this research demonstrates an entirely new method for estimating underlying journal rankings from a set of peer reviewed data such as that produced

by the RAE. To our knowledge the only similar attempt was by Geary at al (2004) which used a fairly crude form of averaging. This represents a novel application of LP.

Table 1 Submission statistics for the last three RAEs

Adapted from Geary et al (2004), Bence and Oppenheim (2004), RAE (2009a)

	1996	2001	2008
No of submissions	100	97	90
No. of staff submitted	2300+	3000+	3300
Total no. of outputs	8000+	9942	12575
No. of journal papers (% of total)[a]	5494 (69%)	7973 (80%)	11625 (92%)
No. of journal titles	1275	1582	1639
Mean outputs/journal	4.3	5.0	7.1
Mean outputs/institution	80.0	102.5	139.7

[a] Totals differ slightly between different sources. Figures for 2008 are after data cleaning as described later

Table 2 Number of publications by output type
Adapted from Geary et al (2004), Bence and Oppenheim (2004), RAE (2009a). Categories with zero entries have been suppressed

Output Type	Description	1996	2001	2008
A	Authored book		431	285
B	Edited book		77	60
C	Chapter in book		863	332
D	Journal article		7973	11374
E	Conference contribution		295	85
G	Software		3	1
H	Internet publication		24	318
N	Research report for external body		80	98
T	Other form of assessable output		184	22
	Total		9942	12575

Table 3 Dispersion of submitted journal outputs
Adapted from Geary et al (2004), Bence and Oppenheim (2004), RAE (2009a)

No. of entries per journal title	1996 % titles	2001 % titles	2008 % titles
1	51.0%	49.0%	42.9%
2	15.5%	16.0%	14.6%
3	8.5%	7.0%	8.7%
4	5.0%	5.0%	4.9%
5-10	11.5%	12.0%	14.0%
11-25	5.5%	7.0%	8.5%
> 25	3.0%	4.0%	6.3%

Table 4 Twenty journals most frequently submitted in the 2008Business and Management RAE Adapted from Geary et al (2004)

2008 Rank	2001 Rank	Journal Name	2008 Total	2001 Total	ABS Grade	ABS subject area
1	2	Journal of Management Studies	219	116	4	GEN MAN
2	8	Human Relations	171	78	4	ORG STUD
3	3	Journal of the Operational Research Society	153	113	3	OR&MANSCI
4	5	European Journal of Marketing	146	90	3	MKT
5	10	Organization Studies	144	75	4	ORG STUD
6	14	European Journal of Operational Research	137	61	3	OR&MANSCI
7	6	International Journal of Operations and Production Management	134	85	3	OPS&TECH
8	9	International Journal of Human Resource Management	133	76	3	HRM&EMP
9	1	Journal of Marketing Management	125	127	3	FINANCE
9	11	Journal of Business Finance and Accounting	125	65	3	MKT
11	4	British Journal of Management	108	107	4	GEN MAN
12	13	Work, Employment and Society	103	64	4	HRM&EMP
13	7	British Journal of Industrial Relations	99	84	4	HRM&EMP
14		Regional Studies	97		4	SOC SCI
15		Research Policy	95		4	SOC SCI
16	16	Service Industries Journal	92	59	2	MKT
17		Critical Perspectives on Accounting	89		3	ACCOUNT
18	20	Organization	83	55	3	ORG STUD
19		Accounting, Organisations and Society	82		4	ACCOUNT
20		Journal of Business Ethics	81		3	ETH-GOV
20		Journal of Business Research	81		3	GEN MAN

Table 5Ranking of journals reconstructed from the RAE (available from http://www.kent.ac.uk/kbs/community/staff/profiles/mingers_john_bio.html)

Table6 Correlations between reconstructed RAE grade and journal rankings (no. of observations)

	ABS 2009	Kent 2007	Geary median	ABS 1* or 4*
RAE grade	0.42 (574)	0.37 (575)	0.42 (416)	
ABS 2009		.69 (574)	0.48 (394)	
Kent 2007			0.49 (394)	
Geary median				
RAE grade 1* or 4*				.61 (183)

Table 7 Reconstructed grades for non-journal outputs

Type	Modal grade	Mean grade	No of outputs	No. of institutions
Authored book	2	2.437	285	68
Book chapter	2	2.21	333	65
Edited book	2	2.10	61	28
Other	1	1.41	169	47
External report	1	1.13	102	35

Table 8 Top 30 journals from RAE (>=50% 4 and >= 12 entries)*

Journal Title	4.0	3.0	Submissions	ABS 2009 1* to 4*	Kent 2007 1* to 4*	5-yr CIF 2* to 4*	ABS sector	Geary Median	FT Top 40 y/n?
Management International Review	100.0	0.0	27	3	2		IB&AREA	6	y
Journal of the Academy of Marketing Science	100.0	0.0	18	4	3	3	MKT	6	
Accounting, Organisations and Society	100.0	0.0	82	4	4	2	ACCOUNT	6	y
Harvard Business Review	100.0	0.0	19	3	3	3	GEN MAN	7	y
Journal of Applied Psychology	100.0	0.0	33	4	4	3	PSYCH	6.5	y
Abacus	100.0	0.0	25	2	2		ACCOUNT	5	
Journal of Marketing Research	100.0	0.0	17	4	4	4	MKT	7	y
Journal of Financial Economics	100.0	0.0	39	4	4	4	FINANCE	7	y
Academy of Management Review	100.0	0.0	30	4	2	4	GEN MAN	6	y
Fiscal Studies	100.0	0.0	13	2	1	2	ECON	5	
Review of Financial Studies	100.0	0.0	27	4	4	4	FINANCE	7	y
International Journal of Industrial Organization	100.0	0.0	14	4	3	2	ECON	6	
American Economic Review	100.0	0.0	17	4	4	4	ECON	6.5	y
Information Systems Journal	100.0	0.0	18	3	3	3	INFO MAN	6	
Organization Science	100.0	0.0	28	4	4	4	ORG STUD	6	y
MIT Sloan Management Review	100.0	0.0	23	3	3	2	GEN MAN	7	y
Marketing Science	98.2	0.0	12	4	4	4	MKT	7	y
Communications of the ACM	91.8	5.2	13	3	4	3	INFO MAN	6	
Financial Accountability and Management	87.5	4.3	30	3	1		FINANCE	5	
European Journal of Information Systems	85.7	13.6	49	3	2	2	INFO MAN	5	
International Journal of Logistics	82.4	15.3	15	2	2		OPS & TECH	5	

Management									
Journal of Business Venturing	79.3	17.2	32	4	3	4	ENT-SMBUS	6	y
New Technology, Work and Employment	77.4	17.3	35	3	2	3	HRM&EMP	5	
Management Science	69.9	29.0	29	4	4	4	OR&MANSCI	6	y
International Journal of Management Reviews	62.2	36.5	37	3	2	3	GEN MAN	6	
Journal of Econometrics	62.0	36.5	26	4	3	4	ECON	6	
Industrial and Corporate Change	60.4	10.6	30	3	2	3	SOC SCI	6	
British Journal of Industrial Relations	60.0	36.4	99	4	3	4	HRM&EMP	5	
Journal of Empirical Finance	58.5	3.6	14	3	3		FINANCE	6.5	
Accounting and Business Research	57.4	39.0	52	3	2		ACCOUNT	6	

Table 9 Proportions of journals in particular ranks comparing ABS with RAE grades
Note: we show the proportions in terms of % for ease of comparison but all Chi-Square tests were performed on the underlying frequencies

	Grades given in ABS			**RAE Estimated Grades**		
	All journals	**Journals not in RAE**	**Journals in RAE and our list**	**All our list**	**Journals not in ABS**	**Journals in ABS and our list**
4*	10%	4%	15%	18%	13%	19%
3*	24%	12%	31%	28%	19%	30%
2*	37%	39%	37%	28%	25%	28%
1*	28%	45%	17%	22%	29%	21%
0*				3%	13%	2%
GPA	2.17	1.74	2.43	2.34	1.9	2.42

Table10 Differences between RAE 2008 and ABS

RAE 2008 Grade	**ABS 2009 grade minus RAE 2008 Grade (absolute value)**			
	0	**1**	**2**	**3**
0*		2	9	
1*	51	51	18	2
2*	69	73	14	
3*	70	86	16	
4*	31	38	35	3
Grand Total	226	248	93	7

Table 11 Distribution of dispersion of journal grading

% outside the modal grade	Frequency
0	432
<=10	56
10<=20	43
20<=30	55
30<=40	46
40<=50	55
>50	8

Table 12 ABS 4 journals with a high degree of dispersion*

Journal Title	4.0	3.0	2.0	1.0	0.0	RAE grade	% dispersion	Entries
Organizational Research Methods	0.0	43.4	25.9	12.4	18.3	3	56.6	13
Journal of Organisational Behaviour	33.7	22.7	43.7	0.0	0.0	2	56.3	26
Industrial Relations: A Journal of Economy and Society	48.0	0.0	16.0	36.0	0.0	4	52.0	25
Organization Studies	48.8	50.4	0.8	0.0	0.0	3	49.6	144
Business History	0.0	51.4	48.6	0.0	0.0	3	48.6	37
Journal of Business	46.5	51.4	2.0	0.0	0.0	3	48.5	16
Transportation Research Part B: Methodological	0.00	51.96	48.04	0.00	0.00	3	48.0	5
Entrepreneurship, Theory and Practice	12.4	52.0	35.5	0.0	0.0	3	48.0	28
Journal of Marketing	44.9	53.8	1.3	0.0	0.0	3	46.2	18
Journal of Finance	54.2	45.0	0.8	0.0	0.0	4	45.8	49
Administrative Science Quarterly	0.00	55.92	44.08	0.00	0.00	3	44.2	6
Public Administration Review	38.1	56.5	5.4	0.0	0.0	3	43.5	17
Organizational Behaviour and Human Decision Processes	0.00	42.18	57.82	0.00	0.00	2	42.1	8
British Journal of Social Psychology	58.08	0.00	0.00	41.92	0.00	4	41.9	4
Journal of Economic Literature	32.32	8.57	59.10	0.00	0.00	2	40.8	4
British Journal of Industrial Relations	60.0	36.4	3.6	0.0	0.0	4	40.0	99
Journal of Econometrics	62.0	36.5	1.5	0.0	0.0	4	38.0	26
Annals of Tourism Research	0.0	63.0	37.0	0.0	0.0	3	36.9	42
Environment and Planning A	0.0	66.8	33.2	0.0	0.0	3	33.2	46
Journal of Operations Management	30.7	68.0	1.3	0.0	0.0	3	32.0	31
Work, Employment and Society	0.0	68.3	31.7	0.0	0.0	3	31.7	103
Journal of Economic Geography	0.00	68.72	31.27	0.00	0.00	3	31.2	10
Management Science	69.9	29.0	1.1	0.0	0.0	4	30.1	29
Regional Studies	0.0	29.0	71.0	0.0	0.0	2	29.0	97

Table 13 Journals in OR/Management Science ranked by reconstructed RAE grade and number of entries

Journal Title	RAE Grade	Entries	4.0	3.0	2.0	1.0	0.0	ABS 2009 Grade
Management Science	4	29	69.9	29.0	1.1	0.0	0.0	4
Group Decision and Negotiation	4	4	100.00	0.00	0.00	0.00	0.00	3
Operations Research Letters	4	4	100.00	0.00	0.00	0.00	0.00	2
Computational Optimization and Applications	4	3	100.00	0.00	0.00	0.00	0.00	#N/A
SIAM Journal on Optimization	4	3	100.00	0.00	0.00	0.00	0.00	#N/A
Journal of Heuristics	4	3	100.00	0.00	0.00	0.00	0.00	#N/A
Decision Sciences	4	3	100.00	0.00	0.00	0.00	0.00	3
European Journal of Operational Research	3	137	0.0	98.2	1.8	0.0	0.0	3
Computers and Operations Research	3	20	0.0	67.9	32.1	0.0	0.0	2
Operations Research	3	20	0.0	98.9	1.1	0.0	0.0	4
Journal of the Royal Statistical Society Series A (Statistics in Society)	3	9	26.24	44.89	28.87	0.00	0.00	3
Theory and Decision	3	9	0.00	80.01	19.99	0.00	0.00	2
IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans	3	7	0.00	84.46	15.54	0.00	0.00	#N/A
Mathematical Programming	3	7	0.00	98.34	1.66	0.00	0.00	3
Advances in Applied Probability	3	5	0.00	81.04	18.96	0.00	0.00	3
Annals of Operations Research	3	5	0.00	96.26	3.74	0.00	0.00	2
Journal of Combinatorial Optimization	3	4	0.00	96.56	3.44	0.00	0.00	#N/A
Computational Statistics & Data Analysis	3	3	0.00	82.06	17.94	0.00	0.00	#N/A
Discrete Applied Mathematics	3	3	0.00	96.10	3.90	0.00	0.00	#N/A
Applied Mathematics and Computation	3	3	0.00	59.35	40.65	0.00	0.00	#N/A
Interfaces	3	3	0.00	86.26	13.74	0.00	0.00	2
Journal of the Operational Research Society	2	152	0.0	37.0	63.0	0.0	0.0	3
Omega: The International Journal of Management Science	2	37	0.0	47.0	52.0	0.0	1.0	3
International Journal of Forecasting	2	23	0.0	8.3	91.7	0.0	0.0	3

Journal of Forecasting	2	14	0.0	5.9	94.1	0.0	0.0	3
Systems Research and Behavioral Science	2	11	0.00	4.98	95.02	0.00	0.00	2
International Transactions in Operational Research	2	6	0.00	5.63	82.92	11.44	0.00	2
Journal of Optimization Theory and Applications	2	5	42.90	4.36	52.74	0.00	0.00	#N/A
OR Insight	1	4	0.00	1.70	22.06	76.24	0.00	#N/A
Simulation Modelling Practice and Theory	1	3	0.00	0.00	0.00	100.00	0.00	#N/A
IMA Journal of Management Mathematics	0	7	0.00	0.00	0.00	36.56	63.44	2
Naval Research Logistics	0	6	0.00	4.27	13.15	17.17	65.40	3

References

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

Bazerman, M. H., Moore, D. A., Tetlock, P. E., & Tanlu, L. (2006). Reports of solving the conflicts of interests in auditing are highly exaggerated. *Academy of Management Review, 31*, 43-49.

Bence, V., & Oppenheim, C. (2004). A comparison of journal submissions to the UK's Research Assessment Exercises 1996 and 2001 for UoA 43 (Business and Management Studies). *European Management Journal, 22*, 402-417.

Bessant, J., Birley, S., Cooper, C., Dawson, S., Gennard, J., Gardiner, M., Gray, A., Jones, P., Mayer, C., McGee, J., Pidd, M., Rowley, G., Saunders, J., & Stark, A. (2003). The state of the field in UK management research: Reflections of the Research Assessment (RAE) Panel. *British J. of Management, 14*, 51-68.

Clark, T., & Wright, M. (2007). Reviewing journal rankings and revisiting peer reviews: Editorial perspectives. *Journal of Management Studies, 44*, 612-621.

de Rond, M., & Miller, A. N. (2005). Publish or perish: Bane or boon of academic life? *Journal Of Management Inquiry, 14*, 321-329.

Doyle, J., & Arthurs, A. (1995). Judging the quality of research in business schools. *Omega, Int. Journal Mgmt. Sci, 23*, 257-270.

Doyle, J., Arthurs, A., McAulay, L., & Osbourne, P. (1996). Citation as effortful voting: a reply to Jones, Brinn and Pendlebury. *Omega, 24*, 603-606.

Dulek, R. (2008). Academic research: Two things that get my goat--and three that offer meaning. *Journal of Business Communication, 45*, 333-348.

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.

Hackett, E. (1997). Peer review in science and science policy. In M. Frankel & J. Cave (Eds.), *Evaluating Science and Scientists* (pp. 51-60). Budapest/Oxford: Central European University Press/OUP.

HEFCE. (2009). *Research Excellence Framework: Second consultation on the assessment and funding of research.* http://www.hefce.ac.uk/pubs/hefce/2009/09_38/09_38.pdf.

Jones, M., Brinn, T., & Pendlebury, M. (1996a). Journal evaluation methodologies: a balanced response. *Omega, 24*, 607-612.

Jones, M., Brinn, T., & Pendlebury, M. (1996b). Judging the quality of research in business schools: a comment from accounting. *Omega, Int. Journal Mgmt. Sci., 24*, 597-602.

Macdonald, S., & Kam, J. (2007). Ring a ring o'roses: quality journals and gamesmanship in Management Studies. *Journal of Management Studies, 44*, 640-655.

Marsh, H. W., Jayasinghe, U. W., & Bond, N. W. (2008). Improving the peer-review process for grant applications: Reliability, validity, bias, and generalizability. *American Psychologist, 63*, 160-168.

Mingers, J., & Harzing, A.-W. (2007). Ranking journals in business and management: a statistical analysis of the Harzing dataset. *European Journal of Information Systems, 16*, 303-316.

Moed, H. (2008). UK Research Assessment Exercises: Informed judgements on research quality or quantity? *Scientometrics, 74*, 153-161.
Moore, D. A., Tetlock, P. E., Tanlu, L., & Bazerman, M. H. (2006). Conflicts of interest and the case of auditor independence: Moral seduction and strategic issue cycling. *Academy of Management Review, 31*, 10-29.
Nkomo, S. (2009). The seductive power of academic journal rankings. *Academy of Management Learning and Education, 8*, 106-112.
Ozbilgin, M. (2009). From journal rankings to making sense of the world. *Academy of Management Learning and Education, 8*, 113-121.
Paul, R. (2007). Challenges to information systems: time to change. *European J. Information Systems, 16*, 193-195.
Paul, R. (2008). Measuring research quality: The united Kingdom Government's Research Assessment Exercise. *European J. Information Systems, 17*, 324-329.
Peters, D., & Ceci, S. (1982). Peer-review practices of psychological journals: The fate of published articles, submitted again. . *Behavioral and Brain Sciences, 5*, 187-255.
RAE. (2004). *RAE 2008 : Initial decisions by the UK funding bodies.* HEFCE, http://www.rae.ac.uk/pubs/2004/01/.
RAE. (2005). *Guidance on submissions.* HEFCE, http://www.rae.ac.uk/pubs/2005/03/.
RAE. (2006). *RAE 2008 Panel criteria and working methods.* Report no. 01/2006, HEFCE, http://www.rae.ac.uk/pubs/2006/01/.
RAE. (2008). *RAE 2008: the outcome.* Report no. 01/2008, HEFCE, http://www.rae.ac.uk/pubs/2008/01/.
RAE. (2009a). *RAE2008 subject overview reports: I 36 Business and Management studies.* HEFCE, http://www.rae.ac.uk/pubs/2009/ov/.
RAE. (2009b). *RAE 2008 submissions: UOA 36 Business and Management Studies.* HEFCE, http://www.rae.ac.uk/submissions/submissions.aspx?id=36&type=uoa.
RAE. (2009c). *UoA 36 Business and Management Studies: Quality Profiles.* HEFCE, http://www.rae.ac.uk/pubs/2009/pro/uoas/uoa%2036%20-%20all%20submissions.pdf.
Scott, A. (2007). Peer review and the relevance of science. *Futures, 39*, 827-845.
Select Committee on Science and Technology. (2004). *Tenth Report, Section 9: Integrity of the publishing process.* Parliament, London, http://www.publications.parliament.uk/pa/cm200304/cmselect/cmsctech/399/39912.htm.
Starbuck, W. (2003). Turning Lemons into Lemonade: Where Is the Value in Peer Reviews? *Journal Of Management Inquiry, 12*, 344-351.
Starbuck, W. H. (2005). How Much Better Are the Most-Prestigious Journals? The Statistics of Academic Publication. *Organization Science, 16*, 180-200.
Stewart, J. (2005). The UK research assessment exercise. *Human Resource Development International, 8*, 489-494.
Suchan, J. (2008). How academic organizational systems and culture undermine scholarship and quality research. *Journal of Business Communication, 45*, 349-356.
Syed, J., Mingers, J., & Murray, P. (2009). Beyond Rigour and Relevance: A Critical Realist Approach to Business Education. *Management Learning, 41*, 71-85.
Thomas, P., & Watkins, D. (1998). Institutional research rankings via bibliometric analysis and direct peer review: A comparative case study with policy implications. *Scientometrics, 41*, 335-355.
Truex III, D., Cuellar, M., & Takeda, H. (2008). Assessing scholarly influence: Proposing new metrics. Proceedings of the ICIS 2008. Paris: Association for Information Systems.