A critical review of the application of citation studies to the Research Assessment Exercises

Julian Warner

The Queen's University of Belfast, Northern Ireland, UK

Received 5 May 2000
Revised 7 July 2000

Abstract.

The Research Assessment Exercises (RAE) conducted in the UK have attractedvarious types of published response. These include citation analyses and a review of the public reception of the RAE 1996, which included a brief critique of the citation studies.

This paper develops the critique. Largely unexplored issues in the theory or assumptions of bibliometrics, e.g. the level of citation which corresponds to a quantum of research quality, are found to emerge in the studies. A weak, and unsatisfactorily treated, correlation between citation aggregates and research quality for individual entities is revealed. The proposal to replace informed peer review by citation analysis is regarded as highly unrealistic.

Productive uses for citation analysis in research evaluation are suggested. A historically rare instance of correlation between rankings derived from citation aggregates and from real world peer review has been revealed by the studies. The future value of citation analysis could be to inform, but not to determine, judgements of research quality.

A combination of methods is advocated for future studies of the RAEs. Information science must attend to considerations of value, as well as using established techniques, if it is to avoid marginalisation.

Correspondence to: J. Warner, School of Management and Economics, The Queen's University of Belfast, Belfast BTY 1NN, Northern Ireland, UK. Tel: +44 2890 245133. Fax: +44 2890 328649. E-mail: j.warner@qub.ac.uk

Introduction

Research Assessment Exercises (RAE) were conducted in the UK in 1992 and 1996 to determine the ratios with which public funds were to be distributed to universities, on the basis of grades obtained by their constituent departments, to support research infrastructure. Other geopolitical regions would have comparable exercises, although with different approaches to research evaluation [1]. Initial information about specific criteria and procedures used in the RAE could be obtained from the websites of funding councils [2]. The significance of the grades obtained, and the cogency of the issues in academic consciousness, led to a variety of published responses. Many focused on the discipline of assessment [3, 4]. A few had broader concerns: within information science, citation analyses, arguing that this technique should replace established assessment procedures [5–7], surveys by the funding councils [8, 9] and a review of the public reception of the RAE 1996 [10].

In the review of the reception of the RAE 1996, a brief critique of the citation studies is presented. The information science literature had:

> . . . taken the research assessment exercises as a given and concentrated on adapting established [bibliometric] techniques, particularly, although not exclusively, citation analysis, to the evaluation of research. The use of these techniques in assessment is then advocated as more economic than the procedures which have been used in the research assessment exercises.
>
> A broader responsibility to inform discussion of the research assessment exercises and their effects from understandings of scholarly communication and of the development of knowledge . . . [had] not yet been accepted [10].

The aim here is to develop the critique of the citation studies. A degree of complementarity may be revealed between apparently antithetical approaches.

Review of citation studies

Some crucial issues have remained largely unexplored in the theory or, perhaps more accurately, the assumptions of bibliometrics:

(1) how a single scholar or author is to be understood and differentiated;
(2) what level or number of citations constitutes a quantum of research impact or quality;
(3) whether research quality as indicated by the consensus of a scholarly community can be differentiated from quality as perceived by specific members or groups within that community; and
(4) whether citation counts provide objective measures or intersubjective indicators of research quality.

More familiarly, there has been an equivocation on the distinction between research impact and quality. Macro- and micro-correlations between citation aggregates and research impact or quality have been distinguished, with reservations placed on the strength of the correlation for individual scholars. These issues emerge, rather unreflexively, in the studies reviewed here, despite their highly empirical focus and pragmatic approach.

A distinction between levels of description and of value needs to be introduced. Some cultures have not separated description from value: for instance, in chivalric language, *knight* is simultaneously descriptive and evaluative. Post-feudal or modern cultures have distinguished, although not consistently or entirely, description from value [11]: *capitalism* can be received as a description of a set of social relations, which can then be evaluated either positively or negatively or in a more complex combination. Within information science, confusion of description and of value, e.g. characterising apparently unsystematic information seeking practices as amateurish, has inhibited the development of understanding [12, 13]. In this context, description is understood as the correlations established and value as the uses made of those correlations.

Understanding of an author

The concept of an individual scholar or author is introduced in the citation studies, but in an atheoretical manner:

> Another problem arose because, in a number of cases . . . other authors with exactly the same surnames and initials in quite different disciplines were being cited . . . Where articles cited were clearly outside the field of librarianship or information science, the citations were ignored . . . perhaps the count for that author was inaccurate . . . the cited articles rejected, even if they were written by the library or information science academic, would not have been put forward as part of the RAE as those articles are almost certainly not in the field . . . In some cases, there is genuine ambiguity [7].

This corresponds to the common modern understanding of an author as a field of conceptual coherence, detected by Foucault [14]. Analogously, Fleck distinguished the member of a thought community from the personal subject, noting that a single personal subject could be a member of different thought communities and that contradictions between the beliefs of thought communities might not rise to the level of individual consciousness in such subjects [15, 16]. From this perspective, the S.C. Bradford who wrote *Documentation* [17] would be differentiated as an author from the S.C. Bradford who wrote *Romance of Roses* [18] and the Warren Weaver who compiled *Alice in Many Tongues: The Translations of Alice in Wonderland* [19] from the Warren Weaver who collaborated with Claude Shannon for *The Mathematical Theory of Communication* [20]. In this instance, there is a congruence between the ordinary discourse understanding of an individual author reproduced in the citation studies and the more theoretically informed perspectives, themselves derived from observations of ordinary understandings. Reference to the work of Foucault and Fleck (widely recognised as significant) might have resulted in a more sophisticated discussion. For instance, it could have been acknowledged that the most problematic disambiguation decisions as to what constitutes a single cited author arise when different personal subjects publish under similar or identical names and as potentially confusable fields of intellectual coherence.

In the decisions made about what constitutes a single cited author in the first citation study, concerned with library and information science, there is a hidden appeal to contextual knowledge of the field (although some personal subjects would question the splitting of themselves into different authors). The question then arises, but is not addressed, of how these decisions would be made for fields without contextual knowledge. Working with citation data powerfully enforces recognition of its messiness and the difficulty, and potential arbitrariness, of disambiguation decisions [21]. These would be particularly significant for areas with low levels of citation, which would include some reported in the studies. It is asserted that mistaken authorial attributions will have self-cancelling effects, that it is 'plausible' that distortions arising from limitation to first authors will be slight and that any errors introduced by the methods used are therefore 'likely to

be minor' [6]. Here, assertion has replaced demonstration, as no control study was concurrently conducted. The broad congruence of results with an independent study not limited to first authors [5] is later treated as confirmation of these assertions [7]. The issue of the effects of contrasting communication practices of different scholarly fields, with known variations in the incidence of multiple authorship, is not explicitly addressed.

Quanta of research quality

The issue of the number or level of citations which correspond to a quantum of research quality, which must be considered if they are to be regarded as measures, and not just valuable signs, of research quality emerges in the studies. The absence of such a simple correspondence complicates the account of the correlation between citation aggregates and RAE grades. The first study concludes:

. . . the RAE score correlates strongly with the number of citations received and with the number of citations per member of staff [6].

This is inaccurate, as the correlation demonstrated is between rankings derived from citation data and from RAE grades and not directly between the number of citations and the grades themselves. The inaccuracy is more than technical, as no algorithm or procedure is indicated to determine or even to guide the translation of rankings derived from citations into RAE grades. If citation levels did correspond to quanta of research quality, a fully deterministic algorithm would be conceivable. Sceptically, it could be said that the studies depend on a simplifying assumption, whose banality is offensive when made explicit, but evade the reality of that dependence by remaining silent about it. In an unremarked retreat from the impasse created, a later study asserts that citation analysis should be used to suggest the rank ordering of departments, with grades assigned by expert panels [7].

Research impact and quality

Research impact is first nominally differentiated from research quality but then, in a familiar equivocation, the distinction is lost. The view that citation analyses are most useful when contextualised with other indicators is introduced, but, similarly, priority is finally accorded to citation data [6, 7]. The apparent scientific rigour of the studies is, then, partly betrayed by a close reading.

Macro- and micro-levels of correlation

Macro- and micro- levels of correlation can be differentiated. Macro- is here understood as the overall correlation between rankings derived from RAE grades and citation aggregates, and micro- as the correlation for individual entities, usually academic departments, within the Units of Assessment distinguished for the RAE. The studies themselves do not explicitly distinguish macro- from micro-level correlations, but inevitably refer to these different levels. Once isolated, the micro-level of correlation is revealed to be highly unsatisfactorily treated. In the most extensive study, the tabular presentation of results is, curiously and rather unhelpfully, ordered within Units of Assessment by geopolitical region and then alphabetically, rather than by one of the correlating elements [7]. Reordering by total number of citations for departments assessed, the ranking which had the strongest correlation with the order derived from RAE grades reveals that overall correlation conceals gross discrepancies at the departmental level (see Tables 1a–1c). One department that is ranked fourteenth out of eighteen by RAE grades is ranked fourth by total number of citations. Such discrepancies would have major funding implications. The known limited correlation at the micro-level of citation aggregates with peer judgement is confirmed, although no comment is made or even fully implied.

Description and value

Description and value are not adequately distinguished in the citation studies and description is allowed to slide largely unnoticed into value. At the level of description, a strong correlation has convincingly been established at the macro-level, although with substantial and largely unnoticed reservations at the micro-level; at the level of value, no case is made for the substitution of citation analysis for the established research evaluation procedures, other than their lower direct cost. An early study simply asserts the possibility of replacement, while a later study gives less than a page to real-world considerations of replacing fully informed peer review with citation-based procedures. Some crucial assertions in the later article are carefully, if somewhat disingenuously, prefixed by *arguably*. For instance, in an immediate erosion of the position that citation studies should be used to suggest rank ordering, it is asserted:

Arguably, however, the primary function of RAE peer group panel should be to review the citation results, and assign their scores of one to five as appropriate [7].

Arguably, the proposal is nonsense.

Table 1a
Results for Archaeology, ordered by total number of citations

Rank	Higher education institution	Total no. of citations	Average citations per staff member	1992 RAE score
1	Oxford	210	23.30	5
2	Cambridge	176	14.67	5
3	University College	128	4.27	5
4	Sheffield	75	6.82	5
5	Southampton	75	9.90	5
6	Belfast	72	14.40	4
7	Liverpool	66	4.13	3
8	Edinburgh	65	9.29	4
9	Leicester	60	4.00	3
10	Glasgow	60	7.50	3
11	Bradford	40	5.00	3
12	**Reading**	**34**	**6.80**	**5**
13	Birmingham	24	2.40	3
14	Durham	23	1.92	4
15	Exeter	23	4.60	3
16	York	22	2.75	3
17	St David's	21	3.00	3
18	Newcastle	19	2.71	3
19	Manchester	9	2.25	2
20	Nottingham	8	1.00	2
21	Bournemouth	0	0.00	2

Table 1b
Results for Anatomy, ordered by total number of citations

Rank	Higher education institution	Total no. of citations	Average citations per staff member	1992 RAE score
1	Birmingham	1,907	211.89	5
2	Cambridge	1,880	98.94	5
3	University College	1,537	109.78	5
4	**Edinburgh**	**1,119**	**124.33**	**2**
5	Oxford	1,110	100.91	5
6	Glasgow	1,102	91.83	3
7	Nottingham	867	86.70	3
8	Bristol	631	31.55	3
9	Cardiff	527	47.91	3
10	**Liverpool**	**483**	**25.42**	**4**
11	Southampton	411	82.20	2
12	Manchester	340	68.00	3
13	Queen Mary	328	36.44	2
14	Leeds	294	42.00	1
15	Aberdeen	218	36.33	1
16	**Charing Cross**	**166**	**23.66**	**3**
17	Leicester	40	6.67	2
18	Belfast	40	5.00	2

Table 1c
Results for Genetics, ordered by total number of citations

Rank	Higher education institution	Total no. of citations	Average citations per staff member	1992 RAE score
1	Cambridge	1,746	145.50	5
2	Leicester	1,572	131.00	5
3	Glasgow	1,020	78.46	4
4	Leeds	548	42.20	4
5	**Nottingham**	**514**	**73.42**	**5**
6	Liverpool	426	22.42	4
7	UCL	403	57.50	3
8	Queen Mary	29	29.00	3

The proposal is naïve with regard to real-world issues of dispute, possibly of litigation, and the depth of persuasion needed for funding councils and other interested parties to accept this proposal. What, for instance, given the reservations revealed on correlation for individual departments and the known limitations of citation analysis at the micro-level, could be said to a department dissatisfied with outcomes? The prescriptive tone and the failure fully to distinguish description from value could themselves contribute to the sometimes acknowledged poor public image of bibliometrics.

Future development

What, then, can be retrieved from the citation studies? Further evidence of correlation in other subject domains, even when combined with greater technical rigour, would not affect fundamental reservations on the value and realism of proposals to substitute citation-based measures. In this instance, to repeat history would be farce. One productive approach might be to, in a sense, stand existing arguments on their head. A beginning can be made by restoring the distinction between description and value.

At the level of description, more is present than has been sufficiently emphasised. Concerns have been expressed about the relative lack of validating studies for citation analysis [22]. The RAE grades represent a historically rare instance of real-world, and not experimentally elicited or contrived, judgements of research quality, influenced, although not determined [10], by shared definitions and expressed as discrete points on a common scale. The correlation revealed by the existing citation studies can then be regarded as a highly significant validation of citation aggregates as signs or

indicators (not measures) of research quality, while recalling their critical limitations at the micro-level. The largely nominal distinction of impact from quality can then be dissolved and further equivocation avoided. Members of research communities would, however, be abrogating their intellectual responsibility if they simply accepted consensus indicators of research quality, without full direct evaluation. Different aspects of research quality, as established by group consensus and individual scholars, have, then, been differentiated. Citation counts or levels are not regarded as objective measures but intersubjectively validated indicators, amenable to data transformations and forms of graphic representation, difficult to generate from verbal judgements of research quality. Such transformations and representations can assist in giving a concise overview of a field.

At the level of value, citation analysis can then be employed as one element used to inform judgement of research quality, with judgement underdetermined by any single element. For instance, where a significant discrepancy between rankings derived from citations and grades is found, this discrepancy could be investigated, without privileging either indicator as final or objective, but considering the specific circumstances of their production. In the context of a formal research evaluation progress, intended grades could be reviewed; in post-event review, an attempt could be made to understand the sources of contrasts. To give an example of a possible contrast and the different circumstances of production of research judgements and citation aggregates, review articles may not necessarily be regarded as research (and this was the dominant approach of the RAEs) but are known to be highly cited as surrogates for the research they review. Bibliometric indicators other than citation aggregates [23] could also be assimilated to this evaluative model and similarly used to inform judgements. Such a heterogeneous and, from this perspective, balanced approach might reduce hostility to bibliometric indicators.

Further investigative possibilities are opened up for citation analysis, subtler than simple correlation between citation aggregates and RAE grades. One possibility at least is implicit in patterns revealed by the studies. The correlation between rankings derived from sums of citations for entities assessed and from RAE grades is stronger than that derived from mean citations for individuals within those entities: this could imply positive significance for critical mass (an implication supported by discussions of the RAE 1996 [10]) and could be investigated further. Definitions for RAE 1996 grades made precise reference to proportions of individuals at various levels of research attainment; informal communication and public discussions strongly suggested that these were rigidly interpreted by some panels, influenced by their disciplinary culture (e.g. of analytic philosophy), while other panels adapted initial definitions to focus on culture and impression [10]: is this reflected in different extent of deviations from mean number of citations for individuals within the entities assessed by different panels? What is the relative influence of outstanding individuals, both as understood in terms of citations and other forms of recognition, including verbal judgements? The intra-disciplinary status of groups and individuals could be differentiated from extra-disciplinary recognition, by an analysis of the disciplines of citing journals: for instance, information science and, possibly even more, bibliometrics are known to be importers rather than exporters of citations. In these proposals, citation analysis is transformed, at the level of value, from a determining method into a forensic technique.

Discussion and conclusion

Some attention should be given to the implications for information science of the publication of the citation studies in leading journals and their generally uncritical reception. Critiques made in the 1970s and early 1980s, which might have been expected to elicit a stronger counter-response, remain disturbingly relevant. A dominant strain of research in information science was characterised as sophisticated paper pushing, enabled by the computer [24]. Bibliometric studies were regarded as showing more technical than conceptual sophistication [12]. These remarks would be applicable to the studies reviewed and even their technical rigour is questionable. A further, more recent, suggestion has been that a relationship of complicity exists between legitimate uses of citation data and their abuse, despite occasional criticisms of abuse; this would be supported by the elision of critical reservations in the studies reviewed, which have been accepted as legitimate.

The truth is not always exciting and an unexceptionable, although balanced and productive, conclusion is offered here. The various approaches to the RAE, citation analyses, surveys based on questionnaires and focus groups, and studying the reception of an event, are all valuable in themselves, with contrasting methodological reservations and limitations. Why not exploit their potential complementarity in combination? The issues involved in publicly funded research

are significant, if scholarly endeavour is to be regarded as more than a game played for certain exchange-values, a view partly encouraged by the RAE. The review of the reception of the RAE concluded:

> Reflection on contributions does identify a fundamental responsibility for funding councils: to ensure that evaluative methods, while producing meaningful distinctions between units assessed, contribute to the health of the system as a whole. The tenor of the public discourse studied here suggests that this responsibility is unlikely to be assumed by individual institutions. The individual rational economic self-interest of members of a system may not maximize group or wider communal benefit [10].

Fuller understanding of the forces influencing research systems, obtained from whatever methods of investigation, can assist policy formation for communal benefit. In these circumstances, any assistance to understanding is to be welcomed. An ecumenical approach to investigation might also result in a contribution which could be addressed to the wider public, reducing any reciprocated ignorance between the subculture of science and wider society. If information science can contribute a broader theoretical understanding and an awareness of issues of value, as well as established techniques, it might avoid the danger of its own further marginalisation.

References

[1] R.N. Kostoff, *The Handbook of Research Impact Assessment* (DTIC Report ADA296021, 7th ed., 1997). Available at: http://www.dtic.mil/dtic/kostoff/Handweb 71.html

[2] Department of Education Northern Ireland, Higher Education Funding Council for England, Higher Education Funding Council for Wales and Scottish Higher Education Funding Council, *RAE 2001*. Available at: http://www.rae.ac.uk/

[3] G. Whittington, The 1996 Research Assessment Exercise, *British Accounting Review* 29(2) (1997) 181–197.

[4] J. Elkin and D. Law, The 1996 Research Assessment Exercise: the library and information management panel, *Journal of Librarianship and Information Science* 29(3) (1997) 131–141.

[5] L.B. Seng and P. Willett, The citedness of publications by United Kingdom library schools, *Journal of Information Science* 21 (1995) 68–71.

[6] C. Oppenheim, The correlation between citation counts and the 1992 Research Assessment Exercises ratings for British library and information science departments, *Journal of Documentation* 51(1) (1995) 18–27.

[7] C. Oppenheim, The correlation between citation counts and the 1992 research assessment exercise ratings for British research in genetics, anatomy and archaeology, *Journal of Documentation* 53(5) (1997) 477–487.

[8] *The Impact of the 1992 Research Assessment Exercise on Higher Education Institutions in England* (Report Ref: M 6/97) (Higher Education Funding Council for England, 1997).

[9] I. McNay, *The Impact of the 1992 RAE on Institutional and Individual Behaviour in English Higher Education: The Evidence from a Research Project* (Ref: M 6/97) (Higher Education Funding Council for England, Bristol, 1997).

[10] J. Warner, The public reception of the Research Assessment Exercise 1996, *Aslib Proceedings* 49(10) (1997) 263–276. Also in *Information Research* 3(4) (1998), available at: http://www.shef.ac.uk/~is/publications/infres/paper45.html

[11] B. Latour, *We Have Never Been Modern*, translated by C. Porter (Harvard University Press, Cambridge, MA, 1993).

[12] N. Roberts, A search for information man, *Social Science Information Studies* 2(2) (1982) 93–104.

[13] J. Warner, Retrieval performance tests in relation to online bibliographic searching. In: D. Shaw (ed.), *ASIS '92: Celebrating Change: Information Management on the Move* (Proceedings of the 55th ASIS Annual Meeting, Pittsburgh, PA, 26–29 October 1992) (Learned Information for the American Society for Information Science, Medford, NJ, 1992).

[14] M. Foucault, What is an author? In: P. Rabinow (ed.), *The Foucault Reader* (Penguin Books, 1984).

[15] L. Fleck, *Genesis and Development of a Scientific Fact*, translated by F. Bradley and T.J. Treen and edited by T.J. Treen and R.K. Merton (Chicago University Press, Chicago, IL, 1979).

[16] J. Warner, *From Writing to Computers* (Routledge, London and New York, 1994).

[17] S.C. Bradford, *Documentation* (C. Lockwood, London, 1948).

[18] S.C. Bradford, *Romance of Roses* (London, 1946).

[19] W. Weaver, *Alice in Many Tongues: The Translations of Alice in Wonderland* (University of Wisconsin Press, Madison, WI, 1964).

[20] W. Weaver and C.E. Shannon, *The Mathematical Theory of Communication* (University of Illinois Press, Urbana, IL, 1949).

[21] Y. Lindholm-Romantschuk and J. Warner, The role of monographs in scholarly communication: an empirical study of philosophy, sociology and economics, *Journal of Documentation* 52(4) (1996) 389–404.

[22] E. Garfield, From citation indexes to infometrics: is the tail now wagging the dog? *Libri* 48(2) (1998) 67–80.

[23] S. Hodges, B. Hodges, A.J. Meadows, M. Beaulieu and D. Law, The use of an algorithmic approach for the assessment of research quality, *Scientometrics* 35(1) (1996) 3–13.

[24] V. Rosenberg, The scientific premises of information science, *Journal of the American Society for Information Science* 25(4) (1974) 263–269.

Comment

One of the referees of the paper above asked me for a second opinion. While we both agree that the paper should be published and that the author makes some interesting and valid points, I feel that it should not go into print without some comment.

The author has missed a fundamental point. All the research published hitherto has demonstrated that there is a statistically significant correlation (at a very high level of significance) between RAE ranking and citation count. The author points out the very well known problems of citation counting, e.g. the problem of common names, and the problem of using first authors only, but the fact remains that the correlation is there. The author's lengthy philosophical debate on understanding what an author is, and the quoting of Foucault and Fleck, misses the point.

The same point applies elsewhere. For example, the author criticises the argument that distortions arising from first authors will be slight, claiming that 'assertion has replaced demonstration' – but the demonstration is there – RAE scores are highly correlated with citation counts, and so clearly any distortions in the methodology are trivial.

The author falls into the same trap as critics like McRoberts and McRoberts – and others. By focusing on the potential weaknesses of citation counting, he concludes that it must be flawed, when all the evidence shows that, despite its weaknesses, it is a robust and reliable tool for correlating with all other measures of eminence.

The author is correct to pick up the erroneous statement that RAE score correlates with citation counts – it is indeed RAE ranking that correlates with citation counts. No one has yet tried to correlate the citation count with the numeric RAE score. The author notes: 'In an unremarked retreat from the impasse created, a later study asserts that citation analysis should be used to suggest the rank ordering . . .' – this was not a retreat, but a development in thinking the issues through.

The author notes that one department ranked fourteenth out of eighteen by RAE grades is ranked fourth by total number of citations. This betrays a lack of understanding of rank correlation statistics. Imagine a list of RAE scores (A,B,C,D . . .) against a ranking of citation counts (1,2,3,4,5 . . .) that goes as follows:

A 1 (i.e. the highest RAE score is also the highest citation count)
B 2
C 3
D 4
E 12
F 5
G 6
H 7
I 8
J 9
K 10
L 11
M 13

There is no question that the rank order correlation is statistically highly significant here, and picking the one anomaly (fifth highest RAE score has twelfth highest citation count) is quite irrelevant to that overall truth. Taking the point further, this may mean that that particular institution would have a much reduced income if it had been given an RAE score directly linked to its citation count, but, equally, the situation may have been reversed and an institution might get higher funding based on citation counts than based on the current system – in other words, yes, there will be the odd and rare case where an institution would get significantly more, or significantly less, than the monies it would have obtained under the current system of grading. Since, however, the current system of grading is purely subjective (and, judging from the anecdotal comments after the 1996 RAE, there was considerable fury at some of the ratings given, as it was felt that the 1996 RAE panel discriminated against new universities and against interdisciplinary research), who is to say that a system based on citation counts would not be perceived as fairer?

The author notes that no case is made for using citation studies except for their lower cost. There are two methods to assess research quality. The results of the two are statistically highly significantly correlated with each other. One costs 1% of the other. Which one would the author have us use? The author should demonstrate convincingly why cost should not be used as a criterion, when all other factors are equal.

The author claims that use of citation counts followed by peer group award of scores is naive because of the risks of litigation. He seems unaware that the present system is viewed with deep distrust by most academics and that previous RAE scores have been challenged in the courts. I would have thought that the