

University of Bradford

Response to Joint Funding Bodies' Review of Research Assessment

Summary

We believe that an RAE2001-type assessment, which involves regular retrospective peer-based review combined with an informed evaluation of future plans, continues to be the fairest and most cost-effective basis for the future funding system for research. We also believe that every institution should be assessed in the same way – high quality research is not solely the prerogative of a few. We recognise, however, that the current system, although cost-effective, is extremely burdensome on the system, and improvements can be made. These include the following:

- i) Limited use of metrical data in some subject areas.
- ii) Some consolidation of UoA's (although the unhappy experience of the merged Earth and Environmental Science Panel should be remembered).
- iii) Holding it less frequently (possibly every 7 years, with the disparity of assessment period between Humanities and Social Sciences removed).
- iv) The use of a 'light touch' approach, whereby much-reduced submissions can be accepted from areas not seeking re-grading or significant increase in volume, against a penalty of gradually reducing resource for stasis.
- v) Some attempt to minimise the incentives within the current system for leaving out academic staff with 'marginal' profile at a particular grade.

1: Expert review

1.1 We believe that 'expert review' in general, and 'peer review' in particular, must remain at the heart of any research assessment methodology. This is an essential aspect to retaining the confidence of the community, and it cannot be supplanted by more formulaic metric-based methodologies, or opinion surveys. It can, and should, however, be supplemented by other contributions, in particular, from representatives of the appropriate user communities.

1.2 We can see the assessments for teaching and research converging to some extent in terms of methodology, but combination would be problematic. Units of assessment for teaching do not automatically map on to those for research. It would also make it very difficult to construct inter-institutional submissions for research (or teaching).

1.3 Most HEI's support the current assessment methodology (as applied in RAE2001), believing it to be as open, transparent and fair as possible within the constraints. It is also widely accepted as being cost-effective. It is, however, extremely burdensome on those individuals charged with preparing and assessing the material.

1.4 An alternative 'lighter touch' peer assessment system might involve a shorter report (but containing much the same detail as required now) on a periodic basis from all those research units which feel that their existing grade continues to be appropriate. These could be reviewed by an expert panel without full

assessment of each piece of work submitted. Only those units wishing to be considered for grade improvement, or wishing to claim substantial increase in volume, or those deemed to be dropping in quality, from a previous review, would be required to submit full RAE documentation. An audit procedure would be necessary to ensure that quality is maintained in those submissions not requesting re-grading. In order to encourage continuous improvement, the funding model should gradually erode the income from units choosing to remain at a given grade.

1.5 In response to the questions:

a. Should the assessments be prospective, retrospective or a combination of the two?

A combination of the two, although greatest weight should be placed on recent performance.

b. What objective data should assessors consider?

As now: measures of publications, grant income, student training, but also to include broader measures related to public understanding, commercialisation and technology transfer.

c. At what level should assessments be made – individuals, groups, departments, research institutes, or higher education institutions?

NOT at HEI level, since the majority of institutions are heterogeneous in research quality. Nor do all institutions have departmental structures that directly conform to RAE UoA's. The current method of allowing Institutions to define their own research groupings within a prescribed framework allows full expression of the diversity within the system.

d. Is there an alternative to organising the assessment around subjects or thematic areas? If this is unavoidable, roughly how many should there be?

No. There probably should be less than there are now, perhaps reduced to around 50? Fewer than this would begin to force very dissimilar research areas together into the same assessment, with the consequences witnessed within the UoA containing Earth and Environmental Sciences. It is important that RAE structures do not stifle or prescribe academic development.

e. What are the major strengths and weaknesses of this approach?

A major strength of the approach outlined in 1.4 is that it retains the best features of the existing RAE methodology, whilst removing the more onerous, and also that it protects the freedom of individual HEI's to manage their research areas as they see best. A major weakness is that it might have a slower response time to changes in research intensity.

2: Algorithm

2.1 We do not believe that any assessment based solely on metrical inputs to an algorithm can reliably assess research quality, or retain the confidence of the

community. Nevertheless, metrical inputs in some areas are a valuable adjunct to informed assessment.

2.2 Answers to specific questions:

a) *Is it, in principle, acceptable to assess research entirely on the basis of metrics?*

No. Metrical assessment is not applicable to the entire spectrum of research activity, and tends to primarily measure quantity, not quality. How is development (or developing) potential measured by metrics?

b) *What metrics are available?*

- *measures of reputation based on surveys*

It seems to be very dangerous to base the distribution of c. £1 billion of public money on the basis of opinion survey. These would have to be carefully constructed to avoid allegations of bias or collusion, and would need to contain a similarly constructed international dimension.

- *external research income*

This is a valid indicator in many areas, but is often not comparable between disciplines (or even, in some areas, within disciplines). A more robust measure for cross-disciplinary comparative purposes is **number** of equivalent status external research grants and contracts obtained.

- *bibliometric measures (publications or citations)*

Again a valid indicator in many areas, but not all academic areas have meaningful bibliometric indices, particularly when the principle output is the single author monograph. Even with bibliometric indicators as applied to journals, there is the assumption that the average index for the journal applies equally to all articles contained within the journal.

- *research student numbers (or completions)*

This is a broadly based and valid indicator.

- *measures of financial sustainability.*

Research groups with a track record of substantial and continuous external funding should clearly be allowed to use this fact as evidence of sustained research quality. Other measures, such as average income per academic, are also valid indicators of research success, but are extremely discipline-specific. Levels of institutional support are also valuable indicators of long-term esteem, but are equally discipline-sensitive.

c) *Can the available metrics be combined to provide an accurate picture of the location of research strength?*

Yes, they can be combined, but the resulting picture is only a crude representation of research strength, in need of further refinement. It is also almost exclusively retrospective.

d) *If funding were tied to the available metrics, what effects would this have upon behaviour? Would the metrics themselves continue to be reliable?*

If a major criticism of the current system were that it induces particular behaviour and ‘gamespersonship’, then a system based purely on metrical measures would exacerbate this problem considerably.

e) *What are the major strengths and weaknesses of this approach?*

If it worked, a perceived strength would be the removal of informed assessment, and therefore a reduced cost. We do not believe that, across the board, it can do more than inform judgement, although in some areas of science and engineering it might be very helpful indeed.

3. Self-assessment

3.1 Self-assessment is attractive in principle, and the ‘light touch’ model outlined in 1.4 incorporates an element of self-assessment. One has to consider, however, the validity of self-assessment in a system where major funding decisions rest on the outcome.

3.2 Answers to specific questions:

a) *What data might we require institutions to include in their self-assessments?*

This must include the auditable evidence which contributed to the assessment - publication record, grants and contracts, studentships as is the case now, plus evidence of public understanding, outreach and technology transfer activity.

b) *Should the assessments be prospective, retrospective or a combination of the two?*

Self-assessment can only be auditable if retrospective, but some element of future strategy (perhaps in terms of institutional funding priorities) might be useful.

c) *What criteria should institutions be obliged to apply to their own work. Should these be the same in each institution or each subject?*

They should certainly be the same in every Institution, and should be internationally benchmarked. As in the existing research assessment methodology, however, the criteria will need to be constructed to suit each subject area. This will mean differences between Institutions depending on subject mix.

d) *How might we credibly validate institutions’ own assessment of their own work?*

By panel assessment of submissions, and by audit visits to inspect the evidence upon which the self-assessment is based. This should involve a sampling strategy across institutions and subject areas.

e) *Would self-assessment be more or less burdensome than expert review?*

Not for the institution, but it might be so for the Funding Council and those involved in the expert review.

f) What are the major strengths and weaknesses of this approach

One of the strengths of self-assessment is that it applies the principle of 'earned autonomy'. A major weakness is scepticism within the system about the credibility of outcomes and equality of treatment.

4: Historical ratings

- 4.1 A policy that gives each institution a rating on the basis of its historical performance and/or the value of its research infrastructure contains, as stated, a presumption that research is strongest in those departments or institutions with the strongest track record. We certainly recognise the concentration of research excellence in a minority of large institutions, but this proposal would equally have to recognise the evidence provided by previous Research Assessment Exercises. In RAE2001, 97 of the 175 Institutions assessed were found to contain some element of research graded 5 or 5*. Only a handful of Institutions have no submissions rated below 4. Excellence in research is not confined to a minority of elite Institutions, nor is every unit within these Institutions excellent. Any rating based on historical performance should therefore not seek to centralise research any more than is currently the case, for the good of the UK higher education sector.
- 4.2 Institutional assessment on a historical basis is likely to be problematic because any institutional-level indicators of research activity are likely to be heavily affected by the mix of research carried out in the institution. To go below this level would be to effectively carry out a subject-by-subject review of the individual research areas within an institution.
- 4.3 Responses to specific questions:
- a) *Is it acceptable to employ a system that effectively acknowledges that the distribution of research strength is likely to change very slowly?*
In principle, yes, but the issue is how this is implemented.
- b) *What measures should be used to establish each institution's baseline ratings?*
Current and historical RAE grades, and/or R income, and/or Research Council and major charity income are the obvious baselines, but these depend on the mix of disciplines.
- c) *What mechanism might be used to identify failing institutions or institutions outperforming expectations? Could it involve a 'value for money' element?*
Self-assessment and audit might be part of the appropriate mechanism. 'Value for money' is crude and presents a number of difficulties, since these league tables

depend again rather critically on the mix of disciplines (e.g., grant income per staff member is very discipline-sensitive).

d) What would be the likely effects upon behaviour?

It would inevitably lead to over-centralisation of research, and might breed complacency. Although the former might not be unpopular in some circles, it would ultimately fossilise the system and reduce UK competitiveness.

e) What are the major strengths and weaknesses of this approach?

It is clearly cost-effective for the Funding Councils, but it will not protect the concept of supporting research excellence wherever it is found.

5: Crosscutting themes

a) What should/could an assessment of the research base be used for?

RAE data has certainly been used for a wide range of purposes – by outside funding agencies to inform funding decisions, and internally as management information. We believe that it is inevitable that these data will be used in a variety of ways. It would be helpful if the funding councils were more explicit about what the information produced by the exercise means, but it is not necessary to design a national assessment process with the explicit aim of providing reliable management information.

If the funding councils worked with the research councils to develop ‘complementary assessment processes’ (e.g., by using grant success rates to distribute a proportion of the R income), it would minimise the assessment burden. It would, however, undermine the value of the dual support system, which allows a certain amount of research to be carried out without explicit external funding.

The research councils might make more use of the assessment data to streamline some of their processes. It should remain, however, a fundamental principle that proposals in the non-thematic mode are assessed primarily by the quality of the work proposed, not by the context in which that work will be carried out.

b) How often should research be assessed?

All institutions should be assessed at the same time and with the same frequency. If the current model were to be retained, then a 5-10 year cycle would be appropriate. Ideally all subjects should be reviewed at the same time, each with the same assessment period. Possibly to reduce the burden clusters of subjects could be assessed at different times, but this would inevitably lead to game playing. If a ‘lighter touch’ model were adopted, then assessment could be more frequent, say every three to five years, if no case for change were being considered.

c) What is excellence in research?

The judgement of 'quality' in research is a combination of several factors, and not every piece of research will necessarily achieve excellence on all measures. These include cost-effectiveness, novelty and timeliness, magnitude of implications, degree of communication to the public, and potential for technology transfer. There are therefore different aspects of research (for example, genuine interdisciplinarity, creativity and applicability) that demand recognition, and these were not in general captured in RAE2001.

d) Should research assessment determine the proportion of the available funding directed towards each subject?

Not if the intention is to retain a diverse research and higher education base in the UK. It is of course possible to conceive of an algorithm which takes, as input, some or all of the factors listed. The long-term effect, however, would be to gradually divert funding away from those areas in which the UK has demonstrated research excellence, but the economic benefits of which are more removed than in others. The existing external funding mechanisms of the research councils, major charities and industrial collaborators are, we believe, sufficient drivers in that direction without the funding councils needing to take additional action.

e) Should each institution be assessed in the same way?

Yes. The current system has the advantage that it is seen as broadly fair to every institution – each has the opportunity to make a case for recognition in its chosen arena. Whilst it may seem attractive to tailor the assessment mechanism to the scale of activity (presumably a two or three tier system), contradictory arguments can be constructed. For example, should those institutions with strong research traditions be scrutinised more or less intensively than others? On the one hand they might be regarded as having earned the right to be more autonomous, but on the other they are responsible for the expenditure of a large share of the public investment in research. The notion that such a system might 'provide a ladder of improvement so that all researchers and institutions have the opportunity to demonstrate potential' seems attractive on paper, but might be difficult to achieve in practice.

f) Should each subject or group of cognate subjects be assessed in the same way?

It is essential that subject communities are free to define the most appropriate forms of assessment (within guiding principles established by agreement between Funding Councils and the sector). In general, RAE2001 appears to have struck the right balance in this respect.

More homogeneity of assessment in broad subject areas might be appropriate for some groupings, but would cause difficulties for others. For example, subjects such as Geography or Environmental Science, which span the natural science/behavioural science/humanity divide, would be caused severe difficulties if a decision were taken to group the subject with others which are not entirely congruent. An example of this is the problem experienced within the Earth and Environmental Science UoA in RAE2001, where the more broadly based aspects of Environmental Science were felt to be disadvantaged.

g) How much discretion should institutions have in putting together their submissions?

The framework of the RAE has not matched well the diverse organisational structures in all HEI's. If future assessments follow this pattern, then Institutions must retain the freedom to design their submissions in accordance with their institutional strategic plans. It is also essential that, as in RAE2001, the assessment framework is published well in advance, and institutions are allowed to make their preparations accordingly

The question of alleged unfairness to individuals is, however, a very significant issue. The Funding Councils must recognise that the problem was considerably exacerbated in 2001 by the terminology adopted in the Exercise itself. To define those not selected for return, as 'not research active' was extremely unhelpful. Any future arrangements need to allow institutions to manage this issue in accordance with their strategic plans, without damaging the careers of individuals. Difficulties can arise as a result of 'selection' when institutions wish to promote parity of esteem between the teaching and research functions of an academic.

In most institutions and subject areas there is a spectrum of research activity, and a balance has to be struck between volume and quality. The freedom of institutions to trade funding for the prestige of a high rating is constrained in the present system by the steepness of the funding curve, and we have to assume this trend will continue (although if the financial differentials between grades were less, there would be less pressure to restrict inclusivity). A compromise is needed which balances the need to identify and reward genuine research strengths and critical masses, whilst preserving the freedom of institutions to follow their own strategic mission, and the concept of parity of esteem for individual academics.

h) How can a research assessment process be designed to support equality of treatment for all groups of staff in Higher Education?

Any assessment of research, which quantifies individual performance, however executed, has the potential to discriminate against certain groups of staff. It has probably negatively impacted on the appointment to academic positions of those staff who have non-standard career trajectories, or substantial career breaks. It could also, potentially, disadvantage those who publish the results of

their research in non-English language media, either because of their own background or because of their subject of study.

If data on individuals continues to be collected, then more attention needs to be given to individual circumstances than was the case in RAE2001. This will inevitably increase the workload on those compiling the submission, and on those assessing it.

i) Priorities: what are the most important features of an assessment process?

We believe that the three most important characteristics of future assessments are:

- fair to individuals and institutions
- not burdensome
- transparent

6: Have we missed anything?

We invite respondents to tell us whether there are other issues or options not considered here. In particular, we would be interested to hear of any approach to research assessment that could not be described as a variant of the approaches listed above.

Please see summary comments at beginning of document.