

## **Joint Funding Bodies Review of Research Assessment**

*Notes based on discussion of the Research Committee of School of Archaeology and Ancient History, University of Leicester, 13 November 2002.*

A key issue raised by our discussions is that the burden of the RAE should be commensurate with its outcome. It is unsatisfactory that huge effort went into demonstrating major advances in our subject area in 2001, when the end result was an effective decrease in funding. Our response relates to an academic unit comprising archaeologists and ancient historians, with the submission made to the archaeology RAE panel and the work of the ancient historians (and some of the classical archaeologists also) cross-referred to the Classics panel.

### **Expert Review (including peer review)**

We were strongly of the opinion that this is the most suitable of the 4 proposed methods of reviewing archaeology (and many other subjects, especially those outside the science sphere). We felt that it was important that the guidelines for panels should be sufficiently flexible to allow subject panels to determine for themselves the precise mix of peer review of publications, with assessment of the research culture of a department and the use of metric data in arriving at final gradings. That was a clear strength of the 2001 assessment and the publication of the statements and the background data made this the most transparent exercise to date.

We are opposed to the idea that a system combining teaching and research assessment be devised.

If expert review is used in future RAEs, we believe that the selected publications should remain a key diagnostic of individual calibre, but feel that the final department scores should also reflect to some degree the overall research culture of research groups and of the department. In addition to, say, 4 cited publications and metrics on grant income, post-docs and postgraduates, we believe that the individual and group trajectory over a 10-year period, say, should also be profiled, so that individuals' outputs over the RAE period could be set in the context of their long-term career path.

There was some feeling that the impact of an individual judged of sub-national standard could be disproportionately damaging to a unit, in comparison with the uplift factor of internationally-recognised figures – we felt that in future RAEs all staff should be submitted, but with slightly greater tolerance levels for one or two weaker researchers in an otherwise very strong unit. Some colleagues felt that there could be greater differentiation in the process between staff who were involved in teaching, research and administration and those who are purely research staff.

There should be increased cross-referral of interdisciplinary research in archaeology, but we do not think that creating fewer, broader review panels would be an advance – we feel the current system with our ancient historians commented on by the Classics panel has worked well. However, it was felt that the criteria for cross-panel inputs and judgements could be made more transparent in future RAEs. Creating broader subject panels would weaken the reliability and credibility of peer review.

### **Algorithm**

We strongly oppose this suggestion, whilst recognising that it may work more satisfactorily in some areas of the sciences, where metrics such as citation indexes are more commonly used than they are in the Arts and Humanities sector. The fact that some archaeology departments contain a higher percentage of science-based practitioners than others also skews the distribution of external grant income within the subject – grant winning is to some degree a measure of 'research cost' and therefore is not the prime measure of research excellence, though it is a useful indicator of competitiveness in another sphere of peer review. We would expect top-rated departments to be successful in such competition, but distrust absolute grant income rankings to reflect overall quality within the sector.

Our main objection to the use of an algorithm is that it will impact on behaviour in other theatres of research governance. If RAE rank becomes solely linked to grant income, PhD studentships, citations, etc., then all of these areas will become open to manipulation and patronage – something that would be very damaging to the principles of impartiality and peer review in all subjects.

At present the metrics available are not of sufficient quality (or equal authority) in all disciplines and the lack of fit between RAE and HESA units compounds this.

### **Self-assessment**

The experience of the QAA suggests that self-assessment would make the RAE process a more adversarial and disputed process and would increase the negativity of the process – focussing on decreasing the research esteem of individuals and groups. This could have a major impact on morale and an increased danger of litigation and appeals. The process might actually prove more burdensome for departments. There is the possibility, as with QAA, that Universities would become more adept at playing the system. As with teaching, it would probably also make research statements more conservative and risk free – to the detriment of the research culture of innovation.

### **Historical Ratings**

This seems to us an inappropriate measure, unless there is an acknowledgement that there will be no positive change in research funding for the foreseeable future – which would be calamitous for UK research. The funding outcome of the 2001 RAE was a real snub to the academic community, given the scale of improvement demonstrated across the disciplines. To have credibility in the HE sector, any future exercise must offer the prospect of rewards both for improvement and sustained high performance level. Historic ratings could put the research effort in the UK into reverse, providing no incentives for investment and innovation. Whilst the idea might have some attractions to the elite Universities with a high percentage of staff in 5/5\* departments, it would remove the possibility of the majority of institutions improving their position. The results could be complacency in some Universities and loss of morale in others – both damaging to UK research. Recent history suggests that over successive RAEs some institutions have strengthened their research profile considerably and spectacular improvement has been seen at subject level. In archaeology, we can name 4 departments that have risen from 2 or 3 to 5 across the last decade (and one is now 5\*).

### **Crosscutting themes**

- a) **What should the assessment of the research base be used for?** The determination of a significant proportion of the HEFCE QR stream from the RAE results is clearly appropriate. We are not convinced that making some of the RC funding (as has been suggested in some quarters) tied to RAE grades is a good idea. The RAE grades already influence decisions about postgraduate studentships and research grants. However, we believe that this should not be more formally linked. Some 5\* researchers are in departments with overall ratings of 4 or below. The publication of the RAE 2001 baseline data has been a major advance and this is invaluable in HEI internal management and monitoring of research activity.
- b) **How often should research be assessed?** We believe that the minimum period under review should be 5 years, but that even if cited publications are taken from such a period, some review of the longer-term plans and profiles of individuals should also be undertaken. This might be particularly helpful in the case of junior staff, or of staff who have had specific difficulties during an assessment period (maternity leave or illness). Could this be accomplished by including a brief personal statement by each submitted staff member to set their cited publications in the context of their long-term research development? We are firmly of the view that the cross-disciplinary single census date snapshot of the RAE as currently established is the fairest system, rather than assessing subjects at different times on a rolling programme. The latter system would potentially be more burdensome as institutions might devote greater efforts to each unit as it came round, managing the process ever more professionally.
- c) **What is excellence in research?** The RAE process needs to recognise disciplinary diversity here. There is also the difference between individual contributions and the impact of the overall research culture of a unit. The process should reflect both inputs (grant income) and outputs (primarily publications, but not exclusively so). Excellence is not just about absolute quality, but must also reflect scholarly impact in the field (has this research changed the way people think). In some disciplines the impact of research may be slower to show itself than in others, due to the time-lag in for instance journal publication. There was a strong feeling that the breadth of archaeology and ancient history as disciplines means that there are both researchers

with large grant income and interdisciplinary team-based projects and others who are solo researchers producing articles and monographs. The system needs to be flexible enough to recognise the strengths of both those extremes. Expert peer review appears the best way of doing this.

- d) **Should RAE determine funds available to each subject?** A weakness of the present system is that it aims to demonstrate the international significance of UK research, based primarily on the say-so of UK-based experts. Although the 2001 review involved consultation of some international experts, this aspect of the review could be strengthened or at least more formalised. Some subjects, among them Archaeology and Classics, have a very high international standing and have nothing to fear from international scrutiny of the proposed gradings. A distribution based on an algorithm linked to external funding for each subject would penalise the non-Science subjects and would again confuse the measure of cost with that of quality of research.
- e) **Should each institution be assessed in the same way?** Yes, we believe that all institutions entering the RAE should be assessed in the same way, though in the case of Universities competing at the bottom end, this is only sensible if they have some expectation of resources following the award of grades at the lower end of the scale.
- f) **Should each subject be assessed in the same way?** The 2001 system of setting out a broad system, but allowing panels to place greater emphasis on one or another element of metrics and expert review seems an effective model to take account of disciplinary variation – provided of course that the ground rules that each panel will operate under are clear well in advance. Consultation with subject associations at an early stage should be an important part of this process. The burden on panel members is now considerable; might it be preferable to schedule the review meetings as a block of, say, 1-2 weeks at a venue where all the submitted publications would be available for them.
- g) **Discretion of Institutions in putting together publications.** We are in favour of a system that reports more fully the total research activity within the subject area (reducing the tendency for institutions to exclude individuals), accompanied by some slight mitigation of the impact of less productive individuals on the overall score of a unit. The prescription that a single individual of sub-national level will prevent a unit reaching 5\* should be reconsidered if a more inclusive census is demanded.
- h) **RAE and equality of treatment.** We feel that more flexibility should be built into the system to take account of unusual life factors within a review period affecting individual output (maternity leave, illness, illness of a close relative, etc. can all have a major impact). Rather than simply giving more weight to departmental statements about such situations (which in itself would be helpful), we feel it might be more appropriate to profile cited publications against a longer-term research trajectory for each individual. The same applies to very young researchers, where the long-term impact of the research will be very much harder to judge. The system needs to be more adept at identifying the rising stars, as well as the established international figures. Should the panels also include some young scholars as well as senior figures and those reaching their peak in mid-career?
- i) **Priorities.** The main priority we identify is that the funding outcome of RAE should reflect the quality of the result achieved. It is unsatisfactory that overall improvement is rewarded with reduced funding. The HE sector is willing to shoulder a certain burden of work in this regard, provided the rewards are there for improved or sustained excellence of performance. Transparency of the process is welcomed. Although there is some perception that the current system of expert review is time-consuming, it was uniformly felt by us to be the best method of the four canvassed (albeit that it also incorporates some use of metrics and self-assessment in the judgement). Since existing monitoring and record keeping are geared to this system, a radical redesign of the process might prove more burdensome than fine-tuning the current one. The suggestion that the views of the expert panels may be subjective carries less weight now that they are informed by a range of metrics, even when these are not used as part of an algorithm. The quality judgements of people who have read and engaged with all the individual submissions should remain central to the process, at least in subjects like Archaeology and Classics. We believe that the

overall scores for a department should also reflect an assessment of the overall research culture of the unit and not simply be the average of the individual scores, taking due account of the value added that a vibrant research group brings to individual efforts. Finally, we were surprised that the review documents did not make more explicit mention of the scoring system and scale to be used, or the way that individual scores (international, national, sub-national) are to be mapped onto the existing (or any new system) of overall departmental grades. It is vital that this information is made available at an early stage in the cycle.