Reform of Higher Education Research Assessment and Funding

The Royal Society of Edinburgh (RSE) is pleased to respond to the Department for Education and Skills’ consultation on the reform of higher education research assessment and funding. These comments have been compiled with the assistance of a group of Fellows from a range of disciplines under the direction of one of the Society’s Vice-Presidents, Professor Jan McDonald. (The membership of the group is attached).

The Royal Society of Edinburgh is Scotland’s National Academy, with a Fellowship of elected scholars and practitioners from the fields of science, engineering, technology, the arts, humanities, law, social sciences, business and commerce. The Society is, therefore, well-placed to provide a considered, and an independent, view of the implications of the proposals set out in the consultation paper for all areas of academic research.

The consultation paper makes reference to criticism of the RAE, primarily, but not exclusively, in terms of what it costs, but does not attempt to evaluate the costs or to assess the benefits of the current system. It does not present an argument for the principle of moving to a metrics-based system but, on the contrary, assumes that such a scheme will be adopted in some form. The Society considers that it is essential to retain the principle of peer review of research outputs in order to preserve the international excellence of British research.

The Society supports the aim of “reducing substantially the cost and burden of the current Research Assessment Exercise (RAE), while continuing to recognise and reward research excellence in ways that are accepted by the sector as appropriate”. However, we do not consider that the approaches set out in the consultation paper, all based on research income, provide a way forward for assessing research performance. This is not to argue that metrics have no part to play in research assessment and, indeed, the subject based RAE panels already draw on quantitative as well as qualitative indicators where appropriate. Although the scope for employing metrics in research assessment varies between subject areas, the important element of the current approach in all disciplines is that quantitative data are used to inform the review process, not to drive it.

The consultation paper accepts that research income cannot provide a satisfactory measure of research excellence for what are termed non-STEM subjects. However, the Society is concerned about what it sees as a false divide between STEM and non-STEM subjects. Some subjects within the arts and humanities, for example the creative and performing arts and aspects of the social sciences, are equally as dependent on research grant income as are some STEM subjects. On the other hand, some areas of STEM subjects, for example in mathematics and theoretical physics, are more akin to those subjects in the arts and humanities where excellent research is not dependent on significant external income. A robust system of assessment is one that is appropriate to the specific area of research. We do not consider that one quantitative measure, or even a basket of measures, can be an accurate proxy for excellence in all areas, and certainly not without the moderation afforded by peer review.

The Society has consistently supported the principle of the dual support system, and welcomes the Government’s continuing commitment to this approach to research funding. QR funding and research council funding serve different purposes and have different aims, and both are important. We are concerned that if the QR funding stream is directly linked to levels of research council funding, many of the benefits of the dual funding system will be lost to the detriment of the quality and range of research in the UK.

The Society is concerned about the conflation of the issues of assessment of research quality and allocation of funding (QR). This is a confusion that runs right through the consultation paper. While the RAE is a UK-wide exercise, the use of RAE results in research funding is devolved to the different funding councils. It is essential that quality assessment is developed quite separately from any consideration of issues such as funding “pots” for different subjects.

The questions raised in the consultation paper are addressed below.

Which, if any, of the RAE 2008 panels might adopt a greater or wholly metrics-based approach?

As the RAE 2008 main panel chairs have pointed out, they have spent considerable time developing rigorous methods and criteria that have the confidence of the academic communities. Where appropriate, judgements will be informed by metrics. It is too late to introduce changes to the exercise at this stage after institutions have been provided with very specific instructions in preparation for RAE 2008. Such a belated
intervention would be destabilising and impair confidence in the conduct and in the outcome of the procedure.

If, however, there is to be a lighter touch metrics-based approach to research assessment in the future, it will be important to have clear means of developing and evaluating this alternative against established quality benchmarks. RAE 2008 with its rigorous approach will provide a very clear set of benchmarks, and it should proceed as planned with a metrics-based “shadow” exercise conducted quite separately. Giving more weight to metrics in RAE 2008 than is already planned would throw into question the validity of the comparison because it would undermine the independence of the two methodologies.

Have we identified all the important metrics? Bearing in mind the need to avoid increasing the overall burden of data collection on institutions, are there other indicators that we should consider?

There are many input and output metrics that might potentially inform research assessment and, where appropriate, some of these are already taken into account by RAE subject panels. The consultation paper identifies and discusses a number of indicators and measures but then goes on to describe models all based on only one indicator, research income. The Society is firmly of the view that no single indicator can provide a substitute for peer review of research quality, and that even a range of indicators can do no more than inform peer review.

Although indicators may be helpful in drawing comparisons within a UoA, they tend not to be useful across different disciplines. For example, research income fails to take account of research undertaken without external funding. Even within STEM subjects this approach disadvantages the more theoretical areas of, for example, chemistry and physics. The question of which metrics could be employed to address the distinction between ‘scholarly’ and ‘applied’ research which exists in many STEM and non-STEM subjects is unaddressed. Although bibliometric data do include research not supported through specific grants, patterns of publication, for example, the monograph as opposed to the journal article, differ across the range of disciplines and within cognate subjects. In addition, citation practice differs from discipline to discipline. High levels of citation are not necessarily a guarantee of quality, and in some areas there is no acceptable hierarchy of journals. Minority subjects with fewer opportunities for publication in ‘high-ranking’ periodicals will be at a disadvantage, although the standing of the work may be excellent. ‘Fashionable’ topics on which much is published (and cited) in the short term will conversely be privileged. Receipt of prizes requires that the standing of the prize is assessed against other awards, and similar considerations apply to such indicators as membership of societies, invitations to make presentations at conferences, international exchanges and service on boards and committees. Whilst all such data can be useful in research assessment, a fair and meaningful measure of research quality requires that the raw data are interpreted by a process of peer review.

Which of the alternative models described in this chapter do you consider to be the most suitable for STEM subjects? Are there alternative models or refinements of these models that you would want to propose?

As stated above, the Society does not consider that any one indicator can provide a measure of research quality. Therefore, none of the models described is satisfactory.

We have already referred to the shortcomings in the applicability of research income as the principal indicator of research quality across all disciplines. Any metric is not equally valid for all subjects, and it is important that differences between subjects are recognised with indicators selected for their appropriateness to the area being assessed. Although, as the consultation paper says, metrics cannot be expected to reproduce in full the granularity of peer review panel conclusions, the aim should be to develop metrics producing as fine-grained an assessment as possible appropriate to a given area of research.

What, in your view, would be an appropriate and workable basis for assessing and funding research in non-STEM subjects?

As stated above, the Society does not accept the simplistic distinction between STEM and non-STEM subjects. The use of the term ‘non-STEM’, in defining the humanities, arts, and social sciences in terms of ‘what they are not’, is in itself divisive and outdated, and does not adequately reflect growing interest throughout the academic community in interdisciplinary and collaborative research.

The Society considers that one of the strengths of the RAE is that it operates a system of peer review across all subject areas. We consider it appropriate to retain this common approach, albeit recognising the differences that exist between units of assessment. Although we accept that in some areas of the arts and humanities there are few if any appropriate metrics to inform research assessment, this is not true for all areas. If a purely metrics approach were to be adopted for the assessment of only STEM subjects, this would have implications for those subjects that fall in the indistinct boundary between STEM and non-STEM,
such as cognitive science, computational linguistics and informatics. It would also discourage the important and growing body of interdisciplinary research that crosses the boundary.

If there is an argument for a slimmed-down RAE, this should be explored for all subject areas and not just those in the arts and humanities.

**What are the possible undesirable behavioural consequences of the different models and how might the effects be mitigated?**

The utilisation of only one measure as an indicator of research quality effectively sets a target for researchers when QR funding is based on the assessment exercise. If the measure is research income, as in the models, researchers will look to increase income from those sources included in the calculation. This will result in increased pressure on research councils, and significant increases in their costs. At an institutional level there will be an incentive only to appoint established researchers, running counter to the importance placed in RAE 2008 on nurturing the research of the future through the recruitment and support of junior staff as a sign of a quality research environment. As generators of QR funding, the more expensive areas of research will be favoured over those requiring few or no external grants. Areas of research known to be favoured by funding bodies will be pursued, leading to greater emphasis on areas currently in fashion regardless of the real needs of both scholarly and applied research.

The RAE is already considered to disadvantage areas of research where relatively long periods of time are needed for a publishable outcome. Examples include the preparation and writing of major monographs, the compilation and editing of comprehensive dictionaries and reference works, long term environmental studies essential to important projects such as biodiversity conservation, and the elucidation of the taxonomy of important groups of organisms such as parasitic fungi and marine micro-organisms. An approach based only on metrics is likely to discourage activity in such important areas of scholarship.

Traditionally, although admittedly less so in the current Higher Education environment, research and teaching have been regarded as symbiotic, the one feeding and enhancing the other. The importance attaching to the generation of research income will further increase the value placed by institutions on research active staff to the detriment of teaching. Regrettably, there is already a gap developing between research and teaching that the proposals in the consultation paper would only serve to widen.

While we accept that any system of assessment will influence behaviour to some extent, the use of a small number of quantitative indicators would introduce perverse incentives with significant resource costs. Quantitative indicators alone cannot provide an accurate measure of research quality and, therefore, their use in isolation would only serve to distract research effort from the pursuit of quality. The Society is of the strong opinion that metrics have an important role to play in research assessment but that their influence must be moderated by a system of peer review if the recognition and support of quality is to remain the aim.

**In principle, do you believe that a metrics-based approach for assessment or funding can be used across all institutions?**

Again this question conflates the assessment and the funding of research. As argued above, no single measure is equally appropriate to the assessment of quality in all disciplines. The results from any solely metrics-based assessment will, therefore, depend on the discipline-mix within an institution and the choice of metrics. There is a danger that a metrics-based approach would seriously disadvantage small specialist institutions, and adjustments would have to be made to correct such anomalies. However, subject peer review panels (informed as appropriate by metrics) can provide a reliable assessment of research quality across all institutions, taking into consideration the objectives and priorities of each. Whether the same funding formula should apply to all institutions is a separate question.

**Should the funding bodies receive and consider institutions’ research plans as part of the assessment process?**

If the intention is to reduce the burden of the RAE on both institutions and review panels, the inclusion of research plans in the assessment exercise would run counter to this. Research plans are more to do with aspirations than achievements, and are prone to change in response to changing circumstances. It is, therefore, difficult to draw meaningful inferences of research quality from future plans. That said, at UoA level research plans might be helpful in providing a strategic context for the assessment.

**How important do you feel it is for there to continue to be an independent assessment of UK higher education research quality for benchmarking purposes? Are there other ways in which this could be accomplished?**

International benchmarking of research quality is seen as increasingly important at subject level. The Society also recognises that there exists a common interest in working to improve the international standing of UK
research. We welcome the improvements in benchmarking that are being introduced for RAE 2008 with the enhanced role of the main panels in overseeing the work of the review panels.

Although trends in a few carefully chosen metrics can be useful in international comparisons of national research bases, they are less reliable at lower levels of aggregation. At subject level peer review along the lines of the RAE, possibly simplified and streamlined, is an essential component of benchmarking. International input to the peer review process is essential to ensure that the strength of research in the UK compared with elsewhere in Europe is properly recognised.

Additional Comments

The Society recognises that there could be scope for a lighter touch RAE that continues to enjoy the confidence of funding bodies and institutions. However, the models proposed in the consultation paper are fundamentally flawed as measures of research quality. The danger is that the serious concerns that exist about the proposals will lead to an uncompromising defence of the RAE 2008 process, and unwillingness to countenance any reform.

We consider that there remains the opportunity to mount a shadow exercise to the 2008 RAE that would help evaluate how best to incorporate metrics to simplify and streamline the RAE process. However, considerable work would be required to design this shadow exercise that should cover all units of assessment. The Royal Society of Edinburgh, with its multidisciplinary Fellowship and its independent status, would be pleased to work with DfES and the funding councils in taking this forward.

Any enquiries about this submission should be addressed to the Research Officer, Ian Melville (email: imelville@royalsoced.org.uk).

October 2006

MEMBERSHIP OF FELLOWS GROUP

Professor Jan McDonald (chair)  
Vice-President, Royal Society of Edinburgh
Professor Brian Ashcroft  
Fraser of Allander Institute, University of Strathclyde
Professor David Birch  
Department of Physics, University of Strathclyde
Professor Sue Black  
Dept. of Anatomy & Forensic Anthropology, University of Dundee
Professor John Butt  
Department of Music, University of Glasgow
Professor April McMahon  
School of Philosophy, Psychology and Language Sciences, University of Edinburgh
Professor Frank Odds  
Aberdeen Fungal Group, Institute of Medical Sciences, University of Aberdeen
Professor Ian Underwood  
Scottish Microelectronics Centre, University of Edinburgh