

## Response to HEFCE's consultation on *the assessment and funding of higher education research post-2008*

The Royal Society is pleased to respond to the Higher Education Funding Council for England's consultation on the *Assessment and funding of higher education research post-2008* (HEFCE 2007). This submission has been formed with the advice of a small group (membership at the end of the document), and approved by the Council of the Society. We make some broader points before responding to the specific consultation questions.

### Key Points

- The Society strongly endorses the current dual support system of financing research in UK higher education, and we firmly believe that this should continue post 2008.
- We strongly believe that existing and proposed metrics should be used as indicators only, and that to assess fully the quality of research peer judgement is needed, in all science subjects.
- The implementation of a bibliometrics-based research assessment framework is a substantial and complex task, and the proposed timescale for the pilot is unrealistic. We believe therefore that it is necessary to allow RAE2008 to inform funding for longer than is currently proposed, and to take more time, at least an extra 12 months, to learn from the pilot.
- The Society does not believe that the currently proposed division between 'science' and 'non-science' subjects is appropriate or workable.
- The future success and sustainability of the research base involves a number of broader research-related activities such as public engagement, innovation and engagement with user communities, and contributions to policy. The Society believes that an overall research assessment system, that includes the Research Excellence Framework (REF), must properly recognise these activities, and is concerned that the current consultation is not offering an integrated perspective that considers how the REF will link with a consideration of these other aspects. There will be a need for peer review regarding the recognition and rewarding of these activities.

### General points

#### The dual support system

Dual support is an effective mechanism to sustain excellent research. The vital plurality of judgement, which is a central feature of dual support, would be lost if either funding stream is directly dependent upon the other.

As we have said in the past (Royal Society 2006) we agree with Government that dual support is a valuable system that rewards excellence and nurtures promise. Part of the value stems from the way the dual support system allows greater plurality of decision-making in funding allocations and hence greater scope for creativity to flourish. The peer review judgements facing Research Assessment

Exercise (RAE) panels are different from those facing a grant panel looking at the promise of a specific project proposal. If decisions about the quality related (QR) component of the Higher Education Funding Councils' block grant were made solely on the basis of decisions already made by Research Councils and other funding agencies, the effective outcome of an income-based model, some of the value of dual support would be lost.

#### The aims of research assessment

A fundamental aspect of any assessment system is what precisely it assesses. An assessment of "research excellence" can therefore be one of many things. The assessment could focus solely on excellence in terms of published outputs, or could consider excellence as it exists in a number of forms. To encourage a vibrant and diverse research base, all kinds of research and excellence need to be able to flourish. There is therefore a need for an overall assessment system that recognises and rewards excellence of all kinds. This involves both recognising diversity within research types, and recognising the importance of broader research-related activities carried out by researchers.

The research assessment system in place at any time will necessarily be a strong driver of culture within universities, who will modify their behaviour to optimise their gains as much as possible from whatever research assessment is in place. Indeed, a fundamental aim of a research assessment system is to encourage certain behaviours, and/or to discourage other behaviours.

We believe that any overall system must evaluate, and be sensitive to, a number of aspects, in addition to assessing quality in terms of published outputs. Assessment must adequately recognise and reward user-focused research (as discussed further in Q4). Additionally assessment must be able to capture excellence in activities such as public policy involvement or involvement in public engagement or outreach work. We fully agree with John Denham's recent statement that the proper recognition of researcher engagement in policy-related activities is an area that requires further consideration. Researchers who, for example, produce fewer research papers, but who provide excellent evidence and advice to policymakers should not be at a disadvantage in the assessment system. A research assessment system that rewards quantity, rather than focussing on peaks of achievement could therefore be to the detriment of these other activities.

Measuring and assessing excellence in these broader research-related activities is difficult, and devising suitable assessment mechanisms will inevitably be a complex task, requiring a great deal of time and thought. There are no easy metrics that can be used, and we believe that robust assessment will require some form of qualitative peer and other expert judgments. Whatever the outcome of this current consultation exercise, the fair assessment of these broader activities must be a crucial consideration when devising research assessment systems for UK higher education.

The Society has always had an active interest in research assessment and resource allocation, submitting evidence to major consultations, and publishing proactive position pieces. Our key principles lead on from work we have undertaken in the past:

- Any assessment method must include an element of peer judgement (Royal Society 2002, 2003a, 2004 & 2006)
- Quantitative indicators can inform the judgement of these peer panels (Royal Society 2002, 2003a, 2003b, 2004 & 2006)
- Different subjects may have different relevant indicators (Royal Society 2002, 2003a, 2003b, 2004 & 2006)

### Timetable and pilot exercise

Moving a significant part of the research assessment framework to bibliometric assessment is a major undertaking, and therefore requires a thorough and robust pilot exercise. Any shortcomings or problems that emerge from the pilot must then be fully resolved before a full scale exercise takes place. The proposed timescale appears unrealistic. The Society stresses that any move to bibliometric research assessment must only take place once the system is robust, which we believe would involve extending the current timescale, by at least an extra 12 months, to allow for a longer and more thorough pilot exercise. It is vital that the research assessment framework is robust, and accepted as such by the sector: getting it wrong would have far-reaching implications and would seriously damage UK science and academic research.

There are a number of key concerns that must be addressed during and following the pilot phase. The following are what we consider to be the most important to resolve before proceeding further.

- The pilot will need to be sensitive to a climate where researchers and university departments will inevitably alter their behaviour to maximise their gains under the assessment system. While this is inevitable in any system where there is so much riding on the outcome (including the current RAE), specific behaviours that could distort the results of bibliometric assessment will need to be addressed before any full scale exercise takes place. In this regard it should be noted that a single additional citation for each publication from a university department could make a vast difference to the outcome under the new proposals.
- A particular risk is that new emerging areas will not be rewarded due to the fact that there will only be a small pool of individuals working in the area and therefore fewer opportunities for citations. Individuals undertaking high risk research or research lines with no obvious outcomes in terms of publications and citations must know that their work will be subject to a fair assessment.
- The assessment of interdisciplinary research is a challenge that needs to be resolved in the pilot phase, so that researchers working in these areas can enter a full scale bibliometric exercise with confidence that their work will be judged fairly. Care will need to be taken to ensure that both research that straddles multiple science groupings, and research that straddles the proposed science/non-science divide are appropriately recognised and rewarded. An additional point is that we would be concerned if a paper was to be assigned for citation purposes to a subject area on the basis of the names of the journal, rather than the returned subject of the author.
- The pilot must allay the fears of some that bibliometric assessment measures popularity rather than excellence (which may not always be found together). The pilot must convincingly prove that the notion that the quality of research can be indicated by citation level is robust. Care must be taken to ensure that a research assessment framework that uses citation counts as a central measure does not in effect encourage academics to undertake PR activities to ensure their (rather than an equally valid and good) citation is used. Further issues to address when asking whether a citation is inevitably a marker of quality include: papers describing new techniques, which may not equate in quality to fundamental activity are likely to have high citations; and, due to editorial pressure for brevity, authors may cite a single review article rather than the original papers in which the fundamental work appears. There is a need to ensure that there is no incentive against publishing (ie researchers should be expected to have a mix of top-performing

and other papers – the less cited papers should not detract from the existence of their most cited papers).

- The robustness of the Thomson database (and any other databases that may be used) must be demonstrated during the pilot exercise. Unless and until the higher education community is convinced of the accuracy and appropriate coverage of the database a full scale bibliometric assessment will not be accepted by the academic community. A fundamental concern is the information contained in the database: all disciplines subject to bibliometric assessment must have full confidence that the database adequately covers their discipline. If the Thomson database has inadequate coverage of some areas either these areas should not be subject to bibliometric assessment, or databases with adequate coverage should be considered alongside the Thomson database. An inevitable side effect of 'bolting on' other databases however would be to encourage debate as to what databases/measures should also be included, leading to disciplines each arguing for additions that would benefit their subject over others.
- The outcome of the pilot exercise must illustrate that the methodology used does work for all applicable disciplines. Issues to address and resolve include the fact that citation rates vary within disciplines as well as between disciplines - this must be fully calibrated in order for robust bibliometric assessment to take place. Also, in a field with high citation counts bibliometrics may be able to provide a reasonable level of discrimination, but it may be more difficult to achieve this in a field with lower citation rates. The fears that some disciplines may suffer if the citation count is limited to 5 or even 10 years must be allayed.

We believe it to be essential that further consultation takes place following the pilot exercise, in the light of the findings of the pilot. This will enable the sector to discuss the proposals in the light of the pilot's findings. Also, if the pilot is deemed by HEFCE to be successful, this further consultation and discussion of the pilot's outcomes will be essential for assuring the sector that bibliometrics are a reliable and valid method of assessment.

Finally, the Society strongly believes that metrics should only be an indicator, and that to fully assess the quality of research, peer judgement is needed in all science subjects.

#### Subject divisions

Under the proposals mathematics and statistics are not being moved to assessment using a bibliometric system. The Society agrees that mathematics and statistics should be judged by light touch peer review, rather than by bibliometrics. However, there are many science disciplines that are proposed to sit under bibliometric assessment that interact closely with these subjects, and which we believe should be assessed using similar criteria as mathematics and statistics, such as computer science, informatics, engineering and physics. These subjects then have strong links with other science subjects, for instance computer science has links to cognitive science, synthetic biology and systems biology, which have links to other biological sciences and so on. Therefore, it is not simply a case of redrawing the boundary between the two groups: moving disciplines or subdisciplines from one side of the boundary to the other side creates more issues that then require addressing. These issues require further investigation as part of the pilot exercise.

It is vital that the separation of science from non-science subjects does not prevent interdisciplinary work that crosses this boundary from being fairly considered. Under the proposals this work could therefore be subject to either bibliometric analysis or to peer review depending on which discipline submitted the paper for assessment. Some areas of science are very multidisciplinary - for example earth science is carried out by geologists, palaeontologists, microbiologists, chemists and biologists. The assessment of

multi-authored papers where there is a diversity of disciplines involved will require particularly careful consideration.

The Society also has concerns regarding the proposed make-up of the six groups of science-based disciplines. The grouping of subjects means that some groups contain subjects which utilise publications in very different ways. For example physics and environmental sciences are currently in the same group, but the way the two sciences use publications is very different. Careful thought will be required to ensure that all disciplines within each grouping are treated fairly within bibliometric assessment.

Due to the above concerns, the Society believes that there is a need for a fundamental rethink regarding both the boundary between science and non-science categories, and regarding the proposed make-up of the six science groupings. In our view it is impossible to draw an adequate and workable distinction: all science subjects should be on the light-touch peer review side of the boundary. The Society stresses that all panels must have appropriate expertise and the remit to use their judgment in the identification and assessment of excellent research.

#### Early career researchers and postgraduate students

It is vital that early career researchers are not discriminated against under the research assessment system - a healthy and vibrant research base can only be sustained if there is a constant inflow of researchers, who are then able to flourish and progress. Early career researchers are the lifeblood of the sector and are essential for the sector's future health and sustainability.

The Society is concerned that the proposed bibliometric approach will have an adverse impact on those researchers who have been publishing for less than 5 or even 10 years, and who have not yet built up a sizable body of published work: there may well be excellent work amongst the output produced by an individual who has only been publishing for a few years. It is not only early career researchers who would be discriminated against within a bibliometric approach: postgraduate students may be discouraged by their institution from publishing, if it is thought that publication of their work would bring down the citation profile of a particular institution.

Institutions will naturally aim to maximise their position in the assessment, by controlling who or what is submitted for assessment: this is inevitable when there is so much riding on the outcome of the assessment process. Departments will be aware of how staff and potential staff members positively or negatively affect their citation score, and it will be in their interest to recruit accordingly.

### **Responses to the consultation questions**

- 1 a) *Do you endorse our proposals for defining the broad group of science-based disciplines, and for dividing this into six main subject groups, in the context of our new approach to assessment and funding? b) Are there issues in relation to specific disciplines within this framework that we should consider*

As discussed under general points above, the Society does not believe that the division between the 'science' and 'non-science' subjects is an appropriate division. We believe that there are deep issues that cannot be resolved by moving disciplines or subdisciplines from one side of the boundary to the other side, as this creates more issues that then require addressing. We also have serious concerns regarding the proposed make-up of the six groups of science-based disciplines. The grouping of subjects means that some groups contain subjects which utilise publications in very different ways. Additionally, it is vital

that the proposed separation of science from non-science subjects does not prevent interdisciplinary work that crosses this boundary from being fairly considered.

- 2 *Do you agree that bibliometric indicators produced on the basis that we propose can provide a robust quality indicator in the context of our framework? Are there particular issues of significance needing to be resolved that we have not highlighted?*

See comments under 'timetable and pilot exercise' in our general points above. Further to this, we believe that the pilot exercise and its outcomes must reassure researchers who work in subdisciplines with widely varying norms and practices that the assessment system is fully sensitive to their needs. For example, some discipline specific points are:

- In particle physics collaborations are extremely large, sometimes consisting of hundreds of individuals. Dividing the number of citations by the number of authors would not be a suitable solution, nor would using the mean citation on all the papers authored by a specific researcher. Within these large collaborations a sub-set of authors may produce work on a particular aspect on the experiment – how this would be assessed under bibliometrics requires clarification. Additionally, there is a discipline tendency to cite a lot of proceedings or preprints, which would not be noted under bibliometric assessment of the Thomson database. A further point is that it can take years for some work (eg R&D work on detectors) to take place, therefore citations referencing this work will only build up over a relatively long timeframe.
- In computer science conference papers are often rated as highly as or more highly than journal papers – any bibliometric system will need to be sensitive to this. Coverage in Thomsons for computing is limited, and we would stress that no discipline should be subject to bibliometric assessment until researchers are confident that the assessment system will judge them fairly. Moving computer science and informatics from the 'science' side of the boundary to the light touch peer review side of the boundary would solve this issue, however this would then create further challenges as these subjects are fundamental tools that are used across many science disciplines.
- Citation rates can vary widely within a discipline. For example papers on taxonomy underpin much biological science research, but are themselves not highly cited. Instead it is the research that relies upon these taxonomy papers as an underpinning that is cited more widely.

The results of the bibliometric analysis should be available to institutions and others in a useful form that enables analysis at subject or institutional level, including the ability to 'drill down' to identify pockets of excellence.

New and important papers that could represent major steps forward within their field would not be measured due to not having yet built up citations. This is one area where peer review is essential, as it is able to judge the potential impact of new work.

The algorithm to which bibliometric data will be subject will necessarily be complicated, as it must represent fairly all the possible categories and exceptions and normalisations. Therefore only those individuals with a relatively strong grasp of mathematics will understand the workings of the system. This means that for many 'users' of the system there will be a lack of transparency, with people asked to put a lot of faith in a system they may not fully comprehend.

Collecting, managing and evaluating the bibliographic data will require a very large information system. There is inherently a high risk in any project that requires as its basis a complex information system of this sort.

- 3 *What are the key issues that we should consider in developing light touch peer review for the non science-based disciplines? What are the main options for the form and conduct of this review?*

The Society agrees that mathematics and statistics should be assessed by light touch peer review, and that bibliometric assessment is not suitable for these subjects. However, as discussed above the Society believes it would be more appropriate for HEFCE to develop light-touch peer review for all disciplines, and that it would be preferable not to have this distinction between 'science' and 'non-science' subjects, as it is not possible to draw a satisfactory boundary between the two groups.

We would like to see evaluation carried out by panels of experts in each discipline, as opposed to overarching cross-discipline panels. Second, evaluation should be limited to a specific number of outputs per researcher, supported by a consideration of quantitative and qualitative indicators if appropriate.

- 4 *Is there additional quantitative information that we should use in the assessment and funding framework to capture user value or the quality of applied research, or other key aspects of research excellence? Please be specific in terms of what the information is, what essential element of research it casts light on, how it may be found or collected, and where and how it might be used within the framework*

The assessment of applied research is no less about measuring excellence than the assessment of basic research, and it is vital that this assessment is carried out fully and fairly. The wider issue of the assessment of applied research (and other activities not easily included within the proposed REF) is discussed earlier in this response. However, if the assessment of applied research were to remain within the REF then we believe the assessment of this research should combine the use of robust representative indicators with the judgements of informed and expert peers.

It is essential that measuring excellence in business-focused areas considers those measures that are most applicable to the particular subject under consideration. For example some areas of computer science are strongly business-focused and utilise conferences and workshops involving industry to a greater extent than they do publication in journals. Relevant measures of excellence for subjects such as this could include best papers, or invited or plenary talks. Other possible indicators of quality could include patents, or money spent on spin-off companies. If applied research is to be measured in a meaningful way then an element of peer review by relevant experts has to be introduced to judge and grade the excellence of such activity, including the interpreting of outputs, as shown by the quality indicators.

An additional point is that much innovative research is interdisciplinary in nature: it is vital that the research assessment system ensures that this research does not fall between panels and is therefore not subject to adequate evaluation.

- 5 *Are our proposals for the role of expert panels workable within the framework? Are there other key issues on which we might take their advice?*

Panels need to be able to use their experience and expertise to make judgements on the quality of research. This will include using indicators, such as bibliometrics, to inform their judgment, and doing more than rubber-stamping the outcome of a formula.

It is also essential that panels contain individuals with comprehensive knowledge of the subfields including any, sometimes subtle, differences there may be between the way different sub-disciplines operate. We suggest that each panel will need to be quite large to cover all the areas.

The panels will be playing a critical role in the quality assessment part of the process, and should be actively involved right from the beginning of the research assessment exercise.

- 6 *Are there significant implications for the burden on the sector of implementing our new framework that we have not identified? What more can we do to minimise the burden as we introduce the new arrangements?*

The burden on the sector will be immense, especially at the stage of setting up the new system. Universities will understandably want to check their submissions, for example ensuring that the citation rate attributed to particular papers in the assessment exercise equals the citation rate institutions believe these papers to have.

The research assessment system is not just a measuring system; it is also, inevitably, a means of changing behaviour. A successful assessment system will be carefully designed so that it encourages positive behavioural changes, rather than negative ones. In order to minimise negative behavioural changes, it will be necessary for the system to include safeguards, which may result in an increased burden on universities.

- 7 *Do you consider that the proposals in this document are likely to have any negative impact on equal opportunities? What issues will we need to pay particular attention to?*

In addition to the strong concerns we have regarding the impact on early career researchers and postgraduate students, as discussed above, safeguards will need to be put in place to prevent unfair negative outcomes for a number of other groups:

- The proposals as they stand could penalise all those who have not worked full time on research-related work – including those who have family responsibilities or those who have chosen to broaden their experience through working outside of academia.
- Care will need to be taken to ensure that researchers from other countries, or researchers who have been working abroad, where there may be different habits in publishing, are not penalised.
- There are potential equal opportunity issues to overcome regarding researchers engaged in different types of research. For example staff engaged primarily in user-focused research should not be at a disadvantage when applying for posts because their 'type' of research would not be properly recognised and rewarded fairly under whatever research assessment system is in use.



## References

HEFCE (2007) *Consultation on the assessment and funding of higher education research post-2008*.

Royal Society (2002) Roberts review of the RAE: submission from the Royal Society.

Royal Society (2003a) Response to the Roberts review of the research assessment exercise.

Royal Society (2003b) *Supporting basic research in science and engineering: a call for a radical review of university research funding in the UK*.

Royal Society (2004) Submission to the House of Commons Science and Technology Committee follow-up enquiry into the RAE

Royal Society (2006) Response to the DfES consultation on the reform of higher education research assessment and funding.

## Group membership

Professor Adrian Smith FRS (chair)	Principal, Queen Mary, University of London
Professor Anne Dell FRS	Professor of Carbohydrate Biochemistry, Imperial College
Professor Dianne Edwards CBE FRS	Professor of Palaeobotany, and Head of School of Earth, Ocean and Planetary Sciences, University of Cardiff
Dr Jason Hall-Spencer	Faculty of Biological Sciences, University of Plymouth
Professor Andrew Hopper CBE FREng FRS	Professor of Computer Technology and Head of Department, The Computer Laboratory, University of Cambridge
Professor Richard Jones FRS	Professor of Physics, University of Sheffield
Dr Cristina Lazzeroni	School of Physics and Astronomy, University of Birmingham
Jeff Skinner	Director for Corporate Alliances, University College London
Professor Philip Wadler	Professor of Informatics, University of Edinburgh

## Secretariat

Nicola Berkley Science Policy Officer, Royal Society

## Any enquiries about this submission should be addressed to

Nicola Berkley  
Science Policy Section  
The Royal Society  
6-9 Carlton House Terrace  
London  
SW1Y 5AG

[nicola.berkley@royalsociety.org](mailto:nicola.berkley@royalsociety.org)

+44 (0)20 7451 2219