

Response to the consultation on the Research Excellence Framework (REF)

Background

1. Proposals for a new research assessment and funding framework for quality-related higher education research were announced in November 2007. The consultation document can be found at:

http://www.hefce.ac.uk/pubs/hefce/2007/07_34/

2. Responses to the consultation document were sought by 14 February 2008 and a response form with defined questions was provided.
3. The College's response to the consultation is provided below. As well as responding to the questions, we also prepared a letter in response since the questions were unduly restrictive.
4. The response was informed by discussions with the College Operational Research Committee, meetings with each of the Faculties, and a meeting with colleagues from Oxford, Cambridge and HEFCE.

14 February 2008

Michelle Coupland BA
Director of Planning

Research Excellence Framework consultation
HEFCE
Northavon House
Coldharbour Lane
Bristol BS16 1QD

Dear Sirs

HEFCE Consultation 2007/34: Research Excellence Framework

Imperial College London welcomes the opportunity to comment on the proposed Research Excellence Framework to assess and fund research in UK Higher Education Institutions. Whilst we recognise that the consultation and the proposals have arisen from previous Government decisions and consultations, we would urge HEFCE to encourage the Government to rethink the proposals since their implications will be significant, particularly in terms of the perception about the quality of the UK research base that any system will create. The College would be willing to support HEFCE in doing this and has volunteered to be a pilot institution.

We have responded to the specific questions posed and, in addition, we have a number of more significant general points which can be summarised as:

- The main policy objective must be that any system is able to sufficiently recognise and assess the highest quality research and to fund it accordingly. A desire to reduce the administrative burden should not be the main determinant of any system.
- The creation of two distinct assessment systems and the move away from a single common assessment framework is unnecessary and not supported.
- The consultation document and the supporting studies suggest that there are many issues that need to be considered and addressed appropriately before any new system is introduced. RAE2008 outcomes are well-timed and can be used for many years to come, thus any new system must be developed and tested for longer than is currently proposed and implementation must not be rushed.
- Any system is both an assessment and a funding mechanism, rather than merely being a means to enable HEFCE to distribute its funds for research.
- The outcomes will provide important indicators and messages about the quality of the UK research base and the subjects and institutions within it. Any new system should be developed with this in mind.
- Institutions should be free to determine who and what is included in the assessment process.

To elaborate further on these points, one of our main concerns with the proposals is that the main determinant seems to be a desire to reduce the administrative burden. The outcomes have significant reputational and financial consequences and thus it is reasonable to assume workload implications for institutions. The RAE is not a particular burden when compared with alternatives (e.g. Research Council peer review which, all things considered, is more time consuming and expensive). Much of the information returned in the RAE is important for normal day-to-day management purposes and would be collected and used irrespective of whether it is required to assess research. Certainly, and at least initially significant work must take place to check and validate information and prove the credibility of any new process.

The main objective of any research assessment system must be to sufficiently recognise, assess and fund the highest quality research on a selective basis. We would question whether one metric alone (i.e. citations) is able to capture and assess the full range of research and able to distinguish sufficiently between research which is good and that which is the very best. The outcomes must not lead to a narrowing of the funding differentials or the distinction between good research and the very best research. Further, HEFCE QR funding should be used to support research capacity and capability and to enable speculative 'blue skies' research. Using citations as proposed would tend to encourage 'safe' research, namely that which is likely to be cited or funded. It is therefore likely that really innovative work, which others would then develop, would reduce, thereby impacting UK competitiveness and exploitation.

As a funding mechanism, the results are used to determine future research funding for the subjects and institutions assessed. Any assessment mechanism must therefore be sufficiently prospective and able to identify and fund future potential. It is recognised that the system should be based on fact and therefore require objective, historical information rather than being purely subjective. However the time period proposed for using citation information is too historical and more retrospective than the information returned recently to RAE2008. Equally, and given the use of the results to determine future funding, it is essential that any system assigns credit to the institution employing staff on a given date rather than where the output was published. Each Higher Education Institution (HEI) must be free to determine who and what is included in the assessment.

The UK, and indeed the World, has become accustomed to over 20 years of RAEs meaning that RAE outcomes are recognised as an important measure and benchmark of the quality of the UK research base and of the individual subject disciplines and institutions within it. Separation into two distinct systems and to a higher level of subject aggregation than that to which we are accustomed will mean that important and comparable indicators of research quality within and between subjects will be lost.

The consultation fails to convince of the need for two separate and distinct systems which are developed at different times. The fact that citation coverage is more limited for 'non-science' subjects does not, in itself, provide sufficient justification for two different systems. Indeed, citations have their limitations within the 'science' system, since Thomson ISI coverage is limited for aspects of engineering and citations will fail to identify outcomes which appear in forms which are not able to be cited.

Citations could be a useful indicator for aspects of science, but not on their own. Instead they should be considered alongside other measures of quality and excellence (for example external fellowships, recruitment trends, research strategies) and assessed by peers. For some subjects, the pilots may show that citations could be used to replace the research outputs aspects of the RAE submission (i.e. the information returned on Form RA2). However, even then citations should form only one part of the assessment and should not be used alone. One system could therefore be developed for all subjects with appropriate adaptation for allowable subject differences. The role of peers would be integral to this process.

Any pilot system should therefore be extended in both length and scope of coverage. It should seek to test how and where citations could be used and to understand the limitations of these. Sources of bibliographic information in addition to Thomson ISI should be explored and alternative quality measures should be captured. The subjects that form the pilot should not be restricted to the science disciplines and 'non-science' should be considered alongside this with a view to creating a common assessment system for all with appropriate adaptation. Any outcomes should be compared with the results of RAE2008 to prove their credibility. A detailed and more lengthy period of careful development and iteration, working and developing the types of metrics to ensure their rigour, applicability and fitness for purpose is thus required.

The process must not de-stabilise RAE 2008 since its successful conclusion will provide meaningful outcomes to corroborate and test the new process and outcomes against. RAE2008 enables HEFCE and the higher education sector to have the benefit of not having to rush a system, instead being able to devote time to testing and validating measures and developing a system which is fit for purpose.

To conclude, a credible system is significantly more important and appropriate than implementation in 2009 and thus any new system should not be rushed. Instead any new process must be considered fully and apply (in varying forms) to all subjects. Its main objective must be to provide a robust and reliable system which is able to measure and assess quality in its varying forms and be able to distinguish sufficiently between different levels of excellence. We would urge HEFCE to broaden the scope of the pilots to look at all subjects and other indicators and delay implementation. Any new arrangement should have a steadily increasing influence beyond that period once it has been demonstrated that it is fit for purpose and its rigour has been accepted by the academic sector, by international bodies more widely, and by other stakeholders to the process. If the pilot highlights that the proposed system is not yet workable to accurately assess research quality, time should be taken to find an alternative and the RAE should not be replaced. At the very least, we would hope that the pilots are more extensive and have real influence on the recommendations rather than merely being a means to confirm what is being proposed.

We trust that these comments are helpful.

Yours faithfully

Michelle Coupland

Respondent's details

Are you responding:	On behalf of an organisation
Name of responding organisation/individual	Imperial College London
Contact name	Michelle Coupland
Position within organisation (if applicable)	Director of Planning
Contact telephone number	020 7594 7279
Contact e-mail address	m.coupland@imperial.ac.uk

Consultation question 1a: 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?

No. The proposed subject groupings are too broad and will fail to provide a robust measure and indicator of quality between subjects and institutions. Broad subject groupings mask particular pockets of excellence and do not provide information at a suitable level of disaggregation to be useful to inform the quality of disciplines. They may also disguise particular issues with the data and results.

We note that the consultation suggests that the number of groupings for 'non-science' subjects will be more than for science. We do not agree that there is a rationale for more aggregated subject groupings for those subjects assessed under the bibliometrics (i.e. 'science' system) and more disaggregated groupings for those subjects which are assessed under the proposed light touch peer review system.

We recommend therefore that the number of subject groups be kept as per previous RAEs. Certainly, for the pilot exercise, data should be examined at the RAE sub-panel (UoA) level to enable comparison with RAE2008 and RAE2001 results.

Consultation question 1b: Are there issues in relation to specific disciplines within this framework that we should consider?

Yes. We do not support the creation of two distinct systems and instead believe that a common assessment system with appropriate subject variations should be developed for all subjects. Disciplines do not fit neatly into the 'science' and 'non-science' classifications and the split will adversely affect work which spans the proposed divide. One system would thus serve to alleviate concerns about subjects which span the boundaries between the science and non-science framework.

The proposed subject mapping discourages multi- and inter-disciplinary work, both deliberate cross discipline work and where interdisciplinarity is intrinsic within disciplines (for example Engineering and Maths).

We have concerns that, to an extent, subjects seem to have been defined as science or non-science based on the availability and coverage of citation information. Specific issues also arise for subjects which have more limited coverage within Thomson ISI, for example Engineering and Computing and alternative bibliographic sources must be included.

Consultation question 2a: 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?

No, we do not agree that **on their own** bibliometric indicators provide a robust quality indicator. Bibliometrics could be used in varying forms for different subjects, but this needs to be tested fully through a **longer** pilot process and be compared with the outcomes of RAE2008 before a final decision can be made. For some subjects (namely those with full or near Thomson ISI coverage) they may be a suitable measure to replace the information previously returned on Form RA2 of the RAE submission, namely the assessment of 'outputs' and the need for peers to read them as part of previous RAEs. However, even then they should not be used alone and should not replace other indicators which help to identify, measure and assess research excellence. Most importantly, peer review should continue to form a key and important part of the process.

The issues and concerns with using citations and their possible impact on publication behaviours are well documented. It remains an open debate whether they in fact measure impact rather than quality, whilst an inevitable consequence is that publication practices will vary and concentrate towards producing work which is popular and likely to be cited and contained within Thomson ISI. The over-concentration on journal outlets will thus remove the incentive to produce applied, translational policy-related and practice-based research. This seems to be at odds with Government policy objectives to encourage and support work of direct relevance to economy and society.

Whilst Thomson ISI may be appropriate for some subjects, it has inadequate coverage for others. Thomson ISI is neither robust nor comprehensive enough, being particularly problematic in Computing and Engineering. Hence it will be particularly important that other alternative indicators are developed and used (e.g. Google Scholar and SCOPUS, Proceedings of the ACM). This would have the advantage of also removing dependence on a single data source.

Since the system will be used to determine future research funding, it is important that the methodology is able to adequately identify and fund future potential. There is thus a fine balance to be achieved between capturing citations over a period which enables the impact of the output to be felt and which provides a sufficient indicator of then or future quality. The time period proposed may thus be too retrospective and historical, meaning that a measure of potential will be lost.

In addition, as a measure and funder of potential, the citations must follow staff and be attributed to institutions employing staff on a given date rather than to the institution at which the output was published. The contributions of authors to outputs should also be a consideration.

If, and where, bibliometrics are used, then they will need to be reviewed on a frequent basis to ensure that they remain fit for purpose (e.g. in light of developing multidisciplinary research, changes in publication practice).

Consultation question 2b: Are there particular issues of significance needing to be resolved that we have not highlighted?

We note that the consultation proposes that self-citations be removed from the process. Self-citation in the form of common authorship between cited and citing work can be a positive indicator of strong collaboration. Removal of all outputs containing self-citation will have a strong negative impact on collaborative work, which is essential towards coherence. Excessive author self-citation would be expected to be removed through the refereeing process prior to publication.

Alongside the removal of self-citations, the relative contribution of co-authors is an important consideration. Practices vary in recognising and assigning credit for co-authors, but methods to capture this and analyse the outcomes should be tested. For multi-authored outputs the pilot will need to consider the number of times to credit a publication and, if authors are from multiple HEIs, where to credit the output.

Normalisation between subject groups is essential. Further work needs to take place at field boundaries and with journals that map to multiple fields, i.e. whether it is suitable to use a field average. Also, it should not be assumed that the ~170 ISI fields are fit for this new purpose. We recommend that significant work be undertaken to establish whether it is appropriate to categorise outputs in this way.

Bibliometrics, namely citations, will need to be tested thoroughly before their introduction can be considered viable and appropriate. It is thus important that the pilot exercise is extended in length and in coverage and seen as a means to test thoroughly what is being proposed rather than merely being a means to confirm that the proposals are acceptable. The College is very happy to assist HEFCE in developing and testing any proposals to ensure a credible system.

The pilots would need to test and consider fully:

The means and mechanisms to capture and validate the citation information. This includes validating the accuracy of author names, author addresses, resolving issues where authors publish in group names, assessing how to apportion credit correctly to authors.

Data sources and coverage. This includes identifying alternative publication sources and citation sources and testing the credibility of using them. This is of particular importance for subjects with more limited Thomson ISI Citation coverage and must include the non-science subjects.

Comparison and assessment of outcomes. RAE 2008 outcomes must be compared with the results arising from the pilots and peers involved to understand the reasons for differences. Importantly, an assessment of the impact of the proposals on the UK research base must be considered, since the standing of the UK research base must not be undermined.

Modelling of the implications of different behaviours. This includes assessing the impact of different patterns of citation behaviours (i.e. a group with a fairly consistent citation rate against once which has a spread of rates). The system would also need to look at the impact of assigning staff to outputs and subject groups (particularly where they span subjects) and how to treat journals which span fields. Self-citations and co-author contribution are also important considerations.

Consultation question 3a: What are the key issues that we should consider in developing light touch peer review for the non science-based disciplines?

We do not agree that a separate system should be adopted for the non-science based disciplines. Rather science and non-science should be considered alongside each other and be developed and implemented at the same time. Treating them separately and dealing with

them at different times will at best create confusion and at worst cause inconsistency between subjects and institutions. It is recognised that different indicators with different weightings are needed between subjects but there is no supporting evidence to suggest that a broadly similar system could not be used for all. A common framework for all disciplines would have the advantage of alleviating some of the concerns about the differing treatment of subjects and fields which span the boundaries between disciplines.

The development of light touch peer review system should look at streamlining the current RAE for all subjects. A key component should be to assess how the data collected is used to inform the assessment in order to prevent requests for superfluous data which will not be used by the panels and will not be meaningful to the process.

Consultation question 3b: What are the main options for the form and conduct of this review?

As per our response to question 3a above, we believe that in order to maintain the comparability of disciplines within and between HEIs, assessment should, taking into account appropriate differences between subjects, employ a common framework with all subjects assessed within the same timescale. Difficulties with defining and capturing bibliometrics for disciplines labelled as non-science should not be sufficient justification for a separate system. As with the RAE, simultaneous assessment by peer review must continue for all subjects.

Consultation question 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.

Citations cannot be used alone. Instead they should be supplemented in varying forms by other indicators within a broad framework. Of crucial importance is that any measure should genuinely be able to measure and reflect excellence and quality in its varying forms. Possible additional measures include:

Other 'output' indicators which reflect the full range of user and applied research that is not produced in forms able to be captured within Thomson ISI (or other bibliographic sources). This includes patents granted, spin out companies established, other aspects of technology transfer, confidential and expert witness reports.

Measures of Esteem and external recognition. These include external, prestigious fellowships, significant honours and awards, plenary and key note addresses, memberships of committees, recruitment patterns and trends.

Sustainability and vitality measures which include research student registrations and progression (including subsequent destinations once graduated), key investments in research (for example for research projects, for buildings, for sponsored posts), number of PhDs awarded, research income from external grants and contracts, most notably those that are peer-reviewed.

Future plans, including mechanisms to develop research and support and nurture staff.

The majority of the items above are included in the current RAE assessment and could continue to be used where appropriate. Institutional representatives, HEFCE and the Higher Education Statistics Agency (HESA) must work together to develop appropriate measures.

Such indicators cannot be considered alone and must be considered by experts in the field.

Consultation question 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?

No. The current proposals envisage panels playing only a minor role, namely to oversee the initial implementation of a purely bibliometrics based approach. We are unconvinced that the metrics proposed so far obviate the need for a light touch review from experienced researchers within the disciplines on an ongoing basis.

Panels should be used consistently across all subjects to inform the process and assess the full range of indicators provided. Their role must be to interpret the results of the metrics based assessment and to bring their subject expertise to the exercise. This is of particular importance in the initial stages. To be workable, this would require panels at a lower level of aggregation than is proposed.

Consultation question 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?

An accurate and robust assessment of quality is essential to maintain and develop the standing of the UK research base and to represent its quality on the international arena.

The outcomes have significant reputational and financial consequences and thus it is reasonable to assume workload implications for institutions. The RAE is not a particular burden when compared with alternatives (e.g. Research Council peer review which, all things considered, is more time consuming and expensive). Much of the information returned in the RAE is important for normal day-to-day management purposes and would be collected and used irrespective of whether it is required to assess research.

Implementing two distinct processes at different times will increase the burden significantly, thereby providing further justification for one framework across all subjects.

Consultation question 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?

HEFCE will need to consider and monitor any potential equal opportunities implications.

Of particular note is that a mechanistic process via the proposed algorithm will fail to include a means to consider specific local circumstances which may have impacted on research outputs and productivity. Inclusion of peer review in the process would enable this. Further, a separation into science and non-science systems will introduce additional and potentially complex, managerial and equal opportunities issues within departments and for staff who straddle the subject division.

The framework does not recognise emerging areas and disciplines, or staff who are establishing their research record, being biased towards established researchers and disciplines. Researchers may be perhaps encouraged to reduce their publication volume and only publish work which is likely to be outstanding and highly cited. This will be a particular issue for newer researchers including, for example, students who wish to publish papers from their PhD thesis.

Consultation question 8: Do you have any other comments about our proposals, which are not covered by the above questions?

Please see our comments at the beginning of this document. It is important that institutions are able to continue to select who and what is included in the assessment process.