Research Excellence Framework
Consultation on the assessment and funding of higher education research post-2008 (ref 07/34)

Are you responding:	• On behalf of an organisation
Name of responding organisation/individual	**The Council for the Mathematical Sciences** (comprising the Institute of Mathematics and its Applications, the London Mathematical Society, the Royal Statistical Society and the Edinburgh Mathematical Society and Operational Research Society).
Contact name	Martin Smith
Position within organisation (if applicable)	Secretariat
Contact telephone number	020 7927 0803
Contact e-mail address	cms@lms.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?

1. **The Council for the Mathematical Sciences welcomes the exclusion of mathematics and statistics from the group of 'science-based disciplines'** that are to be assessed by a metrics-driven approach.

2. It is not clear from the consultation document which process is expected to apply to **Operational Research**, but we hope that it will be subject to a 'light touch process'. Currently Operational Research tends to be submitted through mathematics, business and management and computer science.

3. **Cross-disciplinary research remains an issue**, particularly where the work crosses the science/non-science divide. How this is dealt with is critical, and is something that needs to be addressed.

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

4. There are many sub-areas of Mathematical Sciences (and presumably other disciplines in the "non-science" group) where individual researchers might equally well be included in one of the six 'science-based' subject groups. This illustrates the potential for game-playing in this system.

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?

5. The Council for the Mathematical Sciences believes that **bibliometric indicators will be very problematic for assessing research quality in mathematical sciences**, and is sceptical about using such measures as the basis of assessment for other subjects.

6. We do not believe that the proposed bibliometric indicators measure research quality in any reliable sense, or are robust. If it is essential that panels use bibliometric indicators they should be instructed to do so with caution.

7. Bibliometric indicators necessarily refer to research that has appeared in a specified previous time period. Different time periods will be appropriate for different disciplines.

8. The suggested bibliometric approach for science-based subjects **implicitly assumes that citation levels, time-scale and practice are the same *within* each sub-area**. We doubt that this is true and it certainly would not be true at a similar level of aggregation for mathematics.

9. It is also clear from *What do citation counts measure? A review of studies on citing behaviour* (L Bornmann and H-D Daniel, to appear in the Journal of Documentation) that no-one really knows why researchers cite the papers they do, and **a worryingly high proportion of papers cited (up to 40%) appear to be so-called "perfunctory" citations**. The nature and purpose of citations can vary considerably from field to field within a subject.

10. Problems with the Leiden recommendations include:

    - Availability of data - We are concerned by the prominent role that Thomson Scientific, a private company, will have on UK research evaluation. Over-reliance on one private company to provide essential information for determining funding of HEIs is dangerous and could become very costly.

      It is possible that, as a result of this reliance, the UoAs will need to be categories chosen solely by Thomson.

    - Accuracy of data - As noted on p 17 of the Leiden Report, there are many errors in Thomson CI data which would need to be corrected. This is likely to be very difficult.

There is also no reason to assume that the world citation rate c_f is accurate, or even that it is inaccurate to the same extent for each discipline given that citation practices vary widely. It is not clear who would be responsible for checking this information.

11. Another important issue is that the proposal seems to be based on "citation rates per paper". This appears to disadvantage a group with x well-cited papers and y less well-cited ones, when compared with a group with x well-cited papers and nothing else. **Whatever is introduced, it must not penalise an academic for having written a larger number of papers.**

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

12. Of greatest concern is that **UK research and the hiring system would be skewed towards 'fashionable' sub-areas with high citation rates within a discipline**. The result will be great pressure from within universities on their staff not to undertake research in small fields that are currently unfashionable and not to move into new fields.

13. Rather than "publishing only the best papers" [Leiden Report, p36], **researchers are more likely to publish their best work in CI journals, and their lesser work in non-CI journals** to avoid lowering their average citation rate.

14. **The concept of removing 'Self-citations' does not seem to be well-defined** in the context of multi-author papers. It is not clear how a multi-author paper citing one of the author's previous papers would be treated, or even whether one author citing a paper which he contributed to with others counts as a 'self-citation'.

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

15. **It is very important that mathematical sciences receives due attention** when the separate 'light-touch' process is developed for the arts and humanities subjects.

16. Our response to this question focuses on key issues for the mathematical sciences and the factors that support HEFCE's decision to exclude mathematics and statistics from the 'metrics-driven' approach for science-based subjects. **We would be very pleased to discuss these further with HEFCE during the development of the light touch process for the 'non science-based' process** (see response to question 3b). We do not believe that it is appropriate to rely on the HEFCE/AHRC group's report when deciding how to assess research in the mathematical sciences.

17. Light touch or not, important funding decisions will be made on the basis of this assessment. As noted even in the Leiden Report "Peer review is and has to remain the principal procedure for judgment of quality." (p36). It may be difficult to achieve this without input from a substantial external group of experts if the panels are to be broader than at present.

18. **Citations in mathematics are relatively slow to emerge and low in number**. Ideas can continue to be relevant for decades, and there are many fields that have found totally unexpected applications long after they started. "New" research fields may take twenty years to mature. The overall numbers of active researchers in some fields is very low, and the publication process is relatively slow owing to detailed refereeing processes.

19. Grant income is not a robust metric of research excellence in Pure Mathematics, although it is believed to be a better measure of research vitality in some other mathematical sciences subjects

20. **Citation practices vary widely according to subject and even by field _within_ a subject** - for example, in mathematics it is often the case that papers are cited more to provide information for the benefit of the reader, rather than to assign credit to previous authors, and the phrase "see [X] and the references therein" (where [X] is a recent paper or one which contains a good bibliography) is extremely common. This contrasts with clear rules governing citation practices in experimental science subjects.

21. **Average citation counts are likely to vary substantially from field to field _within_ a Unit of Assessment**; for example, this is certainly the case in the many sub-disciplines within Pure Mathematics (number theory, analysis, etc.), where average citation counts can vary by more than an order of magnitude. Averaging only at the resolution of 'Pure Mathematics' (for instance) would disadvantage departments with a majority of researchers in sub-disciplines that naturally fall below the broad average for Pure Mathematics.

22. **The number and length of papers produced also varies greatly by field within the subject areas covered by the mathematical sciences.** Four papers may represent anything from 10% to 100% of a researcher's output during the period depending on the field, although it is desirable for a researcher to continue to be able to select their best work for assessment.

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

23. The CMS working group charged with formulating a response to this consultation has considered many different options, but is not yet in a position to provide a complete proposal for an assessment process. **The Council for the Mathematical Sciences would like to offer to work with HEFCE to design a light-touch process that is suitable for the mathematical sciences**. We note that the HEFCE/AHRC group will not have considered aspects particular to the mathematical sciences, and we would be pleased to work with HEFCE to ensure that there is an appropriate focus on mathematical sciences during these important developmental stages.

24. We welcome paragraph 87 of the Government's response to the 2006 consultation '_Reform of Higher Education research assessment and funding_':

> _"Assessment will be based on a rating derived from a basket of metrics containing research income and postgraduate student metrics, together with expert review of selected research outputs, and with expert advice_

> *on the weighting of all these elements. The review of outputs will be significantly less burdensome for higher education institutions and their researchers than the current RAE process."*

25. We agree with the HEFCE/AHRC group that a less burdensome approach to assessing output quality could involve either sampling of submitted outputs from individual researchers, using larger bodies of reviewers such as Research Council peer review colleges, or a combination of these – these are among the possibilities that we would look at as part of a future discussion. We might also be interested in considering whether a structured reduction of the reading of outputs might usefully be achieved by asking panels to read only the abstract to each paper or the first two pages; further reading could take place where panel members disagree by more than a set amount. It is important, however, that whatever peer review mechanism is adopted does not introduce unacceptable variability into the assessment process.

26. A 'light-touch' procedure must reduce burden on departments as well as panels – **this could be achieved by reducing the effort which departments have to put into producing narratives such as the RA5** (see also our response to question 6). One possible replacement would be a list of esteem factors that the panel could grade by their significance, but this needs further thought.

27. **If there are to be fewer panels than current Units of Assessment then it is important that efforts are made to ensure that the panels have a good overview of the subject**.

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.

28. The three main criteria (publications, external funding, research students) seems about right, with the latter two providing a measure of group vitality and arguably some reflection of applied research quality. Precisely how one defines each indicator is less important than how they are combined (a matter on which subject-specific experts should advise in each case). Nevertheless, we remain sceptical that any combination of output metrics, grants income and student numbers measures 'quality' and advocate the use of peer review moderation of any such process in the mathematical sciences.

29. The impact and quality of public policy advice and advice to business and industry does not seem to be amenable to simple quantification, although a list of esteem factors could tackle this – see our response to question 3b.

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?

See elsewhere in our response.

30. The bulk of work occurs in HEIs but the proposals seem to focus on decreasing burden on the panels.

31. **Validation of the HEI's own citation data is likely to be very burdensome and difficult to implement**, much as HEIs will attempt to take this on to improve their chances of a good score - Evidence Ltd's report notes that linkage of articles to Oxford University was increased by 40% by careful data checking (p32).

32. To an extent there is some inevitability that HEIs will spend large amounts of time attempting to optimize their tactics – **some of the burden is self-imposed given the significance of subsequent funding decisions** and the comparative rarity of the event.

33. **Burden could be reduced by decreasing or removing the need for narratives such as the RA5**; it was the experience of some members of the CMS working group that a large amount of time was spent refining these for relatively little attention from the panel.

34. **Replacing a large-scale quinquennial exercise with annual fine-tuning (with suitable smoothing of results) could potentially help** with reducing self-imposed burden on HEIs as the consequences would in theory be less serious. As one academic has phrased it: "elementary control theory suggests that small changes in response to timely feedback are preferable to the occasional application of a large sledge-hammer". However, it is difficult to envisage a process that would be 'light' enough to undertake every year.

35. The detailed design of any metric-based methodology has to be able to adjust for career breaks which, in practice, still impact more on women than on men. Similarly, there is the danger that the proposed process will have the effect of discouraging institutions to take on early career researchers. **Visibility in a research area appears to be especially important for "perfunctory" citations** (which can make up to 40% of total citations, according to *What do citation counts measure? A review of studies on citing behaviour* (L Bornmann and H-D Daniel, to appear in the Journal of Documentation). **Established academics who receive frequent invitations to speak at conferences will naturally be at an advantage**. A researcher's availability to speak at conferences is affected by family commitments. The research quality of young academics (or those who have had career breaks) can be directly compared with that of established academics under the current RAE, since they can submit fewer than four pieces of research. Comparison seems to be more difficult under the proposed system and use of citation data would be discriminatory due to the time lags in the system. If citation data are to be used alongside other metrics with a

light touch peer review then special provision would need to be made for early career researchers.

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

36. Since it is likely that output from the pilot exercise could be "tuned" to give a broadly similar quality profile for an institution to that from RAE2008 (through judicious use of enough "fine tuning" parameters), **pilot institutions will have a distinct advantage over their (untuned) peers** - who could potentially lose or gain significant amounts of research income from the REF compared to the RAE.

37. Fine-tuned metrics that happen to give a good fit to a light touch peer review procedure in 2013 are not necessarily better than other metrics, and their adoption might well lead to serious distortions at a later date. A good 'fit' with 2008 data does not imply that the process is fit for purpose, and the process will inevitably lead to changes in behaviour.

38. The CMS is grateful to Graeme Rosenberg (Policy Officer, HEFCE) for his assistance to our working group.

Council for the Mathematical Sciences
February 2008