

Annex A

Consultation response form for the Research Excellence Framework

1. Respondents should complete the form below.
2. Responses should be e-mailed to refconsultation@hefce.ac.uk by **Thursday 14 February 2008**. HEIs in Northern Ireland should send a copy of their response to research.branch@delni.gov.uk
3. Institutions wishing to express an interest in taking part in the pilot of the bibliometrics indicator should e-mail their details to refconsultation@hefce.ac.uk by Thursday 31 January 2008.
4. We will publish an analysis of responses to the consultation. Additionally, all responses may be disclosed on request, under the terms of the Freedom of Information Act. The Act gives a public right of access to any information held by a public authority, in this case HEFCE. This includes information provided in response to a consultation. We have a responsibility to decide whether any responses, including information about your identity, should be made public or treated as confidential. We can refuse to disclose information only in exceptional circumstances. This means responses to this consultation are unlikely to be treated as confidential except in very particular circumstances. Further information about the Act is available at www.informationcommissioner.gov.uk.

Respondent's details

Are you responding:

(Delete one)

On behalf of an organisation

As an individual

Name of responding organisation/individual London School of Economics and Political Science

Contact name Professor Sarah Worthington

Position within organisation (if applicable) Pro-Director (Research and External Relations)

Contact telephone number 020 7955 6389

Contact e-mail address s.worthington@lse.ac.uk

Consultation questions

(Boxes for responses can be expanded to the desired length.)

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?

By way of general preface to this response, LSE reiterates its endorsement of the Russell Group response, to which it contributed. In particular, the new regime must be fit for purpose and sufficiently rigorous to command the respect of the research community nationally and internationally. We support the clear and more limited purpose of the REF as a mechanism by which QR funding can be distributed based on excellence (as determined by a proxy measure that is fit for purpose). The REF does not need to provide a detailed, primary assessment of research quality for use by institutions as a basis for their own internal management or funding decisions.

LSE, as a purely social science institution, has no strong views on how the science-based disciplines are divided for the purposes of REF. We understand that the methodology required to draw up the bibliometrics necessitates the consolidation of large bibliometric datasets in order to derive statistically significant results. LSE can, however, envisage that such broad groupings may not be well received by those in the STEM disciplines. The inevitable averaging across various sub-disciplines included in these large groupings may favour certain discipline-mixes that exist in some HEIs/research groupings over those that exist in others, even where the quality is objectively equivalent.

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

We assume this question relates to the internal divisions within the STEM subjects. On that basis, LSE strongly supports the decision to include Mathematics and Statistics in the non-STEM peer review methodology, for all of the reasons put forward by these two subject communities during the 2006 consultation exercise. Psychology will need to be considered under both the bibliometrics and peer review systems. There will undoubtedly be other subjects, either newly emerging or interdisciplinary, which will straddle the STEM / non-STEM divide (health policy and bioscience for example).

But the wider issue is the more important one. We feel there should be no sharp division in the sector between STEM and non-STEM subjects. The pilot exercise should explore an overall framework for research assessment within which principles can be extended to other families of research (e.g. social sciences, humanities, creative and performing arts). This is essential if the UK is to retain a reputation for excellence in research across the entire discipline mix. A two-tier approach, with

different modes of assessment for different classes of research, will inevitably lead to allegations of different quality standards between them. In addition (although LSE will not be so affected), under current proposals most HEIs will need to run two processes, and make difficult decisions if one is seen to be more advantageous than the other. This added burden on HEIs forced to manage two schemes on two different timelines, and deal with the inevitable game-playing and personnel difficulties, seems counterproductive.

REF needs to provide a single overarching framework for assessment of research excellence, not a division between two systems that will prove difficult and divisive for institutions and may create barriers to collaboration and a negative impact on multi- and inter-disciplinary research.

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?

This question relates only to the bibliometric aspect of the proposed REF, but metrics (including bibliometrics) are a proxy for quality, not an actual measure of quality, so a robust REF may need to combine these metrics with peer review to verify and moderate outcomes, even in STEM subjects. In addition, the balance between different metrics needs to be one that provides an adequate proxy for quality. This is likely to favour bibliometrics over metrics measuring research funding or PhD completion rates, but these metrics and any balance between them need to be rigorously tested. A one year pilot seems inadequate for this purpose, and in any event does not permit proper testing against RAE2008 assessments and outcomes.

Bibliometrics themselves are a crude measure of the *use* by researchers of research outputs, not necessarily the quality of those outputs (see the Leiden Report's proposed exclusion of review papers and data sets). Such bibliometrics reinforce 'mainstreaming' effects. Widely cited articles often reflect fashionable topics, not necessarily the truly ground-breaking, exciting and experimental work which should be rewarded under an assessment of research excellence.

Even at the level of production of papers, there may be perverse incentives both in relation to publication at all and in relation to joint publications with junior colleagues. HEFCE seems settled on a 'citations per output' key metric. If this is averaged for the unit of assessment, then researchers within that unit will be motivated to withhold publication of all but the best papers, declining publication of a descriptive or reporting nature, or declining to collaborate with research students or junior colleagues. The impact might be reduced by measuring total citations over the unit of time for the unit of assessment, or the average citations per paper of the top 50% (or other) of outputs. It would add to the individual burdens of institutions if this new regime required selection of

outputs by the HEI—and this is in any event arguably counterproductive in assessing QR entitlement.

There is a dramatic variation in the order of magnitude of bibliometrics between disciplinary fields, often related to the number of journals per field and behavioural differences. For instance, surveys are highly cited but not included in engines such as the Web of Science, so skewing the bibliometrics. Additionally, there are wide variations in citation timelines between disciplines, whereby many non-STEM works may not be cited for several years, whereas some STEM subjects, such as the medical sciences, are cited much earlier. An assessment based mainly on bibliometrics will have to carefully consider citation patterns in each field of research if it is to produce meaningful data.

The Leiden Report commissioned by HEFCE disavows the use of journal impact factors, and we agree with the reasoning. Equally, we consider that other attempts elsewhere to produce journal rankings for different disciplines (e.g. the European Reference Index for the Humanities, produced by the Humanities in the European Research Area project for the European Science Foundation) are even more flawed.

On metrics other than bibliometrics, using research income as a quality indicator for determining QR fails to acknowledge excellent unfunded research and undermines the dual support system by effectively double-counting Research Council income. Research income volumes are discipline-dependent, so at the very least they must be treated as such if this is to be used as a metric. In any event, research grants provide a measurement of inputs, not outputs, so correlation with quality of the outputs needs to be tested. Moreover, competitive research grants are frequently given to ‘popular’ topics, so may not necessarily incentivise or correlate with quality in outputs. Finally, including Research Council grants in the metrics will inevitably increase the number of grant applications, leading to an increase in the overhead costs of research council grants.

If postgraduate student numbers are to be used as a quality indicator, these will need to be discipline-specific (i.e. ranked against the expected average numbers of students in each subject area). It is unclear in the consultation whether postgraduate student numbers will count for quality or just in the QR funding calculation.

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

It is unclear how frequently HEFCE plans to run the bibliometric exercise for STEM subjects under REF. If it is to be run more frequently than peer review for the non-STEM subjects, this could lead to disparities in QR funding between disciplines and to complications for HEI planning and management. There will also be considerable confusion for non-academic stakeholders if the outcomes of REF are published

piecemeal in different years. There will be confusion in the data used by third parties in constructing national and international league tables for UK HEIs.

It is currently unclear what eligibility criteria for staff will be applied under REF. The consultation implies a much wider definition of eligible staff than that used under RAE. The consultation also indicates that HEIs will be responsible for selecting staff for submission. We suggest that the preferable approach is to consider all eligible staff, however eventually defined (see below).

LSE believes that the proposed bibliometrics should include all research outputs from eligible staff within an institution, and that outputs should be assigned to the employing HEI at the date of publication, rather than at the census date. These data would be easier to collect and verify under the bibliometrics system, would be fairer and more transparent, would considerably reduce the burden on HEIs of selection of both staff and outputs imposed by the current RAE, and would reduce the 'RAE transfer market'. It would also reward HEIs for attracting and nurturing talent, even where that talent has migrated between HEIs during the assessment period. An approach that assesses all research outputs from all eligible staff within an institution will eliminate the need for HEIs to 'cherry-pick' research-active staff. This will be easier to manage and significantly reduce 'game-playing' by HEIs. These two approaches should apply equally to STEM and non-STEM subjects.

Finally, LSE urges that the same principles should apply to STEM and non-STEM subjects, and that the pilot system should not simply focus on the STEM subjects but should examine how REF can develop coherent principles of assessment that will apply across all disciplines as far as possible.

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

LSE strongly argues that the timetable for a full consultation on peer review should be brought forward significantly. Deferring consultation until late 2009 means that the new system will not be known until 2010 with, presumably, data collection taking place in 2012 for assessment in 2013. HEFCE should be running the consultation on peer review in parallel with the consultation on bibliometrics, given that peer review will undoubtedly be more labour-intensive and burdensome than the collection of metrics alone.

LSE repeats its view that there should be no sharp division in the sector between STEM and non-STEM subjects, and that the pilot exercise should explore an overall framework for research assessment within which principles can be extended to other families of research (e.g. social sciences, humanities, creative and performing arts). It urges that the peer review system should not lead to non-STEM subjects being viewed as the 'poor relations' of the STEM subjects. Peer review should be, as far as possible, comparable to bibliometrics both in terms of reducing administrative burden and in the

robustness of its systems and outcomes. Peer review for non-STEM subjects should allow for international benchmarking, as bibliometrics will for STEM subjects.

The number of panels needs careful consideration. If panels are too large it will be difficult for HEFCE to find pockets of excellence. Following the 2006 consultation on the RAE, HEFCE found it impossible to reduce the number of units of assessment, which remained almost the same in 2001 and 2008. Its plan to reduce the number of peer review panels for 2013 AND reduce the burden on panel members seems therefore to be highly ambitious. Reducing the number of panels will not itself reduce the burden or the costs of peer review and could undermine the credibility of the system.

The implication in paragraph 53c is that the non-STEM subjects are immature and need to 'catch up' with the STEM subjects in the development of bibliometrics. In fact, non-STEM disciplines all have their own epistemologies and will not all lend themselves to the same narrow bibliometrics proposed for the STEM subjects. Some subjects (such as economics) would lend themselves to bibliometrics, while others which focus on producing books (such as history and anthropology) will not.

The RAE2008 panels should be consulted on what data they found most useful (and least useful) and their feedback used to help determine how to make the peer review light touch.

It is imperative that the 'light touch peer review' should significantly reduce the burden for HEIs, not just for HEFCE and panel members, and should be sustainable. A number of options for achieving these aims are set out in the response to 3b below.

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

Data collection could be revised and considerably reduced. For RAE2008 HEIs were asked to provide either a DOI or a hard copy for all outputs and yet were asked for a staggering amount of detailed information on publications, which the provision of a DOI and/or hard copy rendered unnecessary. Greater use could be made of existing data sources such as HESA returns and the Research Activity Survey to prevent the duplication of data on research income and postgraduate students. The data collection software should be designed to allow HEIs' own systems to provide data more directly, without the need for large amounts of manipulation.

The equivalent of the RA5 could be eliminated. The requirement for a narrative section under the present RAE is burdensome for both HEIs and panel members, and the RA5 will become increasingly meaningless if larger panels are established. Esteem factors, if deemed an appropriate indicator of research quality, could be captured using a simple form.

If HEIs are still required to select research outputs (contrary to the LSE's suggestions in (2b) above), or if 'light touch peer review' requires some form of output selection (see below), and given that the census period for REF is shorter than for previous RAEs, HEFCE may wish to reduce from four the number of outputs required per research-active member of staff. This may be particularly appropriate for those non-STEM subjects where books are the prevalent form of output.

HEIs' own publications catalogues/institutional repositories could be used to provide data on a much wider range of research outputs, with 1 or 2 outputs of special significance per academic highlighted for peer review. While panels would presumably not welcome a complete list of all the outputs produced over the census period (even if all of these were included in the bibliometrics computations), a snapshot of the HEIs' research outputs in a given year may prove informative and would mirror the bibliometrics process more closely.

As we recommended for STEM subjects, research outputs under any peer review system should be attributed to the HEI which employed the member of staff at the point of publication, not at the census date.

If the goal is to have the same set of principles governing assessment of research quality across the disciplines, then the trial period should explore the possibilities for some limited use of metrics for the non-STEM subjects, especially those which would best lend themselves to this (mostly the more quantitative social science subjects). Such an approach would need to recognise that some social science research can take 10 – 15 years to be fully recognised and widely cited, so the citation window for each discipline needs to be carefully considered. It follows that careful consideration needs to be given to the use of this type of data to measure present research quality in an HEI and to inform future funding decisions.

LSE is currently looking into the potential use of metrics in the social sciences and will be able to provide evidence at a later stage in the consultation process.

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.

Public engagement and 'user value' have been mentioned as important features of the value for money in research coming from HEIs. REF could arguably find a way to incorporate bibliometric data from user-focussed journals and other types of research outputs, rather than focussing too narrowly on purely 'academic' journals. However, this may prove to be a fruitless exercise in trying to quantify the unquantifiable. In addition, the time delays in delivering user value for most quality research is probably random both

within and across discipline areas, so compounding the less random problem of citation lifetimes in traditional bibliometrics.

In any event, if REF is simply to be a light-touch means of allocating QR income to HEIs based on research *quality*, then economic user value should not form part of the assessment.

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?

LSE has no specific comments to make on this issue.

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?

The cost of management of bibliometric data may be significant, both for HEFCE and for individual HEIs. The proposed move to bibliometrics will entail HEIs having to develop more sophisticated institutional repositories and internal mechanisms to ensure the capture of all research outputs. For some HEIs this will be more labour-intensive and costly than HEFCE seems to envisage.

HEIs with both STEM and non-STEM disciplines would have to manage two different assessment systems in sequence. It is unclear whether this particular burden would be reduced even if the two systems were timed to run in parallel. There are also HR implications for how the two systems would affect recruitment, promotions and research support.

Given that the results are likely to be much finer grained under peer review than under metrics, there are potentially divisive issues for the external perception and internal management of HEIs in the UK.

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?

At the level of operating principles, should HEIs be left to devise their own equal opportunities codes, or should HEFCE provide a template code for HEIs to adopt as appropriate?

Early career researchers are likely to be disadvantaged under bibliometrics in comparison with peer review. The latter system is potentially able to make adjustments for the career development stage of individuals, as in RAE2008.

As bibliometrics is a blunter tool than peer review, the effects of taking a career break could be more damaging in some subjects, depending on the citation patterns in each discipline.

It is not clear how individual staff will be affected by REF, particularly when their research straddles STEM and non-STEM subjects.

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

Whatever the final version of REF, it must continue to command the confidence of the research community, external investors in research, and our international competitors. There are such substantial differences between REF and the earlier RAEs that any trial period should be long enough to enable the rigorous testing of the robustness of the proposed new system. There is time to do this, since RAE2008 is available to determine funding allocations on an already accepted basis for the test period.

LSE repeats its objection to the proposed sharp distinction between STEM and non-STEM subjects. There is no sharp divide between subject groups, but a spectrum within which some (but not all) STEM subjects are more amenable to metrics as a proxy for quality in the assessment regime. More importantly, having two different systems running to two different timetables is likely to lead to negative quality perceptions of the inevitably lower-funded non-STEM subjects and a reduced perception both nationally and internationally of the overall quality of the UK's research base.

LSE regards the use of some form of moderating peer review as essential in identifying research quality, even in the STEM subjects.

REF will inevitably create a different incentive system from the RAE, and this should be recognised and monitored both generally and across disciplines to identify any discipline idiosyncrasies.

The intention in replacing the RAE with a lighter-touch system is to reduce the burden and cost of the exercise while retaining its robustness and respect. The scale of savings should be assessed over the trial period, as well as the effectiveness of the proposed assessment mechanisms.

While LSE acknowledges that the timeframe for REF was dictated by the government, it is nonetheless unfortunate. The huge data collection exercise within HEIs and the work now being undertaken by the panels for RAE2008 could be used more effectively to permit a proper trial period of the proposed new system, thus ensuring it attracts the confidence of the sector. In addition, HEFCE should utilise the data collected for RAE2008 and any learning from the panels' experiences, to modify and test the proposed new REF.