

Response ID ANON-CVDF-YW8B-D

Submitted to **Research Excellence Framework review: call for evidence**

Submitted on **2016-03-23 16:36:41**

Introductory Section

A Name

Please enter here:

Professor Dorothy Bishop, FBA, FRS, FMedSci

B Email

Please enter here:

dorothy.bishop@psy.ox.ac.uk

C Address

Please enter here:

Department of Experimental Psychology, University of Oxford, OX1 3UD

D Name of organisation

please enter here:

Council for the Defence of British Universities

E Please tick the box that best describes you as a respondent to this consultation

Representative Body

F If you selected 'Individual,' please describe any particular relevant interest; teaching staff, student, etc. 'Please expand or write NA if not applicable.

Please enter here:

NA

G If you selected 'Other,' please give details. 'Please expand or write NA if not applicable

Please enter here:

NA

Section 1

1 What changes to existing processes could more efficiently or more accurately assess the outputs, impacts and contexts of research in order to allocate QR? Should the definition of impact be broadened or refined? Is there scope for more or different use of metrics in any areas?

Please tell us your thoughts in no more than 500 words:

The Council for the Defence of British Universities (CDBU) strongly supports the need for universities to be accountable to taxpayers who ultimately provide funding. We do, however, have concerns about the way the REF has evolved: it does not give value for money, and by shaping research priorities it risks damaging the very thing it tries to assess, namely quality research.

We feel the starting point should be an appraisal of the added value provided by REF compared to alternative approaches, taking into account time, money and opportunity costs.

For instance, a case can be made that QR allocation would not be all that different if it were determined simply by the number of full-time equivalent research-active staff (see: <http://deevybee.blogspot.co.uk/2014/10/some-thoughts-on-use-of-metrics-in.html>). The obvious objection is that this measures volume and not quality. In practice, however, volume is linked to quality, because institutions are reluctant to take on unproductive researchers, and indeed many obtain their own funding through competitive schemes. This approach would raise new issues about how to define 'active researcher' to prevent gaming – e.g. it would be wise to restrict this to people who had been employed by the institution for a significant period (e.g. 3 years). Other criteria would need to be agreed separately for each discipline.

We do not support proposals to link QR to research income, as this would discourage institutions from supporting inexpensive research, with whole areas such as humanities and mathematics becoming unattractive. It would further encourage institutions to pressurise academics to obtain research funding for its own sake (see Q6). A better option might be to link QR to cost-effectiveness, e.g., research outputs relative to grant income, though weightings would then be needed to take into account discipline-specific costs.

Regarding impact, it seems perverse to exclude scientific or academic impact. As the Nurse report (page 7) says "More consideration needs to be given to highly

significant scholarly impact, i.e., work which has a major influence on the field." Further, no one type of measure is appropriate for all disciplines. There are many subject areas, especially in the humanities, where research does not have an obvious real-world application, but where the goal is rather to increase and refine our knowledge and understanding. There are areas of science, such as particle physics and mathematics, where it is virtually impossible to predict the nature and kind of impact that might arise. Furthermore, research in areas such as biomedical science and social sciences often has a practical impact as its direct aim, yet very good research may nevertheless fail to achieve that goal. Evaluation of research purely by impact achieved runs the danger of incentivising overhyped results or fraud.

Immediately after the last REF, many institutions set up committees to promote and monitor impact of new research, aware that evidence for impact would be easier to demonstrate if anticipated. There are already signs that this is shaping choices about which research should be done; this runs the risk of encouraging researchers to pursue short-term, politically expedient goals.

2 If REF is mainly a tool to allocate QR at institutional level, what is the benefit of organising an exercise over as many Units of Assessment as in REF 2014, or in having returns linking outputs to particular investigators? Would there be advantages in reporting on some dimensions of the REF (e.g. impact and/or environment) at a more aggregate or institutional level?

Please tell us your thoughts in no more than 500 words:

We hope that some attention will be given in the review to the phrase 'If REF is mainly a tool to allocate QR at institutional level'. We think that many of the problems of REF arise because of the mission creep that has occurred over the years, so that an exercise that started as involving allocation of QR now serves many additional functions. We agree with James Wilsdon that revision of REF should include a clear analysis of its function: see <http://wonkhe.com/blogs/time-for-a-stern-hard-look-at-the-ref/>. It may not be possible to have one solution for all these functions.

With regard to QR allocation, we agree that, if REF is to continue more-or-less in its current form, there is no need for linking outputs to particular investigators, and that an institution-wide exercise would be more economical and effective. However, we think serious consideration should be given to a more radical approach: a return to the pre-REF/RAE days of the 1980s when the funding allocation was institution-based and where QR allocation was not bundled in with other functions. The end result would not be all that different from that obtained with the far more complex and expensive REF, and there would be less risk of distorting research by altering incentives.

Section 2

3 What use is made of the information gathered through REF in decision making and strategic planning in your organisation? What information could be more useful? Does REF information duplicate or take priority over other management information?

Please tell us your thoughts in no more than 500 words:

N/A

4 What data should REF collect to be of greater support to Government and research funders in driving research excellence and productivity?

Please tell us your thoughts in no more than 500 words:

The question presupposes that without external motivators to 'drive' researchers to be both excellent and productive they will cease to be so. We disagree. There are already strong incentives for researchers to be excellent – both intrinsic (in terms of intellectual curiosity, pride in achievement and natural competitiveness) and extrinsic (in terms of needing to be excellent to obtain funding and publications, and career advancement). Further, at the institutional level, there are already drivers present, in terms of reputational advantage of a strong research base.

It is frequently remarked that many Nobel laureates of previous generations would have been fired for not meeting the criteria of the REF (e.g. <http://kingsreview.co.uk/magazine/blog/2014/02/24/how-academia-and-publishing-are-destroying-scientific-innovation-a-conversation-with-sydney-brenner/>), and a recent simulation study showed how research assessment may fail to capture Nobel quality (<http://www.economic-policy.org/62nd-economic-policy-panel/a-tale-of-two-metrics-research-assessment-vs-recognised-excellence/>). We maintain that researchers achieve most when given free rein to follow their best ideas without needing to participate in bureaucratic evaluation exercises. The REF is a particular burden for panel members who have to spend many hours conducting evaluations of research outputs, only to find that the end result is a single rating that seldom gives any surprises.

Of course, we appreciate that, if government money is awarded, then those receiving the money have a responsibility to use it wisely and well, and in the interests of society. However, the responsibility for ensuring research excellence and productivity should be the business of academic institutions. If that responsibility is handed over to government there is a risk of political interference in academic agendas, with researchers discouraged from pursuing controversial research that challenges existing orthodoxies.

Section 3

5 How might the REF be further refined or used by Government to incentivise constructive and creative behaviours such as promoting interdisciplinary research, collaboration between universities, and/or collaboration between universities and other public or private sector bodies?

Please tell us your thoughts in no more than 500 words:

Again, we question whether this is the business of the REF. Insofar as interdisciplinary or international research are deemed desirable, this can be achieved by funders or institutions promoting schemes to develop such work.

It also cannot be assumed that collaboration is always better. There are examples of disciplines (e.g. economics) where it could be argued there has been too

much uniformity, with a single theoretical approach predominating. There can be value in diversity; a field can be strengthened when individual differences are nurtured.

Section 4

6 In your view how does the REF process influence, positively or negatively, the choices of individual researchers and / or higher education institutions? What are the reasons for this and what are the effects? How do such effects of the REF compare with effects of other drivers in the system (e.g. success for individuals in international career markets, or for universities in global rankings)? What suggestions would you have to restrict gaming the system?

Please tell us your thoughts in no more than 500 words:

One of the factors leading to the creation of the CDBU was a widespread concern among academics about the adverse effects of REF. Specific concerns have been the setting of targets for grant income, as documented in the Times Higher Education (<https://www.timeshighereducation.com/news/grant-income-targets-set-one-six-universities-poll-suggests>), which in one extreme case led to the suicide of an otherwise successful academic. To quote from one of our members: "The competitiveness is now so intense that decency and compassion have gone out the window. I know few people who do not have terrifying stories from their workplace of staff who have been treated badly due to REF pressures, despite being hardworking and conscientious all the years of their career and frequently working very long hours."

We have seen a culture shift with researchers treating obtaining funding as an end in itself, rather than a means to an end; we are aware of some researchers who hoard numerous grants, which they end up unable to manage successfully. The notion of a scholar who does careful work in a given area is eroded by a push for high impact publications and more and more grant income. Many outstanding UK researchers from the past century would not have been appointable in the current climate because they focused on meticulous work over a long time period with little grant income. An emphasis on research has led to devaluation of teaching in many institutions.

Gaming of the system could be avoided if we rewarded sustained contribution rather than excessive productivity. As noted below (Q9), in science, a focus on cited papers needs to be complemented with attention to reproducibility of research findings.

In a paper in PLOS Medicine entitled 'How to Make More Published Research True', Ioannidis (2014) suggests changing the reward system so that "obtaining grants, awards, or other powers are considered negatively unless one delivers more good-quality science in proportion."

Global university rankings have had a pernicious effect, with virtually all institutions having the goal of being 'internationally world-leading'. This is simply not attainable, and it would make more sense to have a diverse higher-education sector where institutions could aim for more realistic goals of excellence that involve participation in large-scale collaborations, development of excellence in specific fields, and/or research activities that address local or national needs. We note that the first of these, collaboration between institutions, is not incentivised by the current REF (see Q7), and work of national significance is regarded as poor quality.

A more radical suggestion is that a reconfigured REF might actively strive to reduce the polarisation in our universities, by moving away from giving the lion's share of money to the most successful universities and instead providing funding more equitably so that research-active middle-ranking institutions could develop a stronger research base. This is not a view shared by all CDBU members, but reflects a concern that currently the 'golden triangle' is a magnet for research talent that makes it hard for other areas of the country to compete.

7 In your view how does the REF process influence the development of academic disciplines or impact upon other areas of scholarly activity relative to other factors? What changes would create or sustain positive influences in the future?

Please tell us your thoughts in no more than 500 words:

It is difficult to disentangle the influence of the REF from historical influences, including the increasing marketization of higher education. The REF is a major driver of excessive competition between institutions. For instance, we know of a case where a junior researcher was told she could not collaborate with a senior scientist in another institution, even though he was far more knowledgeable about her area than her local supervisor. The explanation was that it would not be acceptable for her publications to be affiliated with another institution.

We have also seen cases where institutions went to extraordinary lengths to get an academic 'on the books' in time for the REF, but then offered them little support afterwards. While large sums of money are spent on the salaries of these 'stars', at the other end of the career ladder many early career staff find themselves stuck in casual employment, with heavy teaching loads, unable to produce the 'REF-able' research which would gain them a permanent post. Thus we have seen polarisation within, as well as between institutions.

In some institutions, academics have been discouraged from doing public engagement work because this does not count as 'impact'.

The division between academics who are or are not entered for the REF has been painful for many of those who find themselves omitted; we are aware of cases where a lack of impact case studies restricted the number of staff who could be entered for the REF, leading to many junior scientists being omitted despite having adequate publications. This adversely affected their career prospects as well as their morale.

Our suggestions to overcome some of these problems are:

1. Reward institutions which have high rates of staff satisfaction. Institutions have been obsessed with assessing and improving student satisfaction, but little attention has been paid to the staff. REF-based incentives have led to a view of researchers as disposable, short-term employees who can be hired and fired as needed, rather than as people to be invested in.
2. Take into account the length of contract of research staff. Since the REF is supposed to affect funding for the future, the value of a productive academic should be weighted according to how far their contract extends into the future. As things stand, many institutions hire people for one year, or even for one term, just for the REF; by the time the results appear those people have already lost their jobs.
3. Reward collegial behaviour among researchers, including training and mentoring of younger researchers.

4. Stop rewarding those who hoard scarce resources – e.g. who have far more grant income than they can use effectively (see point 6.)
5. Encourage and reward collaboration between institutions
6. Resist the drive towards awarding an ever greater proportion of funding on the basis of impact.
7. Reward institutions that foster the careers of early career researchers. Younger researchers without a track record are disadvantaged by the current emphasis on impact and new talent is hindered in coming through into the next generation.

Section 5

8 How can the REF better address the future plans of institutions and how they will utilise QR funding obtained through the exercise?

Please tell us your thoughts in no more than 500 words:

Future plans already form the bedrock of all proposals to funders, including proposals for large-scale research institutions as well as individual research proposals. This is a function therefore that is already achieved and we are not clear that it is necessary for the REF to try to encompass it. If QR funding were made conditional on future plans, one would need some mechanism to then ensure that the plans had been followed through: otherwise there is a risk that this approach will generate a lot of hyperbole that will not deliver.

Final thoughts

9 Are there additional issues you would like to bring to the attention of the Review?

Please tell us your thoughts:

Issues specific to humanities: In the humanities much of the best work is published in books which are the result of mature reflection over a prolonged period of time. The provision for double weighting of monographs in the REF attempted to allow for this, but it is clear from the reports of the sub-panels under Main Panel D that the effect and extent of double weighting varied from one sub-panel to another. Moreover, while such double weighting made provision for monographs, only 1.9% of the outputs submitted to the Classics panel, for example, were scholarly editions, a form of output which has traditionally been the vehicle for some of the best and most lasting research in that discipline. (On this and other related issues see <http://cucd.blogs.sas.ac.uk/files/2015/01/Woolf-REF-2014-in-Classics.pdf>.) The temptation to deal with this problem by further complicating the rules for REF submission should be resisted. Instead, we suggest that the solution lies in moving to a genuinely lighter touch system, in which there is recognition that 'outputs' vary tremendously in substance, so that a distinguished book can be worth numerous journal articles. CDBU members in humanities also have concerns that the move toward Open Access requirements for the REF is not practicable for subject areas where books – which are seldom made Open Access - are a highly valued form of output.

Issues specific to sciences: In scientific fields there is concern that REF has been among the forces that have encouraged researchers to publish in journals that focus on ground-breaking, surprising results. Even though use of journal impact factor was expressly vetoed in HEFCE guidelines, it was widely used by departments in selecting submissions to REF2014 – perhaps due to awareness that no panel would be able to properly read and evaluate the huge number of submissions.

One suggestion has been that, rather than focusing on metrics such as citation counts, we should reward research that is reproducible. This is not straightforward, as we obviously want to encourage original, exciting findings, and reproducibility is hard to demonstrate prospectively. However, greater attention to methodological rigour rather than publication outlet could help rebalance this aspect of research assessment.

Final remarks: As will be evident from our responses, CDBU membership finds the previous methods used in the REF inadequate. We are between the devil of simple metrics, whose inadequacies are well documented in James Wilsdon's report, *The Metric Tide*, and the deep blue sea of 'peer review', which has all the problems listed in Derek Sayer's book *Rank Hypocrisies*. While most disciplines accept peer review as the 'least bad' way of appraising research quality in everyday academic life, in the context of the REF adequate peer review is hard to achieve. Instead, because of both the volume of work and limited range of subject expertise of panels in many disciplines, the review that is done runs the risk of being both subjective and superficial. On the other hand, if metrics are seen as a valid alternative, we could end up relying on inappropriate indices such as journal impact factor.