

Stern Review Response

The University of Cambridge welcomes the opportunity to provide evidence to the Stern Review of the Research Excellence Framework (REF). The University and its peer UK research-intensive institutions benefit substantially from the unhypothecated Quality-Related funding (QR) that they receive as a consequence of the REF. QR is critical to their success. It supports the existence of a diverse research portfolio across the country, one that is better able to respond to current and future challenges; it supports “blue skies” research that can lead to innovation and impact; it provides fundamental funding of key research in the arts and humanities disciplines; it provides a degree of stability for institutions and encourages long-term investment; and it enables institutions to support their early career researchers and consequently to develop capacity and skills for the future. It is no exaggeration to state that QR underpins the international standing of the UK Higher Education Sector.

The way in which QR is allocated is therefore of considerable importance. Although the Research Assessment Exercise (RAE) was originally established to fulfil this purpose, it, and its later iteration, the REF, have been used in ways that have progressively deviated from, and extended beyond, this core objective. Cambridge therefore welcomes that the Review has chosen to focus its initial questions on the primary role of the REF as a mechanism to determine the distribution of QR funding.

In addition to this digression from its original purpose, the REF in its current form has many weaknesses. It is widely accepted that the RAE had become increasingly burdensome and expensive to operate and this trend has accelerated as a result of the changes introduced for the REF. Key features of the RAE and REF, for example the focus on individuals as opposed to institutions, detract significantly from the value of the exercise and contribute substantially to its cost. It has also been the case that, since the funding formula has not been known in advance, the financial outcome of the exercise has often been unexpected. This uncertainty has greatly complicated the process of preparing effective financial plans in the run up to RAE/REF census dates. Whilst Cambridge benefits from its strong performance in the REF, the University is of the view that its international reputation and standing, and hence its ability to recruit to the highest quality academic staff and attract students, owes relatively little to its REF performance. Cambridge therefore endorses the aspiration of the Stern Review to seek to establish a “simpler, lighter touch system for the REF” and encourages the Review Panel to consider radical change.

It should remain a fundamental principle of any successor to the REF that dual support funding, including QR, should be allocated on the basis of research excellence. Cambridge is also of the view that the successor to the REF should reflect the original concept of a high-level process designed to determine an institutional-level block grant for research. Further, any successor to the REF should not aim to influence, or have the effect of influencing,

academic behaviour; nor should it have the effect of creating incentives for ‘gaming’ the exercise. Rather it must remain true to its core purpose; it should be designed simply to provide a cost-effective and rigorous way to allocate QR funding while providing assurance of the quality of research that QR supports.

Against this background, the “simpler, lighter touch system” envisaged by the Stern Review is not difficult to conceive: it should be undertaken at institutional level and it should be all-inclusive. The evaluation at institutional level of the quality of research outputs, the research environment and research impact would have the dual benefit of improving the integrity of the process and substantially reducing its cost. The introduction of a lighter touch process would also enable the panels carrying out the assessment to include a greater proportion of international and non-academic members, thereby providing a much more robust assessment of the international standing of research and its impact. This approach would also enable the establishment of specialist panels to assess narrative and quantitative information on the research environment and representative exemplars of institutional impact.

It is acknowledged that there is an argument for retaining the evaluation of research outputs at Unit of Assessment level. If this were to be the favoured structure, the panels responsible for UoA-level evaluations should have the freedom to use those measures most appropriate to assess quality in that discipline. This might include increased use of metrics but only where appropriate and where there is broad consensus within the discipline concerned that such metrics are an effective measure of research quality. However, it should be noted that, although the retention of UoAs may have benefits with respect to the use of a variety of approaches for the assessment of quality in different disciplines, it will militate against the effective assessment of research that spans more than one UoA, this being a particular difficulty given the increasing importance of interdisciplinary research.

The second element of the “simpler, lighter touch system”, the all-inclusive approach, would require institutions to be assessed on the totality of their research. One way of doing this would be to include the substantive research outputs of all academic staff in the evaluation of an Institution’s research performance. This would ensure a balanced assessment of the full range of research undertaken across institutions; it would substantially reduce the cost of laborious internal assessment relating to which staff and which outputs to return; and it would restrict the ability of institutions to influence the assessment process through staff selection. There are, however, three important details to this approach. First, all academic staff would need to be returned – rather than only those with a research component in their contracts – in order to provide a true measure of the research intensity of an institution. Secondly, assessment of outputs at an institutional-level would need to operate alongside effective mechanisms to ensure equality of treatment of particular categories of staff, such as early career researchers. Thirdly, the inclusion of all HESA-returned academic staff would

also ensure that the staff returned have a demonstrably strong and sustained research link to the institution.

More generally, Cambridge wishes to reiterate the concerns regarding the future of HEFCE that it expressed in its response to the consultation on the Higher Education Green Paper. It is imperative for the success of the UK's future research capability that the QR be fully protected from erosion. The University believes that the proposals in the Green Paper to separate teaching from research risk irreparable damage to the student and academic experience and to the sector's reputation. The institutional framework for the successor to the REF must reinforce the integration and unity of teaching and research for all students and academic staff in the university environment. Finally, in the context of research excellence, it is also worth highlighting once more the serious omission in the Green Paper of any mention of graduate students and the post-doctoral community.

This response to the Stern Review consultation has been prepared and submitted on behalf of the University of Cambridge and its 31 Colleges.

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?*

Process of evaluation

- It must remain a fundamental principle of any successor to the REF that dual support funding, including QR, should be allocated on the basis of peer-reviewed research excellence because this ultimately brings the greatest impacts and benefits. In 2014, only 65% of the REF scores directly measured the quality of research.
- The focus of assessment should be on the institution and not the individual. This principle should extend to all aspects of the current REF framework of output, environment and impact. Crucially, institutions should be assessed on the totality of their research with the substantive research outputs of all academic staff considered as evidence of the institution's research.
- Staff selection is perhaps the most complex element of the REF. It is also the basis of much of the "gaming" of the exercise. This should be addressed by:
 - a. A requirement that all HESA-returned academic staff must be included in the evaluation. In other words, all academic staff, irrespective of whether their contracts are teaching only, research only or both should be returned. This would have the twin benefits of increasing the accuracy and integrity of the exercise and significantly reducing its cost. It would also provide a true measure of the research intensity of an institution. If the staff return is restricted to staff with a research element in their contracts there is a risk that institutions alter employment contracts of their academic staff to enable particular individuals to be included or excluded. This can be hugely detrimental for those staff who have been excluded. It also runs counter to the core principle of the importance of the integration of teaching and research.
 - b. Recognition of the fact that the research environment in universities is also dependent upon and enriched by research staff, in particular college fellows, research fellows and senior postdoctoral researchers. Given the importance of their role in generating research outputs and impact, the successor to the REF must continue to allow their inclusion alongside other academic staff.
 - c. Ensuring that staff submitted to the REF have a strong and sustained research link to that institution. The successor to the REF must exclude researchers who only have only a peripheral or temporary role and therefore a very limited impact on the research environment over the assessment period. Cambridge believes that this change is essential to reinforce the integrity of the exercise.

- Further specific suggestions of steps that should be taken to improve future research evaluation include:
 - a. The measure of research quality through the evaluation of research outputs should be given greater weight in the overall assessment of research performance;
 - b. There should be increased international representation in the evaluation process, for example through the inclusion of a much higher proportion of international members on Panels;
 - c. Controls must be introduced to reduce grade inflation and 'bumping into the ceiling' as was the case during the 2014 assessment of impact. The latter issue could be resolved through extension of scales;
 - d. There must be increased rigour in assessment of impact case studies through, for example, the formation of specialist Panels (see below);
 - e. The limit of the maximum of four papers per individual submitted should be abandoned as this discriminates against highly-productive, world-leading researchers. Instead, institutions should be assessed on the totality of their research. An obvious way to do this would be to include the substantive research outputs of all academic staff in the evaluation of an institutions research performance. The use of approaches such as sampling, with appropriate allowances to ensure equality of treatment of particular categories of staff such as early career researchers, could ensure effective but lighter touch assessment of outputs. Such methods would also negate the need to use metrics except in those disciplines where there is a strong consensus on their validity as an effective measure of research quality;
 - f. HEFCE or its successor should simplify the submission process for institutions, for example through reduction in the number of deadlines.

Evaluation of impact

- Cambridge submitted 242 Impact Case studies to REF 2014. A move to an institutional-level evaluation of impact, as proposed above, would allow an in-depth assessment of a sample of exemplars. Fifty or fewer would be an appropriate number, being both representative and proportionate.
- Dedicated panels should be convened to undertake an institutional assessment of impact. Those panels should contain an appropriate mix of expertise (including more international members) to ensure they are able to assess the wide range of impacts. They should be supported by a dedicated secretariat able to undertake a rigorous audit of the evidence supporting case studies.
- A substantial reduction in the number of impact case studies required would reduce the burden on the external organisations (companies, public bodies, etc.) which are asked

to provide supporting evidence but which do not generally benefit from providing such evidence.

- The definition of impact should be further developed to provide greater discrimination between case studies. There is much anecdotal evidence from panels that there was substantial diversity in the scale and the reach of the impact case studies that were graded 4*.
- It is also not clear that panels undertook any systematic testing or questioning of the claims made in impact case studies. There is, therefore, relatively limited confidence in the sector in the rigour of the assessment of impact in REF2014.
- The current restrictions on the definition of impact and the need to link impact to research outputs in a specific UoA produce major anomalies. For example, Cambridge was not able to submit the Raspberry Pi and the work of Raspberry Pi Foundation as a case study of the impact of the work of the University's Computer Laboratory. The current procedures for the assessment of impact also fail to address adequately institutional initiatives, for example the highly regarded Cambridge Science Festival and Cambridge Festival of Ideas.

The use of metrics

- Although reducing the burden of the REF is essential, this must not be at the expense of its integrity as a measure of research excellence. Therefore, any use of metrics must be robust and appropriate to the discipline being assessed.
- The University is aware that the work of HEFCE REF Expert Advisory Groups, and more recently The Metric Tide, have raised significant questions over the development of metrics of any kind that would be sufficiently robust, particularly in respect of either author-level metrics or article-level metrics.
- If UoAs are retained, the panels responsible for UoA-level evaluations must have the freedom to publish and then to use those measures the panels consider most appropriate to assess quality in their respective disciplines. These might include increased use of metrics but only where this is appropriate and where there is consensus within the discipline concerned that such metrics can be effective in the assessment of research quality.
- The use of alternative approaches, for example the assessment of a major body of published work through sampling, offers a realistic alternative to metrics. They could ensure effective assessment of outputs at either discipline or institutional level.

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?

- Cambridge proposes radical change to the current REF model with a move to a process of institutional research evaluation. Cambridge considers that the totality of an institution's research output must be considered a key element of the evaluation of its research quality and environment. It is not clear that the integrity of the process is improved by the current requirement to assess outputs linked to particular investigators.
- Cambridge believes that the key issue in "gaming" in the REF as currently constituted is the ability of institutions (i) to restrict the inclusion of core academic staff; and (ii) to include staff that may have only a peripheral association with the institution concerned.
- REF2014 has provided an excellent demonstration of the close correlation across disciplines of research outputs, impact and environment. The burden of the REF would be reduced substantially by assessment of at least impact and environment at institutional level.
- Institutions should be assessed on the totality of their research. One approach would be to evaluate the substantive research outputs of all academic staff. As already stated, the use of alternative approaches such as sampling and, where appropriate, the use of some metrics could allow the effective assessment of outputs at institutional level (or, if necessary, at discipline level).
- Simplification of the process of staff selection would provide further protection against unintended bias and discrimination that may arise, for example for early career researchers and staff whose research careers have been affected by specific personal circumstances. However, this approach must be accompanied by effective mechanisms to ensure equality of treatment of particular categories of staff. More specifically, the opportunities that an institution provides through its commitment to the Concordat to support the Career Development of Researchers must inform the assessment of the research environment at institutional-level.
- It is recognised that there is an argument for retaining evaluation at Unit of Assessment level of research outputs. For example, retention of UoAs may have benefits with respect to the use of different approaches for the assessment of quality in different disciplines. If a UoA level approach for the assessment of research outputs is adopted, the UoAs must reflect the boundaries of recognised academic disciplines. However, it is noted that evaluation at UoA level will always militate against the effective assessment of research that spans more than one UoA i.e. much interdisciplinary research.

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?*

- Performance in the REF is rarely used as a metric in decision-making and strategic planning by the University. This reflects the University's concerns with the REF process, for example, the UK-centric nature of the assessment, the restrictions on the outputs assessed and the limited alignment of Units of Assessment to recognised academic disciplines.
- The University has benefited from the investment it has made in IT systems to support the REF submission in terms of its ability to record and analyse information on publications and other research outputs. This management information, which extends far beyond the limited publications information submitted to the REF, is combined with other internal and external data sources, for example on research funding, and is increasingly used to support research planning within the University.
- There is significant duplication of effort as a result of the inconsistencies between the reporting and compliance requirements of REF/HEFCE and RCUK, for example regarding open access publications. These contradictions have increased costs and have placed a significant additional burden on academic time. There are, therefore, significant opportunities to reduce the costs of compliance across the sector, for example by requiring RCUK to accept electronic transfer of data on research outputs in a format consistent with that used in the successor to the REF.

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

- The core purpose of the successor to the REF must be to provide a cost-effective and rigorous way to allocate QR funding. The design of any future REF and any requirement for data collection to support this purpose must not be compromised by the need for it to meet wider Government objectives.
- The single most important measure of the quality, range and reach of the academic contribution that an institution makes to the research endeavour is the totality of its substantive research outputs. The "best 4" publications of a highly productive, world-leading STEM researcher over the entire census period of the REF is not an effective measure of his or her research outputs; nor should a world-leading scholar in humanities be penalised for producing only one or two seminal books or monographs over the same period.
- Key measures of research excellence and productivity include:
 - a. The full range and totality of substantive research outputs. This could be broken down into output categories as appropriate on an institution-wide and discipline-

specific basis and should extend to outputs not currently considered by the REF, for example outstanding PhD theses.

- b. External research income, analysed as required into categories of funder, as a measure of productivity and national and international standing.
 - c. Representative exemplars of institutional impact, building on the experience of REF2014 case studies, and including rigorous evidence of the impact achieved.
 - d. Narrative and quantitative information concerning the research environment, including numbers of research staff and research students, evidence of investment in the development of research infrastructure, evidence of support for the career development of research staff, institutional initiatives supporting diversity and wider evidence of the impact of the institution, for example on international, national, regional and local enterprise and on Government policy.
- There are many opportunities outside the REF for Government to capture data and information on University research (e.g. HESA, HEBCIS). As previously stated, greater coordination and electronic exchange of data captured by and for all stakeholders in the sector including RCUK offers considerable opportunities for efficiency, cost reduction and, critically, a significant reduction in the burden placed on academic staff.

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?*

- Features of research activity, such as interdisciplinary or collaborative working and user engagement, are not ends in themselves; these are characteristics of research that may be necessary or essential to achieve world-leading outcomes or impact.
- The purpose of the REF is to provide a mechanism to distribute QR funding. Its focus must therefore remain on encouraging the highest quality research. This should in turn lead to the greatest impact.
- Any successor to the REF must not aim to influence, or have the effect of influencing, academic behaviour and nor should its design be compromised by any aim to achieve objectives other than a cost-effective and rigorous determination of the distribution of QR funding.

6. *In your view 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 the effects of the other drivers in the system (e.g. success of individuals in international career markets, or for universities in the global rankings)? What suggestions would you have to restrict gaming in the system?*

Influence of the REF

- Clearly the REF influences decision-making at institutions, for example regarding cycles of recruitment timed to suit REF census dates, reorganising and restructuring of departments, and even extending to guidelines for or instructions to staff regarding publication strategies. As previously stated, performance in the REF is rarely used as a metric in decision-making and strategic planning by Cambridge, reflecting the University's concerns with the REF system.
- Any method of assessment that generates a financial outcome or reward is likely to influence institutional and individual behaviour. The REF process, and the league tables and analysis that it generates, have encouraged "gaming" of the exercise. Cambridge believes that this, in turn, significantly reduces the value of the process and its outcome.
- "Gaming" in the REF has been driven, in part, by the way the REF has been managed; as an example, it is entirely inappropriate that the media is informed of the collective outcome of REF before the institutions that have taken part in the exercise. There should, therefore, be significant change in the way that the successor to the REF is publicised and its outcomes reported.

Behaviours and response to the REF

- Research is a highly international enterprise and many of the behaviours and actions attributed to the REF, for example national and international mobility of leading researchers, will continue even if the REF were to be abandoned. Numerous differently configured league tables will also always exist and will remain important to universities who will continue to seek ways to improve their standing.
- However, any reputable league table must be based on rigorous data. It is therefore questionable whether league tables should be based on the activities of only a small minority of the staff at the institution concerned.

Gaming in the REF

- As previously stated, Cambridge proposes radical change to the current REF model with a move to a process of institutional research evaluation. The primary issue in gaming is the ability of institutions to restrict the inclusion of core academic staff and to include staff that may have only a peripheral association with the institution concerned. Further, the assessment of a limited number of the outputs of only a proportion of academic staff is not an effective evaluation of the research activities of a Faculty or Department.
- We therefore propose that institutions should be assessed on the totality of their research. It follows that the substantive research outputs of all HESA-returnable

academic staff and all research fellows and senior postdoctoral researchers undertaking a sustained programme of research at the institution should be considered as evidence of the institution's research.

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?

- The REF process has had an essentially negative impact on academic development because of (i) the time the process takes away from core research and teaching, (ii) the bureaucratic burden imposed on individual staff, academic departments and institutions, and (iii) the significant operational cost to the HE sector (ca £250m), money that should be spent on academic activity, for example to increase the funds available to support postgraduate scholarships and junior research fellowships.
- The REF also impacts on scholarly activities by driving publications towards particular journals, including those where the additional requirement of open access can be punitively expensive (sometimes up to £4000 for an article). It has also perturbed the recruitment cycle for academics, impacting on the career progression of young researchers who are required to wait for the next pre-REF recruitment peak.
- Any method of assessment that generates a financial outcome or reward is likely to influence institutional, collective and/or individual behaviour. Cambridge believes, therefore, that any institutional framework for research evaluation must focus on the evaluation of academic excellence as this will reward the highest quality research outputs.

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

- Financial stability is essential for effective planning in the HE Sector. QR funding provides an essential element of that financial stability and supports an increasing number of posts in Universities; for example many staff dedicated to knowledge exchange, interaction with industry, and enterprise development are dependent on this funding. An early decision to adopt the proposed "simpler, lighter touch" approach would allow institutions to redirect a substantial element of the resources currently dedicated to planning for REF 2021 to support other activities.
- All universities must plan appropriately to ensure that adequate investment can be made in the development of academic staff and research facilities and to take account of natural turnover of senior staff, for example as the result of retirement. Forward plans and strategies may form part of the REF assessment criteria but they should not take precedence over measurable relevant evidence, for example of the investments by the institution in its core research infrastructure. Increased emphasis on future plans

risks the evaluation outcome being driven by the ability of institutions to develop and articulate strategies as opposed to their ability to produce world leading research outputs, to train high quality PhD students and to generate research impacts for the benefit of the nation.

- Cambridge manages its QR allocation as a key underpinning funding stream to support research across its academic activity. It provides critical match funding for RCUK grants and other research funding won in competition and it is essential for the financial sustainability of the University. It would be relatively straightforward to provide detailed information on the past and planned future use of QR at Cambridge should this be necessary as an input to an institutional-level successor to the REF. However, we strongly argue that the strength of QR as a funding mechanism is that it is unhypothecated and can therefore be used flexibly to respond to unforeseen opportunities.

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

- The University wishes to reiterate the concerns expressed in its response to the consultation of the Higher Education Green Paper regarding the future of HEFCE.
- The University believes that the proposals in the Green paper to separate teaching from research risk irreparable damage to the student and academic experience and the sector's international reputation. It is essential that undergraduates and postgraduates are exposed to the constantly expanding boundaries of knowledge that academic staff active in cutting edge research can impart to their students.
- The institutional framework for the successor to the REF must therefore reinforce the integration and unity of teaching and research in the university environment. The University considers that only an overarching body, with responsibility for both research and teaching, can be effective as a regulator for world-leading universities. Cambridge has therefore proposed that HEFCE's successor must assume all of HEFCE's current responsibilities, including those relating to research evaluation and the distribution of QR funding.
- If, notwithstanding these arguments, QR is moved to a body other than HEFCE or its successor, it is imperative for the success of the UK's future research capability that the QR be fully protected from erosion. The two funding streams must be maintained in their current proportions and at not less than their current amounts.
- Dual support has enabled the UK to achieve and maintain standards of the highest research excellence in our universities. Cambridge has expressed in its response to the Green Paper consultation its concerns regarding a number of the proposals in the Nurse Review. Cambridge reiterates its concerns, in particular those relating to the proposals for greater direct and detailed ministerial intervention. The University would be even more concerned should this ability to intervene also extend to the distribution of QR.