Response to the consultation on the future of research assessment

Background

1. The March budget statement announced the intention to consult on a new assessment system for quality-related higher education research. The consultation document can be found at:

   http://www.dfes.gov.uk/consultations/conDetails.cfm?consultationId=1404

2. Responses to the consultation document were sought by 13 October 2006 and an on-line response form with defined questions was provided.

3. The College's response to the consultation is provided below. As well as responding to the questions, we also prepared a letter in response since the questions were unduly restrictive.

4. The response has been approved by the Management Board and was informed by discussions with the College Research Committee, the College RAE Advisory Board, meetings with College staff who are members of RAE2008 panels, and a meeting with colleagues from Oxford, Cambridge, UCL and HEFCE.

Imperial College London

Response to the consultation on the future of research assessment

Consultation questions

1. Which, if any, of the RAE 2008 panels might adopt a greater or wholly metrics-based approach?

We would not support the assessment panels changing their assessment criteria at this stage in the process for RAE2008. The panels went through a lengthy and well-defined process to produce their assessment criteria. This included the consideration of metrics (e.g. research students, research assistants, fellows, research grant and contract income) to inform the assessment. The criteria are appropriate to the discipline and have been accepted by the subject communities and Higher Education Institutions (HEIs). A change at this stage would reduce confidence in the RAE2008 process and could lead to legal challenge. Similarly, a shadow metrics exercise alongside RAE2008 could unduly influence the outcomes by enabling the consideration of information which is not included in the criteria and working methods of the panels.

2. Have we identified all the important metrics? Bearing in mind the need to avoid increasing the overall burden of data collection on institutions, are there other indicators that we should consider?

We are not comfortable with the proposed metrics, either as a measure of research quality or as a mechanism to distribute research funding. Using metrics for generalised benchmarking purposes is substantially different to creating a UK system for research assessment and funding. The UK has limited practical experience (other than the quite different considerations of the recent and somewhat problematic HEIF exercise and HE-BCI survey) of developing entire metric systems. Hence we remain unconvinced that a system which replaces peer review with metrics could be an improvement. The caveats and limitations of particular metrics need to be understood and hence we advocate a period of very careful development and iteration, working and developing the types of metrics to ensure their rigour, applicability and fitness for purpose. Phasing in from 2009/10 would therefore not be appropriate.

A system of research assessment and funding based on one single form of metric – for example on income metrics – would be unsatisfactory. Using research income would lead to increased volume which in itself is not necessarily a measure of quality. Different metrics may be required for quality assessment and funding purposes. A basket of measures should be developed which are sufficiently broad to cover research and which can identify and measure quality. Quality measures should recognise competitive peer review (grants, fellowships etc). The basket of measures proposed in Annex 2 of the consultation document may be applicable for all subjects rather than just the arts and humanities.

We also have concerns about the implications of the statement in the question that one of, if not the main, incentives for replacing RAE2008 with a metrics process is the desire to reduce the overall burden of data collection on institutions. We do not see the RAE as a particular burden when compared against alternatives (e.g. Research Council peer review which, all things considered, is more time consuming

and expensive). Much of the information returned in the RAE is important for normal day-to-day management purposes anyhow.

3. Which of the alternative models described in this chapter do you consider to be the most suitable for STEM subjects? Are there alternative models or refinements of these models that you would want to propose?

We do not support **any** of the models proposed in the document. The models contradict a number of key and existing Government policies. The models do not support, and are at odds with, the stated, and extremely successful, research policy objective of the Government to fund the highest quality research. Every model (except Model D) provides funding to subjects within HEIs which do not at present earn HEFCE grant (those rated 1-3b).

Similarly, the flow of funds from STEM subjects to the arts and humanities (as implied by this modelling which we acknowledge as illustrative) would not support national needs.

The models are at odds with the Government stated policy objective of sustainability and full economic costing of research. The incentives from sustainability are such that, other things being equal, applications for research grants would be likely to reduce. However using research grant and contract income as a metric would be likely to lead to an increased number of applications since volume rather than the recovery of full costs would be the main incentive. Success rates would reduce even further, and an incentive to offer research at less than full economic cost would prevail.

The use of research grant and contract income would penalise blue skies research and other research which does no traditionally attract external funding. In addition, we note in paragraphs 2.10 - 2.13 of the Next Steps document that the Treasury would wish to encourage high risk research which has the potential to be of high impact. The implementation of metrics would be at odds with this desire, since metrics would tend to encourage 'safe' research, namely that which will be funded or cited. This would not be the way to fulfil the Treasury's objective of enhancing exploitation, as the amount of really innovative fundamental work, which others would then develop, would reduce.

The significant fluctuations in the results amongst subjects and HEIs provides a strong indication that these models are not sound and hence that further metrics and modelling are required.

4. What, in your view, would be an appropriate and workable basis for assessing and funding research in non-STEM subjects?

Compared to other HEIs, the College has only a small proportion of its work in arts and humanities subjects (namely history and business). Notwithstanding this however, we feel it is important that a common simultaneous assessment process takes place for all subjects at the same time. It is only through such a process that a reliable and understood method (and thereby benchmark) to assess the quality of research within an HEI can occur. Such a benchmark is important for the UK's global competitive position. Equally a separation would compromise fundamentally the need for a common framework across all disciplines in order to support multi-disciplinarily research.

The consultation appears to ignore the social sciences and the fact that the system will not be UK-wide. Further, disciplines do not always fit neatly into a STEM or arts and humanities classification. Statistics, Psychology and Economics are just a small number of such examples.

Difficulties with defining and capturing metrics for the arts and humanities should not be sufficient justification for a separate system.

5. What are the possible undesirable behavioural consequences of the different models and how might the effects be mitigated?

There are a number of undesirable behavioural consequences, some of which have been documented already. Most importantly, the policy of the selective funding of research would be compromised as would the sustainability agenda.

The loss of the RAE outcomes over time would disadvantage other research funders, not least the Research Councils and other Government departments who use the RAE outcomes to inform and make their funding decisions.

Since all models are based on research grant and contract income, the number of applications for research grants would rise, thereby increasing the burden on academic staff and HEIs in producing them and on the grant body concerned in assessing them. Costs and time for research administration would increase, and at a rate greater than the overall cost to the UK of conducting the RAE. In addition, the release of the peer-review burden on academics by an abolition of the RAE would be countered by an increase in the volume of research grant applications which require peer review. Equally, research grant and contract income measures success in winning and spending grants rather than the quality of the research work produced.

Some of the effects could be mitigated by attaching greater weight to research income which is of perceived higher quality (namely that which forms part of an open, competitive and, depending on source, peer review process). An adequate and robust method would need to be created to validate this so that the sector is being consistent and comparable. Ultimately, peer review would be the best mechanism to mitigate some of these effects and we would urge that peer review continues alongside a metrics assessment, at least for a number of years.

6. In principle, do you believe that a metrics-based approach for assessment or funding can be used across all institutions?

The use of metrics is not yet tested and requires much more work before it can be introduced with any level of confidence. The timing of the consultation is therefore inappropriate. Although a satisfactory metrics-based system might be found, it has not yet been designed. To state that metrics will phased in at the earliest opportunity (ie from 2009/10) threatens the establishment of a benchmark against which any metrics would need to be calibrated as it undermines the RAE2008 panels. Even following more detailed consideration, metrics alone would not measure or indicate research quality. Metrics should not be used independently of peer review.

Whilst we recognise the use of different measures for subjects, we would not want them to be so divergent that a reasonable benchmark measure of quality could not be created.

7. Should the funding bodies receive and consider institutions' research plans as part of the assessment process?

Institutional and subject research plans form an important mechanism to assess quality, since these provide a prospective context to the mainly retrospective information which would be provided in a metrics assessment. These would need to recognise the very different approaches to research management within HEIs and be informed by subject knowledge, such that plans are presented at a subject/discipline level rather than for the HEI as whole. For this reason, it would not be appropriate for the funding bodies to actually assess the research plans. Only experts in a particular subject could conduct the assessments in a meaningful manner.

8. How important do you feel it is for there to continue to be an independent assessment of UK higher education research quality for benchmarking purposes? Are there other ways in which this could be accomplished?

This is extremely important for national as well as international benchmarking purposes. Peer review assessment is the best, and indeed only, way to achieve this. The UK system must be clear, understood and enable meaningful comparisons. Extensive damage to the reputation and competitiveness of university-based research in the UK could ensue from these proposals. The policy of recognising and rewarding the highest quality research must remain. No other system has yet demonstrated that it could provide a common, acceptable and simultaneous assessment of research quality across all disciplines.

College generic response

This letter provides comments in addition to our on-line response to the consultation document 'Reform of Higher Education Research and Funding' and should be read in conjunction with that.

Imperial College London welcomes the opportunity to comment on the consultation document 'Reform of Higher Education Research and Funding'. As with other leading research-intensive universities, the College has very strong reservations about the content of the consultation proposals and on the manner in which the proposals are being taken forward. In particular we have concerns about the timing and immediacy of these proposals, their impact on the UK research policy agenda and their likely effects on the quality of the UK research base. Our comments are therefore framed in this context and seek to expand on some of the factual inaccuracies which may have led to the current consultation.

The College is deeply concerned with the current ill-considered proposals and their possible impact on the quality of the UK research base. Such a major potential change in higher education policy (namely the assessment and funding mechanism for research in UK universities) has been announced without any consultation or discussion involving any interested party (including the Funding Councils). The consultation does not reaffirm the stated and successful, policy and funding objective of selectivity awarding funds to the highest quality research areas. In fact, it does the opposite, since four of the five models upon which comments are sought redistribute funds to at present unfunded units, which are recognised to be of lower quality.

Before any review of the assessment of research is undertaken, the purpose of the Funding Councils' research funding (HEFCE QR) needs to be firmly established. QR is used for a number of purposes and is at risk of being subverted to support too wide a range of policy initiatives. In our view mainstream QR should be used to support research capacity and capability and to enable speculative 'blue skies' research. The research supported should accord with the Frascati definition used for the RAE and

incentivise research of the highest quality irrespective of the type of sponsor (eg peer-reviewed charity, business, EC). Any metrics used to assess and fund excellence should be different to those used to fund or incentivise other policy objectives such as the exploitation of research. Knowledge transfer and exploitation should be supported through third-stream HEFCE funding, namely HEIF and other mechanisms, not mainstream QR. Any system should not hinder, and would ideally strengthen the UK university research base such that the whole spectrum of basic and applied research and exploitation is enhanced for the benefit of UK competitiveness and innovation.

We note in paragraphs 2.10 - 2.13 of the Next Steps document that the Treasury would wish to encourage high risk research which has the potential to be of high impact. The implementation of metrics would be at odds with this desire, since metrics would tend to encourage 'safe' research, namely that which will be funded or cited. It is likely that the amount of really innovative fundamental work, which others would then develop, would reduce.

The proposals and possible metrics are not robust. Indeed the oft-quoted correlation between external research income and QR, which is the basis for the whole metrics initiative, is not statistically significant when a small number of the largest institutions are removed.

Bibliometrics are often cited as an alternative. However, citations are not a sufficient indicator of research quality and could not be used on their own. Simple citation statistics depend on the sub-discipline and whether the articles were reviews or original research.

Research quality, its funding and benchmarks are separate exercises but go hand-in-hand and must be part of one connected system. All involved with higher education have a responsibility to ensure that we are not replacing the RAE just for the sake of it. The RAE is not perfect, but it has benefited from being refined and improved over time. Successive reviews have reaffirmed its acceptance as a robust measure of research quality in the UK. The fact that other countries, e.g. Australia, are adopting systems similar to the RAE, is an argument in itself. In 2003, Sir Gareth Roberts reported in his review of the RAE that "any system of research assessment designed to identify the best research must be based on the judgement of experts who may, if they choose, employ performance indicators to inform their judgement" (Recommendation 1). 302 responses were received on the consultation and the overwhelming majority supported research assessment conducted through a process of expert peer review on a subject basis. We cannot see anything that has changed in the last three years.

The RAE has been supported and acknowledged on a number of occasions to be both equitable and cost effective when compared to the alternatives available. Equally, there is nothing to suggest that the RAE has had a detrimental impact on the UK research base and its quality.

The consultation document has done nothing to prove, or indeed even suggest, that a metrics system would be better. Moreover, the main justification for replacing the RAE, the release of academic time, is likely to prove illusory as the numbers of grant application and publication submissions, all of which need to be written and reviewed, increases.

The use of metrics is not yet tested and requires much more work before it can be introduced with any level of confidence. Although a satisfactory metrics-based

system might be found, it has not yet been designed. To state that metrics will phased in at the earliest opportunity (ie from 2009/10) is reckless and pre-supposes that suitable metrics will be developed. The UK has limited practical experience of developing and implementing a robust metrics system. HEIF and HE-BCI are poor examples, which themselves display many of the limitations of using metrics. A detailed period of careful development and iteration, working and developing the types of metrics to ensure their rigour, applicability and fitness for purpose is required.

There is no one single metric that can be used to measure research quality, or indeed that should be used to distribute funding to HEIs. A basket of metrics which are similar to those in Annex 2 of the consultation should be developed and their limitations understood, though no presumption should be made that any suitable metrics can be devised. Metrics may be an important dimension to funding but many are volume measures only or indeed not a sufficient measure of quality.

Even a casual examination of possible metrics reveals many limitations imposed by the absence of a peer-review mechanism. One of the main functions of RAE panels is to ensure that outcomes are not unjust. A metric driven assessment and funding system could not provide the same assurance, at least not at subject level.

Hence, for each broad subject area, a group of experts should devise, test, and regularly oversee the application of metrics in a form most suitable to the circumstances of their subject grouping. Even then, the breadth of disciplines may be too large to capture all subject specific variation. Expert review of metrics should commence only after the Panels have completed their RAE2008 responsibilities, as they should not be influenced in their RAE role by data that did not form part of the RAE. A comparable arrangement should then be maintained for any formal implementation in the years beyond 2008. The peers should also put the metrics into the RAE (or similar) profile so to facilitate both benchmarking and funding. The use of a profile for funding would ensure that HEIs which are similar in quality terms receive the same funding (other things being equal) as amounts would be determined by the profile not by a metric. In addition, profiles below a certain threshold need not be funded, thereby continuing the successful policy of the selective funding of research determined by quality.

The above points illustrate that the sector, even in STEM-subjects, still has some way to go in devising suitable metrics as proxies for research assessment and funding. We welcome the decision that RAE 2008 will take place as planned. Furthermore, the robust nature of the assessments that will result from the 2008 exercise will be a valuable means of calibrating and moderating any subsequent introduction of alternative arrangements. RAE 2008 panel members need to know that their assessments will be used to inform funding for many years to come. The RAE2008 results must drive the funding for at least five years.

Any new arrangement should have a steadily increasing influence beyond that period once it has been demonstrated that it is fit for purpose and its rigour has been accepted by the academic sector, by international bodies more widely, and by the other stakeholders to the process. Equally, if this cannot be demonstrated than the RAE should not be replaced.

To summarise, our view, which we understand to be supported by a number of research intensive HEIs, is that:

- The policy of selective funding of research, so that the highest quality areas are supported, must continue.
- Speculative, blue-skies research must continue to be enabled by the HEFCE funding formulae.
- Metrics should be developed as a result of a full and detailed consideration and involving an iterative process.
- Metrics should be introduced alongside and complemented by peer review.
- The RAE2008 outcomes should be the determinant of HEFCE QR funds for at least five years subsequent to the results being known.
- The introduction of metrics from 2009/10 onwards is premature.
- All subjects should be assessed as part of a common process, with allowable subject variations, and at the same time.
- RAE 2008 should proceed as planned.

We trust that these comments are helpful. We would be pleased to elaborate further.