Consultation

Launch Date 13 June 2006
Respond by 13 October 2006
Ref: DfES

Reform of higher education research assessment and funding

This consultation proposes a metrics-based research assessment and funding system to replace the Research Assessment Exercise after 2008.

department for education and skills
creating opportunity, releasing potential, achieving excellence
Reform of higher education research assessment and funding

A Consultation

To       Higher education sector
Issued   13 June 2006

If your enquiry is related to the policy content of the consultation you can contact Jim Cutshall on:

Enquiries To       Telephone: 020 7925 6371

Contact Details

Responses to the questions posed in this document should be posted to:

J Cutshall
1E
Department for Education & Skills
Sanctuary Buildings
Great Smith Street
London
SW1P 3BT

Or e-mailed to:

rae.consultation@dfes.gsi.gov.uk

Responses should be received by close of business on Friday 13 October.

Foreword by the Minister for Higher Education.

University research these days looks very different from how it appeared in 1986, when the first RAE was held. For one thing, no one doubts that there is much more of it – the financial rewards in the form of the funding that has flowed from the RAE have helped see to that. For another, we know much more about it thanks to the information that the funding bodies now collect. We also know that UK research is excellent by comparison with that of other countries because one international study after another tells us so. This document proposes that the 2008 RAE should proceed largely as planned in parallel with a “shadow metrics exercise and that after 2008
it should be replaced with a new and lighter-touch system based largely on metrics. The principle of using information that is already collected routinely to assess research quality and allocate funding must be the right one. The savings of time and effort that this can bring for university teachers and administrators alike should be welcomed by all, as should the transparency that a system based on publicly available data potentially offers. The approach that we advocate is measured and takes full account of the differences in the current availability and applicability of metrics to different subjects. Our aim is to ensure that we are able to continue to promote and reward excellence as effectively in the next 20 years as we have done in the past.

The essence of our proposals is to be found in chapters 5 and 6. We have tried to express quite technical issues as plainly as is consistent with giving the required level of detail about what we propose. We have also tried to illustrate the effects of the different options available wherever possible because we want the whole university community to understand clearly what we want to do and how it would affect them and the sector as a whole. Only in that way can you make up your own minds about the best way to move towards a more metrics-based system.

This document has been developed by a working group whose members included representatives of all the UK higher education funding councils and a number of Government Departments. I am grateful to them and to the many external partners who have also sent us their views.

I hope that as many of you as possible will read our proposals and discuss them before deciding how to respond. We will consider your views over the autumn and announce how we intend to proceed before the end of the year.

Bill Rammell
Minister of State for Higher Education
& Lifelong Learning

1 Main Findings and Conclusions.

1.1 The results of our work apply to UK higher education as a whole. We reaffirm the value of a common system that allows the comparative research quality in the UK, both at subject and institutional level, to be benchmarked nationally and internationally.
1.2 We also recognise that higher education is a devolved matter and that policy decisions in this area, notably as regards funding, must ultimately remain issues to be determined by the authorities responsible for higher education in England, Scotland, Wales and Northern Ireland.

1.3 We have sought to devise a system based mainly on the use of metrics rather than mainly on separate academic peer review that should be capable, over time, of reducing substantially the cost and burden of the current Research Assessment Exercise (RAE), while continuing to recognise and reward research excellence in ways that are accepted by the sector as appropriate.

1.4 However, it is clear that the data sources currently available do not make it possible to move immediately to a single research assessment and/or funding system driven entirely by metrics.

1.5 At the level of the 67 units of assessment (UoAs) or even the 15 broad disciplines envisaged for the 2008 RAE, there is a far closer correspondence between Quality-Related (QR) research funding and non-QR research income in STEM subjects, where the research undertaken is largely project-based and receives substantial grant support from the Research Councils and other external sources.

1.6 At institutional level, the correlation is closer for very large institutions than for very small or specialist ones.

1.7 Making these calculations on the basis of a 3-year rolling average rather than a single year’s research income makes the correlation closer still by smoothing the effects of large, once-off grants.

1.8 We present five funding models based wholly or mainly on different types of research income metrics. Technical descriptions and illustrations of the effects of these models may be found online at www.hefce.ac.uk/research.

1.9 For subjects other than STEM, a more differentiated approach will be required in order to ensure that research excellence is recognised and rewarded fairly. One example of the kind of approach that might be applied in the case of the arts and humanities is illustrated at Annex 2 to this document. Alternatively, a separate and slimmed-down peer review exercise may prove the most appropriate way forward. We have an open mind on this question. For the immediate future, and

---

1 - Hereafter, we will use the term “STEM” to denote the sciences, technology, engineering, mathematics and medical subjects. Together, these account for about 60% of QR expenditure and about 50% of research volume (ie numbers of research-active staff, research students, etc).
pending the development of more sensitive indicators for these subjects, notably in the field of bibliometrics, we conclude that non-numerical indicators will continue to need to be taken into account. We suggest that the higher education funding bodies may want to remit this task to a small number of panels established for the purpose. Special attention may also have to be paid to ensuring that small institutions are treated fairly.

1.10 In consideration of these facts and the weight of opinion in those parts of the higher education sector to have expressed a view, we conclude that the 2008 RAE should proceed largely as planned. We recommend, however, that the UoA or main panel-level peer review panels in the RAE should be free to make greater use of the metrics for their areas than is currently envisaged in reaching their judgements, or even to base those judgements entirely upon metrics if, in their opinion, this leads to a reliable assessment of research quality. We will take the necessary steps to enable those panels where there is scope for greater use of metrics and which wish to do so to operate in this way.

1.11 Alongside RAE 2008, we intend to run a shadow metrics exercise, drawing on the conclusions of this consultation in constructing the precise metrics to be used. After RAE 2008, the funding bodies will evaluate the outcomes of the RAE and of the shadow metrics exercise to determine the role that metrics should play in their QR funding formulae.

1.12 The outcomes of RAE 2008 should provide the basis for a steady transition towards the use of research metrics for research quality assessment and funding purposes across all subjects and types of institution. In the case of STEM subjects and larger institutions, this change should be relatively straightforward. It would be phased in from 2009/10 in England.

1.13 These various sources of data could be combined in an institutional research profile. We suggest that, in addition to subject- or discipline-level information, the contents of and delivery against the research plans that institutions already produce for RAE or internal management purposes could form part of this. Such profiles would also be useful for international benchmarking purposes if they included refined bibliometric data.

1.14 In order to avoid financial instability, we recommend that the maximum gain or loss that any institution should be allowed to sustain from one year to the next should be moderated appropriately by the funding bodies. It might, for example, be capped at around 5% of QR income.

1.15 We recommend that the higher education funding bodies should
monitor the effects of any new system on research behaviour and be prepared if necessary to take steps to mitigate any unintended or unwelcome effects.

2 Introduction: the evolution of the Research Assessment Exercise

2.1 In this chapter, we describe how the RAE has evolved.

2.2 From 1919 until 1986, Government funding to support core university research activity was included in a single block grant covering both teaching and research activities.

2.3 The first RAE (then called the Research Selectivity Exercise) was organised by the University Grants Committee in 1986. This required universities to return statistical data (metrics) of various types (including numbers of staff and research students) together with up to four examples of the recent research output of each research-active member of staff. This evidence was organised into units of assessment (UoAs), mainly corresponding to academic departments, considered and allotted to quality bands by expert panels.

2.4 A further research selectivity exercise was implemented by the Universities Funding Council in 1989. In 1992, following a redefinition of the higher education sector by the Further and Higher Education Act, the first RAE was held. This covered not only those institutions that had taken part in the previous two exercises, but also those previously funded through the Polytechnics and Colleges Funding Council, together with a small number of other institutions. The outcome of the exercise was both published and used as one of the main elements in a formula designed to allocate a new, dedicated funding stream for research. This has become known as quality-related (QR) funding. Although higher education and its funding is a devolved matter, the RAE remains a UK-wide exercise.

2.5 Further RAEs were held in 1996 and 2001. The next RAE is planned for 2008.

2.6 Successive RAEs have brought refinements and improvements to the 1986 format, especially in the adequacy of the range of research that is covered and assessment procedures. Nevertheless, despite the passage of 20 years, the massive expansion of the sector, the effects of devolution and several formal reviews, most recently in 2003 by Professor Sir Gareth Roberts, the salient features of the 2008 RAE will still closely resemble those of the 1986 exercise. It will set a benchmark of UK research quality and, like its predecessors, will require submission of up to four research items for each research-active member of staff, together with some metrics and qualitative...
information which will together determine up to 50% of the assessment. There will be 67 UoAs corresponding to principal university subjects, each having an expert peer review panel. Panels will assign a quality profile to each institution submitting staff in each UoA, and these will be used by the funding bodies to help determine future QR funding.

2.7 The net effect of successive RAEs on research practices and on the overall quality of UK research has been beneficial. Since 1986 it has become accepted that robust assessments of the quality of publicly funded research within higher education should be produced and made widely available; and that the processes of institutional research planning and management which the exercise examines and encourages are essential to maintaining a strong research base. Partly as a result, UK research is now not only mature and well managed, but also of a quality that makes it internationally competitive.

2.8 At the same time the exercise has continued to attract criticism. For example:
• While it is acknowledged that the present system has been very successful in raising research quality, it has also been alleged that the scale and cost of it now outweighs its ongoing value; and, therefore, that there may be a more efficient alternative.
• It has also been observed that, at certain levels of aggregation, the distribution of QR funding is well correlated with research income from other sources. This provides a case for believing that a metrics-based system may represent a more efficient alternative for allocating QR.
• There is a further contention: that the workings of the system, and in particular the perceived expectations of panels by the academic research community, inhibits ambition in development of research strategies and may inhibit particular approaches to research.

2.9 Having achieved a considerable impact as a catalyst for change in research management and a credible means of benchmarking institutional research, it may no longer be the best or the most efficient way to allocate resources while driving up research quality. That is why we should explore the change to a metrics-based system now.

2.10 QR is not the only source of public funding for research. It is a component (with Research Council grants) of dual support. Thus, it is also part of a broader settlement under which research of public interest in universities is funded by a partnership of several UK Government sources, charities and the European Union.

2.11 The multiple roles played by QR were summed up as follows in the Science and innovation investment framework 2004-2014: “QR provides a foundation allowing University leaders to take strategic decisions about the research activities of their own institutions. It funds
the basic research infrastructure – including the salary costs of permanent academic staff, support staff, equipment and libraries – that gives institutions the base from which to undertake research commissioned by other funding sources; the flexibility to react quickly to emerging priorities and new fields of enquiry; and the capacity to undertake “blue skies” research.” Research commissioned by the DfES into the uses to which QR is put at institutional level has confirmed this analysis.

2.12 The Government has repeatedly stressed its commitment to the dual support system for public funding of research, and to the deployment of public funds to underpin a broad range of economically and socially beneficial research activity. However, it has also laid increasing emphasis on the need to consider alternative approaches to quality assessment and to allocating funding, and in particular on the scope for substantially greater dependence on metrics. Thus the Science and innovation investment framework 2004-2014 noted an intention to examine the use of metrics as an alternative to peer review, while the more recent document Science and innovation investment framework 2004-2014: next steps committed the government and the higher education funding bodies to consulting the sector on what those metrics should be in the firm expectation that they would form the basis of a new system for research quality assessment and funding after 2008.

2.13 In the remainder of this document, we present the principles that should underpin an alternative and simpler system than the current arrangements. We also present a number of examples of how such a system might work in order to inform consultation responses.

3 Principles of a new assessment system

3.1 The Government recognises that, within the QR element of the dual support system, there is a continuing need for a system or systems to identify the best research. This should:

• allow the research performance of UK universities to be benchmarked by policy-makers and university leaders in a way that commands the confidence of all stakeholders (notably higher education institutions, disciplinary communities, end-users of research and Government); and

• form the basis for the selective allocation of quality-related research funding to universities in England and in any other parts of the UK where the authorities responsible for higher education funding judge this appropriate.

3.2 As the discussion of metrics that follows in chapter 4 will show,
although a metrics-based assessment system and a metrics-based funding system would probably use a number of indicators in common, they need not necessarily be identical. Nor are metrics necessarily applicable to all subjects. It would thus in principle be open to the UK funding bodies to use metrics for either quality assessment or funding allocation, or both. This will be a decision for them to take in due course in the light of their different legal powers, priorities and responsibilities.

3.3 In the light of our evaluation of the available indicators and consideration of representations from the higher education sector, we have endeavoured to devise a system based on research metrics that could potentially offer a more efficient way of achieving this in the future for many subjects than the current method.

3.4 The Science and innovation investment framework 2004-2014: next steps made the presumption that the 2008 RAE would go ahead as planned, incorporating a shadow metrics exercise alongside the traditional panel-based peer review system. However, Next Steps also said that if a clear majority of UK universities were to favour an earlier move to a simpler system, we would be open to that. Initial response from the sector suggests that it is unlikely that there will be a clear majority for wholesale change. We would, however, be interested in views on whether there are specific subject areas where it would be valuable for the 2008 RAE panels to make greater use of metrics, as some indeed did in 2001, or even for a metrics-only approach to be used. All of the RAE 2008 panels already intend to make some use of metrics. However, the funding bodies will shortly seek the views of the RAE main panels on whether they may be able to make greater use of metrics in RAE 2008, particularly in disciplines such as engineering, physics, chemistry, medicine and the biological sciences. Clearly, any such proposals should not impose new data collection demands on institutions.

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

3.5 Paragraph 4.10 of the Next steps paper identified four broad objectives for a metrics-based system:

- to reward excellent research of all types, from curiosity-driven to user-focused, maintaining the international standing of the UK’s research base, encouraging collaboration and supporting interdisciplinary work;

- the dual support system should be preserved so that QR funds can continue to provide institutions with the freedom to invest strategically in research, drawing on their own strengths and reflecting local, national and international needs;

- the burden on institutions should be minimised with no unnecessary information being collected; and
• the assessment and allocation processes should:
  o be open to and apply equally to all institutions;
  o be simple, transparent and cost-effective;
  o result in a funding stream to an institution (not an individual or group); and
  o allow higher education institutions to plan effectively.

3.6 In addition to these considerations, we are aware of how important it
will be to understand the behavioural impacts of whatever system is
eventually adopted and to ensure that changes in behaviour contribute
positively to the objectives. Just as the RAE has influenced academic
behaviour, so any metrics-based system could also be expected to
have behavioural effects. We need to be clear about how these could
be predicted and, where appropriate, obviated.

3.7 In evaluating a new system, we will have special regard to the extent
to which costs and burdens are reduced across the dual support
system. We will therefore seek to model likely outcomes and
incentives to ensure that costs are not inflated for research funders
and institutions, and seek to eliminate perverse incentives.

3.8 In recent times, much work has been done on how QR enables and
supports other research streams. Nonetheless, we are mindful of the
need to maintain the distinct role that QR plays within the dual support
system and that, accordingly, any system of metrics which is used to
allocate all or part of the QR budget must be relevant to its purposes.
Clearly, if inappropriate drivers are used, QR will cease to perform its
strategic function within the dual support system to which all UK
administrations are firmly committed. In particular, QR must continue
to support research capacity and capability; it should support strategic,
long-term research; and it should enable speculative research.

4 Information on research activity and quality

4.1 There is a wealth of information available on research activity and
quality. In this chapter, we review the range.

4.2 First, a group of indicators is derived from research income measures.
In a number of these cases - indeed to an extent in nearly all of them -
there is an element of peer review and so these reflect both volume
and, indirectly, quality. It is important to recognise, however, that in no
case does the scope and purpose of the peer review activity in
question exactly duplicate that of the RAE. The main such indicators
are:

• Income from Research Council grants. This is collected for all
  universities on a robust basis through the annual Higher Education
  Statistics Agency (HESA) return, broken down by 40 subject cost
centres. The data reflect peer review judgements applied by the Councils. They cover all disciplines, but the relative balance of funding from Research Councils and QR varies substantially across disciplines. It might also be necessary to resolve the technical difficulty of mapping subject cost centres onto whichever distribution of subjects (for example, those covered by the RAE 2008 main panels) was used for quality assessment purposes.

- Income from other external sources, such as UK and overseas governments; the EU; the NHS; research charities; industry and other. This is also collected through HESA. It reflects a broader range of user priorities. Although these sources mainly apply some form of competition or peer review, the degree of rigour in this varies considerably. There are similar considerations of subject coverage as for the Research Councils, and the very significant sums awarded by charitable sources are heavily concentrated in medical subjects.
- A subsidiary indicator related to Research Council success rates by subject and by institution could be introduced.
- Any metrics based on a range of research income could be weighted according to the source of the income in question and any outcomes that it is considered desirable to incentivise.

4.3 Then there is a group of volume measures:

- Numbers of research-active academic staff: this could be a strong indicator of total activity levels. At present this data is only collected through the RAE. For an entirely metrics-based approach that required a volume measure, this data, or an equivalent volume measure, would have to be collected in some other way.
- Numbers of postgraduate students. These are collected through HESA. There are two useful indicators: PhDs awarded and numbers of postgraduate research students enrolled. Both are robust and cover all disciplines. The former is arguably a better indicator of quality since it takes into account completion rates. The latter may be a better measure of current research activity and of the reputation of a department or institution within the research community.
- We should also consider whether the estimates of time spent on research activity which academic staff nowadays make to their institutions as part of the Transparent Approach to Costing (TRAC) process offer a suitable volume measure for some purposes.

4.4 We have also considered bibliometric data. A considerable body of such data is available about rates of publication in leading international journals and the levels of citation resulting from this, collected in the USA but covering journals published elsewhere too. It is possible to estimate impact indices for institutions from this data. Bibliometric information is generally regarded as a strong indicator of national research strengths in broad subject groups in the sciences, technology and medicine (and is used as such for example in the annual OSI report on PSA target metrics for the UK research base). There is
however less agreement on the use of these data to measure quality of output from individual departments and universities; their interpretation at this level of disaggregation is contested in many subject areas and there are specific concerns about the choice of journals covered by the database and the adequacy of the coverage of some disciplines. However, steps are currently in train to extend these databases to subjects which currently have poor coverage. What is particularly important about these indicators is that they facilitate international comparisons.

4.5 Another group of indicators mainly relates to the products of user-led or applied research, such as patenting, spin-out companies and activity with industry. This data is collected through HESA and, through the dedicated Higher Education: Business and Community Interaction Survey used in England to inform allocations of Higher Education Innovation Fund (HEIF) grant.

4.6 Finally, there are data sources which offer information about perceived research quality. They include:
• Peer review information from the Research Councils. The UK Research Councils undertake extensive peer review in awarding grants; they also now require the submission of end of project reports (which, in some cases, are graded). This information has only recently started to be collected on a common basis by all the Councils. It is subject to the same considerations as Research Council grants and relates in many cases to work done some years previously.
• Research Councils also, from time to time, commission international panels to review the health of particular subjects.
• Institutional research plans. All universities making RAE submissions have a research strategy and forward plans at institutional level. Many also have such plans at departmental level. These could provide important information about universities' research culture and vitality as well as a snapshot of the state and direction of the national research base. It would be feasible to collect these periodically, although they would then require some interpretation and subsequent monitoring against delivery.

4.7 In the following two chapters, we explore ways in which information can be used first to generate a funding formula – taking this first so that we see if it is possible to tease out a quantitative quality measure en route, through modelling – and then how they can contribute to a research quality profile.

Question 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?
5 Possible models

5.1 Much of the available data is quantitative – the so-called metrics. We have seen from the previous chapter that, collectively, the data, quantitative and qualitative, provide a picture of research activity and quality. The next step is to explore the inter-relationships of the metrics within mathematical or econometric models. The main motivation for this is likely to be the development of models to form the basis of funding allocations for those funding bodies that wish to proceed in this way, but we can also explore the extent to which the analysis can offer free-standing measures of research quality. This analysis involves relating the metrics to funding allocations from a previous RAE. This is because the RAE offers the only current guide to quality by UoA by institution. However, when considering the value of different models it is important to bear in mind the criticism that the RAE has attracted, in particular in relation to applied and multi-disciplinary research.

5.2 The quality measures generated by the models are not very highly correlated with individual RAE UoA ratings. This suggests that while the correlations are good in aggregate for funding allocations, we should not rely on these models to offer quality assessments at subject level. That is especially true of non-scientific subjects. As a new system develops, the correlation between funding and quality of outputs will clearly be a matter of significance for research funders in general and the funding bodies in particular, as will accountability both at an institutional and at a system level for the investment of QR resource. We take up this issue in the next chapter.

5.3 We have identified five possible models, which are described in general terms below. Technical descriptions of each and their illustrative effects on institutional QR allocations are available online at www.hefce.ac.uk/research. All of these models are based on STEM subjects only, with allocations for all other subjects held constant in line with current QR allocations. Options for non-STEM subjects are considered at the end of this chapter.

5.4 All of the models use four-year totals of external research income. Unfortunately, due to data constraints, we have only been able to model data for the period 1996/97-1999/2000. However, if they were to be run using income data from different years, we anticipate that a similar overall pattern of distribution would emerge. All the models include income from charities, business and Government departments, in recognition of the importance of research that supports economic and social objectives.
Model A

5.5 This model involves distributing STEM QR funding across institutions as a function of their external research income (Research Councils, EU, charities, UK Government and NHS, domestic and international businesses). The share of the total external research income across all STEM disciplines received by each institution provides a formula for allocating QR funding. Across the subjects included in the model, income from external sources is added up for each institution to produce a total QR allocation.

5.6 This model is the simplest of those explored here. It directly reflects the priorities of the full range of research project funders, rewarding institutions that are successful in competing for Research Council grants and in meeting the research needs of business, charities and Government departments. The information required to run this model is readily available for approximately 40 cost centres and updated annually so it would support a dynamic funding system. Subjects that receive a relatively small proportion of their income through project or programme grants would receive less QR funding under this method than under the current system. This could be problematic where subjects receive relatively little external research income because of the nature of research involved, but at the margins it is difficult to separate out this effect from the effect of some departments or institutions not being successful at competing for research projects and programmes because of the perceived quality or relevance of their research. Another issue to bear in mind under this model is that medicine receives about 90% of all charitable research funding and is likely to receive more QR funding under this system than it would under the current system or the other models proposed. Similarly “big” science would receive more QR funding under this system.

Model B

5.7 This model takes the same research income data as in Model A in each UoA and divides it by a measure of research volume, largely driven by the number of research active staff with information obtained from the RAE. The results are standardised to achieve a continuous quality rating which is multiplied by a volume indicator and the cost ratios currently used by HEFCE to allocate QR (1.6 for medicine and most sciences, 1.3 for pure mathematics and nursing, and 1.0 for other subjects). These ratios are determined by annual HESA data on departmental expenditure divided by total staff numbers in each subject area. The ratios were implemented in 1993/94 and there has been one change since, when the weighting for medicine and most sciences changed from 1.7 to 1.6.

5.8 This model counteracts the potential in model A for rewarding subjects that do well because they involve relatively expensive research and
receive large external research grants. Instead, it uses the cost ratios devised by HEFCE to reflect the need to compensate subjects that involve higher costs. This produces a distribution of resources between subjects driven by cost and volume indicators rather than by income levels and an allocation that more closely resembles past allocations of QR than Model A.

5.9 However, it relies on assumptions about cost ratios which are currently based on fairly simple analysis and may not closely reflect variations in the true costs of undertaking different types of research, given the relatively static nature of the cost ratios. It relies on identifying a simple mechanism for collecting robust information on research volume in the absence of an RAE process.

Model C

5.10 This model works in a similar way to Model A, but constrains the "pot sizes" for each UoA rather than allowing the unit costs for each UoA to fall from the modelling. Currently, to determine the pot sizes for each of the 68 UoAs, HEFCE creates a total measure of volume across all departments rated 4 and above and multiplies this by the cost ratios described under Model B. It uses this ratio to divide the total QR funding available into pots for each UoA. These pots are then distributed across departments.

5.11 During the period from 2001-02 to 2005-06 overall QR levels in England went up by 36% on a like-for-like basis (from £868 million to £1,178 million). Looking at two subjects that are within the same cost band, chemistry and biological sciences, we see that volume effects are reflected in the relative funding changes taking place. Funding for chemistry went up by 17% (from £39.7 million to £46.4 million), which is below the sector average, due to a reduction in its share of research volume during that period. On the contrary, funds for biological sciences went up by 59% (from £56.1 million to £89.5 million), which is well above the sector average, due to an increase in its share of research volume during that period. Most other pot sizes stayed the same in relative terms.

5.12 This model creates a distribution across UoAs that is identical to that achieved through the current distribution of QR, which reflects the cost, according to relatively crude analysis, and volume of quality activity in institutions. Any change is restricted to changes in allocations between institutions within a particular UoA. As with Model B, it relies on there being a measure of volume available to calculate pot sizes.

5.13 Under the current method for attributing pot sizes in England, this model means that all 68 UoAs receive a pot of funding as long as there are departments rated 4 and above. In theory this method could result in the pot for a UoA with ten 4-rated departments being identical to a
pot for a UoA with five 5*-rated departments and five 5-rated departments, as long as both UoAs had the same volume of research activity. This means that a lack of flexibility to recognise new priorities is inherent in the model.

5.14 It is worth noting that the exact method for calculating pot sizes need not remain precisely as that currently employed by HEFCE and the total number of pots would need to be decided. It may be that a simple distinction between STEM and other subjects (which is essentially what has been done in the modelling that we have carried out for Model A) would be sufficient. Or there may be a case for separating out medical research, in order to mitigate its predominance in attracting charitable funding.

Model D

5.15 This model takes the quality measure from Model B, ranks it and compares the ranking to the 2001 RAE ranks in each UoA. It then allocates funds in a "stepped" way, i.e. reproducing the number of departments in each part of the categorical distribution under the RAE system and dividing funding between them. Like Model C, it retains the volume indicators and subject cost relativities from the current QR allocation method and allows these to drive the distribution of funds between subjects.

5.16 The results of this model retain the current differences in funding between each RAE rating and this is the only model under which some departments would receive no QR funding (because they have been assigned RAE ratings of 1, 2 or 3b) even if they do have some external research income. All the other models, as presented here, provide a continuous funding profile related to a department’s external income.

5.17 This model relies on the results of the RAE to allocate funding in a stepped or categorical manner. As such it is the least dynamic of the options presented here and with a quality profile replacing single RAE ratings for each department in the 2008 RAE, this method would need some re-working. Further, with the absence of an RAE process beyond 2008 it is not clear how the categorisation could be updated periodically.

Model E

5.18 This model is the same as Model A, but effectively gives a different weighting to research income from charities to balance the effect of the predominance of medical research in charity funding. This means that the QR allocation for medical research is lower under Model E than under Model A, though still higher than under the baseline QR allocations.
Question 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?

5.19 The modelling that we have carried out makes clear that research income metrics would currently be of limited applicability for non-STEM subjects. Here, a more differentiated approach will be required in order to ensure that research excellence is recognised and rewarded fairly. One example of the kind of approach that might be applied in the case of the arts and humanities is shown for illustrative purposes at Annex 2 to this document. Another possibility that has previously been proposed by some in the higher education sector would be continuing to run a more streamlined RAE system for these subjects. We have an open mind on this question. In any event, for the immediate future non-numerical indicators will continue to need to be taken into account for non-STEM subjects. In determining how best to do this, the higher education funding bodies may want to seek the advice of a small number of panels established for the purpose.

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

5.20 We note that HEFCE and the Arts and Humanities Research Council (AHRC) are jointly to establish a working group to advise on possible assessment and funding metrics in the arts and humanities. This group will report in time for its recommendations to be taken into account in finalising the new approaches. Membership and terms of reference for this working group will be announced shortly after the publication of this document. The Economic and Social Research Council is also committed to working with the Funding Councils to develop more effective metrics for the disciplines they cover.

5.21 We recognise that all of the funding models under discussion are likely to have an impact on the behaviour of researchers and their institutions. It is important to recognise these effects at an early stage in order to be able to take steps to mitigate any which appear undesirable or perverse.

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

5.22 There is a further, unsurprising, issue indicated by the analysis of the results: that is, the correlations are mainly good for large universities but become progressively less good for smaller universities (or for larger universities with smaller QR allocations). This is another issue
that we take up in the next chapter.

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

5.23 In developing and evaluating future models, the Government and funding bodies would undertake to ensure that they were fully compliant with good practice in terms of equal opportunities, and that groups of researchers or types of research were not inadvertently disadvantaged.

6 **Joining up: the elements of a new system**

6.1 Successive RAEs have provided quality assurance for research funders through quality assessment, the basis for the allocation of QR funding and a rich range of information that can be used for internal management purposes within institutions. It is through their individual RAE submissions, that universities are accountable for their QR-spend.

6.2 The objective of the proposed new system is to simplify the data collection and assessment processes, building on correlations that have long been recognised. While we accept that we cannot expect metrics to reproduce in full the granularity of peer review panel conclusions, we have shown that it is possible to develop a model which could form the basis of this process for STEM subjects. We have also highlighted some options for non-STEM subjects and look forward to the outcome of the HEFCE/AHRC project exploring metrics.

6.3 One consequence of the sequence of past RAEs is that most universities now have good internal processes in place for research management – and indeed most of these involve a continuing assessment of research performance. We would expect these processes to continue to be used for allocations of QR to subjects within universities. Typically, universities will produce research plans. Such plans might include a retrospective section evaluating the degree of achievement of past plans, and would certainly look forward. We would be interested to consider whether these should be available to funding bodies as part of their background information on institutions, or whether institutions would wish to report on QR investment and general research priorities in other ways.

**Question 7: Should the funding bodies receive and consider institutions’ research plans as part of the assessment process?**

6.4 This leaves as an outstanding issue the requirement to provide international benchmarking of research quality, both to guide funders
and to provide assurance to Government that funds are being allocated effectively. We have seen in chapter 4 that there is a wealth of information that is routinely available which can be used for these purposes. We would recommend that the funding bodies should ensure that they have access to the expertise necessary to collate, present and interpret this material to their committees and boards. This would enable the funding bodies to produce a research profile for each institution. They may also wish to seek the advice of expert panels representing, for example, the dozen or so broad discipline areas and, of course, Funding Councils’ Research Committees can bring their expertise to bear. Subject to their own internal governance arrangements and any restrictions to which they are subject, the funding bodies already have the necessary powers to convene advisory panels of this sort. However, we would not expect any such panels to make additional data-collection demands on institutions.

**Question 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?

6.5 We can now draw together the elements of a possible new system.

6.6 RAE 2008 can be used for informing and managing the transition to a future funding model. In the case of STEM subjects, and at least for large universities, this should be relatively straightforward and would be phased in, with appropriate moderation, from 2009/10 in England. In the case of all other subjects, it may be necessary for funding bodies to convene panels to take a view of a broader basket of indicators both for quality assurance purposes and for the determination of funding.

6.7 There will be some detailed issues of variable definition and measurement to resolve for a funding formula. For income measures, for example, something like three year moving averages should be used in order to avoid undue turbulence. It may also be necessary to develop a suitable volume measure. These questions should be resolved by the funding bodies in the light of this consultation.

6.8 The funding bodies may wish to adjust funding models to include some additional policy variables, either to provide incentives or to discourage perverse behaviour. It would probably be sensible for the new funding model to be phased in with a year-by-year moderation to ensure that no institution lost more than, say, up to 5% a year from the QR allocation it would have received as a result of the RAE.

6.9 As the system matures, funding bodies would become accustomed to using institutional research profiles for quality assessment purposes. If it emerges that allocations to smaller institutions by formula are
volatile, then the funding bodies would consider having a size cut-off above which allocation is by formula, below which allocation is made by another mechanism, for example, by the Funding Councils' Research Committees, on the basis of research profiles and institutional research plans. The benefits of such an approach would need to be weighed against the advantages of maintaining a coherent system that applied equally to all institutions.

7 How to respond

7.1 Responses to the questions posed in this document should be posted to:

J Cutshall
1E
Department for Education & Skills
Sanctuary Buildings
Great Smith Street
London
SW1P 3BT

Or e-mailed to: rae.consultation@dfes.gsi.gov.uk

Responses should be received by close of business on Friday 13 October.
Annex 1: membership and terms of reference of the consultation working group

1.1 Membership

Sir Alan Wilson (co-chairman)      Department for Education & Skills
Professor David Eastwood (co-chairman)   University of East Anglia
Professor Philip Gummett Higher Education Funding Council for Wales
Mr John Kingman HM Treasury
Mr David McAuley Department for Employment and Learning Northern Ireland
Mr Roger McClure Scottish Funding Council
Sir Keith O’Nions Office of Science and Innovation
Mr Rama Thirunamachandran Higher Education Funding Council for England

Terms of reference
Aim
The aim of the committee is to identify a metrics-based system to assess research quality and allocate QR research funding at an institutional level.
Objectives
The new system should:

- reward excellence in research of all types, from curiosity-driven to user-focused
- encourage collaboration
- support inter-disciplinary research
- minimise the burden on HEIs
- be open to and apply equally to all HEIs
- be simple, transparent and cost-effective
- result in a funding stream to an institution
- allow HEIs to plan effectively

Timescales
The Review was announced at Budget 2006. A consultation document will be issued in May 2006. The consultation period will close in October 2006. Results will be reported to Government in time for the 2006 Pre-Budget Report.
Annex 2: possible procedures for allocating funding to arts and humanities subjects

1.1 Pending the conclusions of the AHRC/HEFCE working group, one possible solution to the difficulties posed by the use of metrics for non-STEM subjects would be simply for them to continue to be subject to periodic RAEs, whose results would then be used to populate a funding formula for QR very much as at present. Similar proposals have already been made by some people in the higher education sector. Were this course to be followed, however, it would certainly be desirable for the assessment exercise to be streamlined and simplified by comparison with the 2008 model.

1.2 Alternatively, it is also possible in principle to devise a number of funding models that might be applied to non-STEM subjects. The remainder of this annex describes one such model for the arts and humanities.

1.3 This model is confined to the metrics in data sets that are already collected and data sets that exist and may be routinely collected. There is no separate peer review of outputs, though there is panel determination of output metrics. The model is intended to exemplify one approach that could potentially be adopted for subjects where, as noted in chapter 5, a research income-based model is not fully satisfactory. This sort of approach could be developed further.

1.4 Panels would operate at super-panel level, charged with agreeing the metrics within the parameters set by the funding bodies. Weightings agreed by the funding bodies on the advice of panels. Weightings could be varied in order to incentivise (or disincentivise) various forms of behaviour.

1.5 The full basket of metrics would deliver the quality assessment. The same outcome could be used for funding, or a sub-set of metrics could drive the funding formula. This could be varied by panel area.

1.6 The model would draw on information within four categories: input metrics; volume metrics; quality and output metrics; and institutional assessment/plans.

Input metrics

Research Council income
Other research income
User-led income
Research Council success rate

Volume metrics

PhD numbers/completions
Staff with measurable outputs

Quality/output metrics

Bibliometrics
User impact
Research Council evaluation
Peer esteem

Institutional assessment

1.7 Data for most of the above are readily available. The area where panels’ advice might be most significant is in the output/quality metrics. Bibliometrics and user impact would vary by disciplinary domain. This would pick up technology transfer measures, spin-out, patents, and forms of civic engagement etc. Research Council evaluation is essentially the end-of-award data. Peer esteem picks up current RAE measures, but in a more regularised and easily generated form.

1.8 Many of these data-sets could and would be produced annually. The quality data and institutional assessment might be produced to an agreed cycle, which may or not co-include with funding adjustments if these were, say, triennial.
Appendix 3

Annex 3: Summary of consultation questions

1. Which, if any, of the RAE 2008 panels might adopt a greater or wholly metrics-based approach?
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?
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?
4. What, in your view, would be an appropriate and workable basis for assessing and funding research in non-STEM subjects?
5. What are the possible undesirable behavioural consequences of the different models and how might the effects be mitigated?
6. In principle, do you believe that a metrics-based approach for assessment or funding can be used across all institutions?
7. Should the funding bodies receive and consider institutions' research plans as part of the assessment process?
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?