University of Newcastle upon Tyne

REPORT TO THE JOINT FUNDING BODIES' REVIEW OF RESEARCH ASSESSMENT

1. The *University of Newcastle upon Tyne has found the processes involved with the Research Assessment Exercise (RAE) generally helpful* in pursuing and developing its own research strategy and agenda. In essence, a deadline by which internal assessments have to be made together with a series of mensuration hurdles against which our own performance can be compiled and assessed has proved helpful. Not least in this process has been the requirement for all UoAs to assemble clear *strategies* by which they intend to move from their current range of activity and strength to improved future performance. Also, a wider audience of academic colleagues now see more clearly the need to aim for outputs and esteem indicators of the highest quality as hallmarks of international achievement.

2. In the past RAEs have been viewed as *costly* to the institutions in terms of money, staff-time and research momentum. While there is clear evidence in our own institution that preparation is eased both by clear definition in advance of the evidence required by the RAE process and by our own efficient husbandry of records and data, for reasons of cost we would support moves to *assessment with a lighter touch*. However, for the reasons set out in our response to the "Notes for Facilitators" questionnaire, we believe that apparently simple metrical or algorithmically based approaches are not capable of evidencing the true levels of innovation and vitality which, of necessity, should be part of the judgement.

3. Our *perceived problems* with the RAE as experienced in 1996 and 2001 are:

    - The Units of Assessment are too rigorously locked on to traditional disciplinary stove pipes rather than encouraging proper assessment of the wealth of research activity which is currently either across those stove pipes or at the interfaces between them.
    - Because of the large number of Units of Assessment, many submissions involve small numbers of staff leading to significant statistical fluctuations.
    - The present outcome allows universities to emphasise *grade numbers* (i.e. their perceived level of research excellence irrespective of the density researchers) rather than any *measure of their capacity* to undertake research of the highest quality.
    - Different panels still seem to conduct their assessments in diverse ways and we detect little evidence of the passing of *best practice* from one to another. This could improve consistency between UoAs and thus improve the "trust in the system".

4. Whatever changes are made for the future, we believe that *four* issues are paramount. These are:

    a) **Efficiency of the Assessment Process**

    Any move to radically change the RAE process (e.g. to a full challenged audit) could be attractive, but it would risk being highly disruptive, whereas **keeping with the present widely-understood process with refinements would help the drive for a "lighter touch".**

    b) **Transparency of the Assessment Processes**

    The system must be sufficiently transparent and equitable from one UoA to another that it inspires trust in all those taking part in it. Thus clear criteria are required and similarly applied between cognate subjects.

    c) **Reporting the outcome**

    We find the usual research rankings published in the media tend to focus solely on grade-number and thus is misleading. **A better hallmark of research-potential and capablity is the "research power" indicator suggested by Research Fortnight (19 December 2001) in which grade number multiplied by number of staff submitted should be used as a measure of the *potential* for that UoA to undertake high quality research in the future**.

    d) **Funding the Outcome**

    Institutional financial planning against uncertain outcomes is difficult at the best of times and, if, as following RAE 2001, the financial benefits of making real progress either cannot be delivered or lead to a decrease in resource this does more to destroy both motivation and faith than almost anything else. **If future RAEs confirm excellence in universities' research then appropriate funding must be delivered by the Funding Councils, and they should anticipate progress by**

negotiating funding from the Treasury ahead of time. This reaffirms the view that has been expressed by the House of Commons Select Committee on Science and Technology.

5. Whatever system is put in place for the next RAE, **it must be one which not only rewards excellence but allows *rapid response* to reward those whose performance has improved** (e.g. those moving from Grade 4 to Grade 5 to Grade 5*, or even, as in two cases at Newcastle, from 3a to 5*). Thus, to lessen the intrusiveness of the RAE, **a lighter touch review over a longer period could be combined with intermediate reviews of those UoAs which "opt in",** and which have been identified by their institution as having made significant progress. However, this implies that **the funding models used by the Funding Councils for universities will need to be responsive to changes taking place between the major exercises**. Unless this happened, it would militate against *significantly* longer time periods between reviews than now.

6. With regard to a proposed rolling assessment (as opposed to optional intermediate assessments) our feeling is that the "big bang" exercise is preferable in terms of work load – at least based our experience of TQA. **We would prefer a seven-year gap between major RAEs with the provision for interim assessment at between four and five years.** (A seven-year cycle would also allow the Humanities to be judged on the same basis as other disciplines in terms of their research outputs.) Given the uncertainty surrounding the present cycle, it would now be preferable for the next assessment to take place in 2008 with an option for universities to submit for an interim review in 2005-6 any UoAs that they feel have made significant progress since 2001.

7. Much exciting new research is occurring at the interfaces between, and across, the traditional academic disciplines**. Thus we would like to propose a move away from the current Units of Assessment (UoAs).** There is some feeling that this imposes unnecessary stove pipes on the assessment of multidisciplinary research by panels primarily chosen from traditional disciplines. This move could do much to promote the clear articulation of interdisciplinary research in the RAE. **Thus there could be much to be said for mapping RAE Units of Assessment on to those areas in which research expenditure is concentrated** (e.g. the longer-term programmes supported by the major research-funders)**.**

8. **We would welcome a move to considerably *fewer* Units of Assessment** though this needs to be addressed with care. In some areas - for example, Engineering - a General Engineering UoA has always been available, and we can see that similar umbrella units could be created elsewhere - e.g. Clinical Studies, Biomedical Sciences, Biological and Environmental Sciences, Natural Sciences, Social and Economic Sciences, Languages, Humanities, Creative Arts. The advantages for this are:

    - The UoAs would be larger and it would be more likely that standards were harmonised across the range by there being a coherence of judgement;
    - The exercise would be simpler to administer;
    - It would allow for more informed assessment inter-disciplinary research.

    A perceived disadvantage is that it might be difficult to assemble panels who could do justice to the submissions made in a large Unit of Assessment, at least without recourse to creating a multitude of sub-panels. Individual or non-mainstream research might also find itself at a disadvantage. However, the system of sub-panels worked well in RAE2001 with regard to the three Units of Assessment that dealt with Clinical Subjects (1-3) where research overlapped and submissions might have been made to alternative UoAs.

9. We feel that *research excellence* is an *intrinsic/intensive* property while many facets regarding the outputs of research and its applicability are *extrinsic/extensive*. Thus we would like to suggest that besides the usual Grade number (4/5/5* etc), some new indicator of "capacity" is needed. We believe that the current letter system (e.g. 4A, 5B, etc) is discredited since few institutions publish the grade-letter with the number. Thus potential research-funders, only perceiving a 5/5* UoA, have no immediate way of knowing this unit's *capacity* for undertaking research (is it one person or 50?). **Rather than ranges of proportions of staff submitted as research-active being expressed as letters appended to the grade, the Funding Councils should insist that the *number* of staff submitted should be published and this expressed as "Research Power".** (See Section 4 above.)

10. We would strongly like to see the current focus on *research outputs* maintained. **However, we believe there is room for sections 5 & 6 of the submission to be more heavily structured and to carry**

more weight. (In 2001 we understand that the sections were only used to modify and check assessments primarily based on metrics and outputs.) Thus more weight could be placed on the *strategies* articulated in the RA5 section of the submission. Calls for a *critical self-assessment* of strategies, output and esteem should also be made.

11. The current metrics for the RAE are well understood and, as an addition to these, we would welcome the indicators of "research applicability" to be assessed in addition to the usual esteem indicators and published outputs. However, numbers of patents, spinout companies, citation indices, etc., can all be open as much to misuse as positive use. **Thus an evidence-supported commentary in RA 5 regarding applicability to health, wealth and quality of life might suffice to introduce this important feature into the assessment in those subject-areas where it is relevant**. However, so as not overly to confuse the enterprise and research excellence activities of universities (which can be complementary and will vary from subject to subject), we believe that **a separate assessment of third strand activities** should be considered. This could open the way for each university to decide where its balance of activities lay on the teaching, research and enterprise axes and to have them assessed and rewarded accordingly.

12. With regard to moving from research expenditure to research income, we consider that the ability to attract large sums of money may not be carried through to the ability to spend that money effectively on excellent research. We believe this move could create confusion and **we would prefer to remain with research expenditure as a measure - provided that it is called thus.**

13. We would welcome better reproducibility in the use of criteria across different Units of Assessment – though we are aware that the assessment needs of different UoAs may differ. At present it does seem that the various assessment panels have considerable room for creating criteria independently of the rest and we detect little move towards the sharing of best practice between UoA panels. We believe that this should be pursued to allow better consistency in outcomes. For example, we believe some panels read all outputs, others only looked at a sample; some graded the outputs of individuals – and thus graded individuals – while others graded the outputs of research groupings. **We would urge a consistency of approach across all subject-areas.**

14. **We would support the retention of the practice of referring all potential 5* and some 5 and 4 grades to international panels in order that the decisions of assessment panels can be moderated and the system held up to scrutiny.** However, in order that this external audit can be effective, sufficient time needs to be set aside for these international panels to operate. There also needs to be scope for these panels to recommend revision *downwards* (in the case of 5* where the international panel feels that the subject-panel has erred on the side of generosity) rather than merely to confirm the present grade or recommend elevation, as we believe happened in 2001.

Notes for facilitators

1. These notes are intended to guide those responsible for producing responses.

2. We have divided the topics for discussion into six groups. We hope this will be helpful to those organising discussions within their organisations or groupings. Four of the groups relate to the approaches to assessment outlined in paragraph 19 of the main document, and the fifth relates to crosscutting issues which will have to be addressed whichever approach is pursued. Group 6 prompts discussion of any topics that we have missed.

Group 1: Expert review

3. We have used the term 'expert review' to describe a system in which experts (possibly but not necessarily peers) make a professional judgement on the performance of individuals or groupings[1], over the previous cycle, and/or their likely performance in the future.

4. In such a system, assessors may make use of metrics, but the ultimate responsibility for decisions rests with them. Assessment may be undertaken entirely by peers or may incorporate others (such as representatives of user groups, lay people, and financial experts). The 2001 RAE was an example of this type of assessment.

[1] A grouping might be (for example) a research group, network, department, faculty, institution or consortium.

5. A variant of this system would be a combined assessment of teaching and research.

6. Suppose the funding councils have decided that they wish to retain the judgement of experts as the cornerstone of the research assessment. They are, however, willing to consider any system, however different from the 2001 RAE, so long as that condition is met. How would you advise them?

7. In providing your advice, you are asked to consider the following questions:

a. Should the assessments be prospective, retrospective or a combination of the two?
A combination; prospective assessments cannot be made without taking into account past track-record, while past track-records may not be sustained without a clear strategy for the future based on current expertise and capacity.

b. What objective data should assessors consider?
The present range of research outputs (publications), postgraduate research students - absolute numbers and funding sources as esteem indicators - grant income, and esteem indicators seem appropriate, but might be refined to include postgraduate completion-rates.

c. At what level should assessments be made – individuals, groups, departments, research institutes, or higher education institutions?
Subject-areas - but fewer UoAs than hitherto.

(See Sections 7 and 8 of reply)

d. Is there an alternative to organising the assessment around subjects or thematic areas?
Not really, in order to give fine enough tuning.

If this is unavoidable, roughly how many should there be?
Maybe 10-15: enough to cover subject-themes and to provide adequate assessment of interdisciplinary research. The use of sub-panels to give cross-cutting and subject-specific advice should be developed.

e. What are the major strengths and weaknesses of this approach?
Strengths
- **It would be able to assess different areas of universities where strength may be unequal;**
- **It would better evaluate interdisciplinary research;**
- **It would make the exercise simpler to administer;**
- **Standards of assessment are more likely to be harmonised across the range.**
- **Less opportunity for games-playing as choice of UoAs would either be reduced or removed.**

Weaknesses
- **Individual or non-mainstream research could be put at a disadvantage;**
- **A small number of UoAs might place an intolerable burden on assessors - unless the use of sub-panels were used. (But would this remove some of the advantage of fewer panels?)**

Group 2: Algorithm

8. Suppose the funding councils have decided to use an algorithm to assess research quality. The assessment must be 'automatic', leaving no room for subjective assessment. Metrics might include:

- measures of reputation based on surveys
- external research income
- bibliometric measures (publications or citations)
- research student numbers (or completions)
- measures of financial sustainability.

9. Assume the councils have not, however, formed a view on what metrics should be used or how they could be combined most effectively in an algorithm. How would you advise them?

10. You have been asked in providing your advice to consider the following questions:

a. Is it, in principle, acceptable to assess research entirely on the basis of metrics?

No. Assessment of research in the humanities and social sciences in particular would be placed at a severe disadvantage.

b. What metrics are available?

Publications and citations indices, student data, research income and expenditure, but, depending on the subject narrow bibliometric measures can provide a highly problematic basis for assessing the quality of research.

c. Can the available metrics be combined to provide an accurate picture of the location of research strength?

No. Some qualitative as well as quantitative judgement is essential so that "risky" or "innovative" research is not placed at a disadvantage.

d. If funding were tied to the available metrics, what effects would this have upon behaviour? Research effort might be distorted. Would the metrics themselves continue to be reliable?

No. Most data can be manipulated to suit a purpose.

e. What are the major strengths and weaknesses of this approach?

Strengths

- **Simple, transparent.**

Weaknesses

- **Open to manipulation**
- **Penalises humanities and some other disciplines**
- **Militates against medium and long-term research**
- **Cannot assess innovation, strategy or applicability**

(See Sections 10 and 11 of the reply)

Group 3: Self-assessment

11. Suppose the funding councils have decided to pursue a self-assessment model in which institutions, departments or individuals assess themselves. A proportion of the assessments are reviewed in detail. In a self-assessment model, the assessment is made by the assessed, although its reliability may be challenged by the validators.

12. Assume the councils have not, however, formed a view on how the assessment should be structured and how self-assessments will be validated. How would you advise them?

Any self-assessment would need to provide evidence to support the claims made, and would need to be rigorous. Any validating challenged audit would need to be carefully structured in order that the assessment is equally imposed on all universities. This could be cumbersome.

13. In providing your advice, you are asked to consider the following questions:

a. What data might we require institutions to include in their self-assessments?
Most of the current indices, but also well structured research statements that follow a prescribed format. This would assist panels and go some way towards ensuring that the assessment is equal to all universities.

b. Should the assessments be prospective, retrospective or a combination of the two?
A combination (see Group 1 above), but as it is the future research potential being funded, research strategy should be tested.

c. What criteria should institutions be obliged to apply to their own work. Should these be the same in each institution or each subject?
A range of tests could be applied for each grade, increasing in rigour for higher grades.

d. How might we credibly validate institutions' own assessment of their own work?
By a challenged audit.

e. Would self-assessment be more or less burdensome than expert review?
More cumbersome.

f. What are the major strengths and weaknesses of this approach

(See Sections 10, 11 and 13 of the reply)

Group 4: Historical ratings

14. Suppose the funding councils have decided to pursue a policy that gives each institution a rating on the basis of its historical performance and/or the value of its research infrastructure. Research would, in effect, be presumed to be strongest in those departments or institutions with the strongest track record.

15. The councils recognise that such an approach could only be used in conjunction with another system: there would need to be some way of identifying institutions whose performance was sharply improving or declining, even if the presumption was that the distribution of excellence would remain stable. It would also be possible to alter the share of the total pot provided for each institution on the basis of what had been achieved with the investment provided (a 'value for money' rating).

16. Assume you have been asked to advise on how such a system might work. In developing your advice, you have been asked to consider the following questions:

a. Is it acceptable to employ a system that effectively acknowledges that the distribution of research strength is likely to change very slowly?
No. Such a system would not encourage developing research in new areas and might encourage complacency in universities which had done well previously. There must be credit for strategic planning and dynamic change that anticipates and responds to the health, wealth and quality of life agenda.

(See Section 11 of the reply.)

b. What measures should be used to establish each institution's baseline ratings?

c. What mechanism might be used to identify failing institutions or institutions outperforming expectations? Could it involve a 'value for money' element? **How is this measured? Does this include also teaching, CPD and "third strand activities"?**

d. What would be the likely effects upon behaviour?
It might engender complacency in the case of universities that had done well in the past and a lack of ambition in universities who had not done well in the past whose efforts to improve performance would not be rewarded.

e. What are the major strengths and weaknesses of this approach?
Strengths
If used with another measure of research the use of historical data against which to make assessments could identify universities, which although performing well, were nevertheless under-performing given past research investment. Similarly universities outperforming expectations could be rewarded.

Weaknesses
- **Could lead to ossification of the system.**
- **Long lag between prosecution of the research and tangible rewards in terms of grade and funding.**
- **Difficult to provide transparency in decisions and therefore to command respect for the outcome in all universities.**

Group 5: Crosscutting themes

17. You have been asked to provide advice to the funding councils on the following fundamental issues:

a. What should/could an assessment of the research base be used for?
Investment in research-infrastructure, the shape of universities - provided that regional needs are recognised. Also economic development within regions.

b. How often should research be assessed?
Frequent enough to reward ongoing research; the five-year period seems suitable, but allowance might need to be given for the humanities and to achieve a "lighter touch". Therefore, see the next comment below.

Should it be on a rolling basis?
Advantages of a rolling assessment include a cushion against greatly fluctuating QR income between research-assessments, and the possible lessening of the impact of league-tables that can be misleading. Disadvantages include the fact that at large multi-disciplinary institutions like Newcastle some part of it will always be under scrutiny. We would generally prefer the "big bang" approach on the grounds of cost and least disruption. This should preferably take place over a longer period of time - say 7 years - which would enable assessment of research outputs in the arts and the sciences to be harmonised. In order not to penalise dynamic areas of research between RAE dates, however, we would also welcome the provision for intermediate reviews for UoAs which have been identified by their institution as having made significant progress. In order not to have standing RAE panels we suggest that this intermediate review might take place half way through the new RAE cycle.

(See Section 6 of the reply)

c. What is excellence in research?
Demonstration that the output of a subject-area is subject-leading, enjoys peer esteem, ultimately at international level, and sets the agenda for activities in the field.

d. Should research assessment determine the proportion of the available funding directed towards each subject?
Yes, provided that before any assessment the Funding Councils extract from government that increases in funding due to demonstrable improvements in research quality will be met in full. Otherwise morale can be easily damaged if excellence is not rewarded - or, as after RAE2001, actually penalised because the Funding Councils were not able to fund research fully at grades less than 5* even though this was judged as nationally and internationally leading.

(See Section 4d of the reply)

e. Should each institution be assessed in the same way?
Yes.

f. Should each subject or group of cognate subjects be assessed in the same way?
Yes, in a similar way, allowing for some variation between disciplines. Otherwise the system is open to question.

g. How much discretion should institutions have in putting together their submissions?
Very little, if gamesmanship is to be avoided.

h. How can a research assessment process be designed to support equality of treatment for all groups of staff in Higher Education?

i. Priorities: what are the most important features of an assessment process?
Transparency
Reporting that is consistent and tells the whole story
Funding that rewards excellence at national and international levels

18. We have elaborated on each of these questions below. Respondents may wish to use these notes as a basis for discussion.

a. What should/could an assessment of the research base be used for?

For the funding councils the immediate purpose of research assessment is to provide the information necessary to calculate funding levels. RAE ratings are, of course, used by others, including institutions themselves, for a variety of purposes.

What should research assessments be used for and by whom?
Should the funding councils be more explicit about what the information produced by the exercise means, and what it ought to be used for?
Yes. Universities (and the media) use and misuse the grading system for a variety of reasons. A linkage between grade and number of staff submitted needs to be made explicit. The letter used to designate a range of proportions of staff submitted as research-active (A-F) is underused and discredited. Instead absolute numbers and percentages of staff should be published, which can be translated into a "Research-Power index" (a function of Grade and staff totals) to indicate the depth of research strength as well as its quality. Perhaps the outcome of the RAE can be announced in terms of grade and Research Power.

(See sections 4c and 9 of the reply)

Should we look to design a research assessment process with the explicit aim of providing reliable management information for academic communities, institutions and other funding agencies?
Yes.

Is it the responsibility of others if they use ratings for purposes that may not be appropriate?
Yes, but better quality of information provided by the Funding Councils would lessen this effect.

Is there scope for the funding councils to work with other funding agencies– particularly the research councils – to develop complementary assessment processes which minimise the total assessment burden?

Yes.

Could the funding councils and research councils make more use of data produced by their respective processes? If so, how?

b. How often should research be assessed?

How often should research assessment take place?
Every 7 years, with provision for intermediate reviews halfway through the cycle for institutions that request it in the case of their units of assessment that are considered to have made significant progress.

(See Section 6 of the reply)

Should all subjects and all institutions be assessed at the same time or with the same frequency?
Yes.

Should clusters of subjects be assessed separately?
No. All major assessments (i.e. the heptennial ones that we suggest rather than special intermediate assessments of subject-areas of a university that request it) must involve all subjects. Not only will this ensure that subjects are assessed on a national basis periodically, and thereby equality of treatment between institutions and subject-areas, it will also allow for better assessment of interdisciplinary research.

(See Sections 7 and 8 of the reply)

c. What is excellence in research?

The purpose of research assessment is to provide information about the quality of research – but what is quality?
Demonstration that the output of a subject-area is subject-leading, enjoys peer esteem, ultimately at international level, and sets the agenda for activities in the field.

Another way of asking this question would be “what is it that distinguishes the best research”? Some might feel that this begs the question, “Is it helpful to speak of the ‘best’ research, in a way which implies that there is a magic ingredient that separates it from the rest”?

Are there different aspects of research activity (for example creativity and applicability) that each demand recognition? Did the 2001 RAE capture this?
Yes, but each can be judged by the above definition. RAE 2001 seemed mainly to capture this - at least in the UoAs with which Newcastle was involved, given the different subject-panels' ability to set appropriate guidelines.

d. Should research assessment determine the proportion of the available funding directed towards each subject?
Yes, but the system needs simplification in order to avoid wide fluctuations in the basic unit or resource between cognate subject-areas. A reduction in the number of UoAs will assist in this. The Funding Councils must anticipate that the outcome of the RAE might confirm that research has improved and negotiate appropriate additional funding from their governments.

(See Section 4d of the reply)

In devising a system of research assessment, it is important to know whether it will be required to inform the distribution of funds between subjects as well as between institutions.

There are a number of ways in which ‘subject pots’ might be determined. These include:

- the quality of UK research in the subject, benchmarked against international competition
- the volume of research in the subject that meets a given quality threshold
- a strategic judgement on the importance of the area to the UK
- a metric based upon external funding in the subject
- an overtly historical distribution which aims to retain the current balance
- a mixture of the above.

If the relative quality of research in different subjects is to be used as the basis for generating subject pots, how is this to be assessed?

A mixture of the above - especially as the third and fourth options could, if used alone, discriminate against some subjects which, while not contributing to the national economic wealth, contribute to the overall social and cultural fabric of the nation which is difficult to quantify.

e. Should each institution be assessed in the same way?

The 2001 RAE obliged all institutions to submit to the same assessment. The research outputs of a large multi-faculty institution with a strong research tradition were assessed in the same way as those of a small college with no tradition of large-scale investment in research.

Some would argue that this is an unfair competition; others that it is important for those with minimal resources to see where they stand in relation to leading units. A middle position would be that it is sensible not to compare institutions that are very different but that the system should provide a ladder of improvement so that all researchers and institutions have the opportunity to demonstrate potential.

The 2001 RAE seemed to provide this, with several instances of smaller universities holding their own against the bigger battalions.

f. Should each subject or group of cognate subjects be assessed in the same way?

Yes.

How far should the nature of the assessment be allowed to vary between subjects? **Very little.**

Should each subject community be free to define the sort of assessment most appropriate to it? **No: there should be some commonality of approach. Otherwise there may be the accusation of different standards between subjects.**

Should the funding councils go further in standardising assessment practice? **No.**

Or is the current balance about right?
The 2001 RAE was about right where panels worked to their published criteria. However, there was some deviation and perhaps different interpretation between panels. Some degree of moderation might be helpful, though a reduction in the number of panels may solve this problem.

This is not necessarily a simple choice between a greater or lesser degree of standardisation. One approach might be to define a small number of broad subject areas, and to make assessment methods within each area as similar as possible while allowing the broad groups to diverge from one another.

g. How much discretion should institutions have in putting together their submissions?

At present, institutions have a large degree of control over the content of their submissions, over who or what is assessed and by whom. This ensures that planning decisions do not make it impossible for the particular nature of an institution's research to be appropriately assessed, but it also brings significant disadvantages.

There are two alternatives: a more rigid system, or a system in which submissions are made and controlled by individuals, research groups or networks rather than by the institutions. The former risks the disadvantages of any inflexible bureaucratic procedure; the latter would arguably be unfair to institutions, as their funding would be determined by an assessment into which they had minimal direct input.

Both, however, would provide more objective results: ratings, scores or shares of the funding pot would depend entirely upon the quality of research activity as measured by the exercise, rather than reflecting the willingness of the institution to trade funding for the prestige of a high rating. They would also close the question of alleged unfairness to individuals who perceive that the decision not to include their work in RAE submissions has damaged their careers.

h. How can a research assessment process be designed to support equality of treatment for all groups of staff in Higher Education?
Submission by individuals or groups could cause all sorts of problems of quality-control and relevance.

It could be argued that the current system unfairly penalises staff at the start of their academic careers - not only those direct from postgraduate study but also those who have entered universities after successful careers outside, e.g. teachers, accountants, lawyers and other professional groups. Others who might be penalised by the assessment process might include staff whose research-careers have been interrupted by personal circumstances.

The impact on morale of staff who are not submitted as research-active is recognised. A solution might be to submit all staff, but this could be just as embarrassing for staff who carry out exceptional jobs in support of research but who themselves have chosen not to be research-active to do so. If all staff are submitted the grading system will need to be more responsive, and instead of an average grade for the University being returned a range of grades with proportions of staff in each will need to be reported.

The funding councils are committed to ensuring that their research assessment process is non-discriminatory. They are also committed to ensure that it does not reinforce a culture, wherever such a culture may exist, in which staff are disadvantaged on the grounds of sex, sexual orientation, race and ethnic origin, disability, age, religion or any other irrelevant characteristic.

Are there features of past research assessment processes which discriminate or which can be abused by those seeking to discriminate against any group?
Not as far as this University has been concerned.

Are there subtler effects, adversely affecting the legitimate interests of groups of staff, to which the design of the process contributes? What are the essential design features of a research assessment process that encourages genuine equality of opportunity for all.

All staff at this University were advised in 1997 of the importance of RAE2001 to the University, their department and themselves, and this message was reinforced over the next two to three years. By the date of submission almost all had "bought in" to the aims of the University and the RAE. As a result many were able to celebrate marked success in December 2001. In order to achieve this the guidelines of the RAE and its outcome must be published well in advance, be held to, and the outcome be transparent.

It is unfortunate that much of the spirit engendered by the preparations for RAE2001 has been dissipated by HEFCE being unable to deliver the QR funding to more than twenty UoAs at Newcastle that had risen by at least one grade to 4 or 5, with only those moving to 5* gaining their full reward. The Funding Councils, therefore, must anticipate that there will be some improvement in research quality and make appropriate provision to reward excellence.

(See Sections 4d and 5 of the reply)

i. Priorities: what are the most important features of an assessment process?

Most people would agree that a successor to the 2001 RAE ought to strive to be all of the following (and many other things besides):

- not burdensome
- rigorous
- fair to individuals and institutions
- informative
- transparent
- resistant to games-playing
- administratively efficient
- flexible (so that changes in policy can be accommodated without redesigning the entire process)
- minimally expensive.

We invite respondents to identify the three most important characteristics of an assessment process. These need not be taken from the list above but should reflect characteristics of the process rather than the philosophy underpinning it (we have asked elsewhere what constitutes excellence in research).

- **Fairness**
- **Transparency**
- **Responsiveness to interdisciplinary research**

(See Section 4 of the reply)

Group 6: Have we missed anything?

19. We invite respondents to tell us whether there are other issues or options not considered here. In particular, we would be interested to hear of any approach to research assessment that could not be described as a variant of the approaches listed above.

a) A criticism of the outcome of the 2001 exercise was the perception that there had been "grade-drift", notwithstanding the conclusions of independent reviews that quality of research output in British universities had improved since 1996. Given that it is not unreasonable to suppose that in many universities the overall quality of research will continue to increase, perhaps the present seven-point scale and its definitions need to be replaced by an alternative measure in order that degrees of excellence and research strength can be more clearly identified. This would suggest that the link between research quality (the grade) and staff numbers involved (research capacity) should be strengthened in order to give an *Index of Research Strength*.

(See Sections 4c and 9 of the reply)

b) With the growth not only of interdisciplinary research within institutions, but also of co-operative research between institutions a system of assessment needs to be devised to take this into account.

c) Funding Councils must anticipate the financial outcome of the RAE in terms of rewarding institutions which have demonstrated improvement in research, and plan accordingly. If a

longer-term, lighter touch is adopted, with provision for selected intermediate assessment, the financial model must be responsive.

(See Section 4d of the reply)

d) International panels should be retained to moderate judgements of excellence at the highest level and be empowered to recommend downward revision of grades if they feel that subject-panels have been too generous.