Sub-group to consider the role of quality assurance and evaluation

Final Report

Main findings and recommendations

- Research assessment is an essential element in delivering the policy of selectivity. It informs the allocation of funds on the basis of quality. It has also been an instrument in improving quality and it provides a degree of accountability for the research funds allocated by the HEFCE.

- There is evidence of significant improvement in the research carried out in the UK over the period of the operation of the Research Assessment Exercises (RAEs). Although direct causation cannot be proved, the timing of the change is compelling evidence of a direct link.

- If research funds are to continue to be allocated selectively, then some basis will continue to be required for distinguishing between the quality of outputs. The RAE is widely accepted within the academic community and beyond as a reasonable basis for making judgements of quality, and the results are regarded as being broadly credible.

- As far as the costs and the benefits of the RAE are concerned, it is, relatively speaking, not an expensive exercise. Returns from all institutions throughout the UK after the last exercise suggested that it cost slightly under £30 million. Extrapolation across the sector of a more in-depth analysis in one institution, including opportunity costs, suggests an upper limit for the cost of the RAE of £37.5 million. However, this still represents only 0.8 per cent of the total funds whose allocation it will directly determine.

- The group also believe that the evidence demonstrates that, far from being a leaden process, the RAE supports the dynamism of the sector as it is associated with substantial shifts of funding at the institutional and departmental level.

- A number of concerns have been expressed during the period of operation of the RAE about its possible distorting effects. In terms of the management of research and of research training, the main concerns are:

1) The RAE is thought by some to have caused the poaching of staff and the creation of a “transfer market”. We looked extensively at relevant statistics, and concluded that there is little evidence of any widespread movement of staff connected with the RAE. Compared with movement in research-intensive industries (such as pharmaceuticals), the movement of academic staff is low. We believe that staff movement is a positive consequence of institutions actively seeking to support the development of their research strategies in an increasingly competitive global research effort, and that the level of movement in the UK system may in fact be lower than is compatible with sustaining its dynamism.

2) Concern has been expressed that the desire of higher education institutions (HEIs) to demonstrate high levels of achievement, and in particular for staff to present an impressive publication record, has dissuaded them from recruiting young staff. We

considered an analysis of staff submitted for the last RAE along with other information, and have found no evidence for this.

- However, we do recognise the concern that researchers in the early stages of their career may not be well served by the current assessment process, which encourages a focus on research output rather than career development. We have three recommendations in this respect:

    i) The introduction of a facility for researchers to submit a personal statement where it will better enable them to illustrate their contribution. We believe that this will demonstrate fairness towards individuals whose achievement or potential is not adequately captured by an assessment of published outputs. Beneficiaries of such an approach would include those new to a research career, researchers taking career breaks, and those engaged on very long-term or risky research projects.

    ii) That the current assessment of research culture should become a much more explicit and focused assessment of research training and career development

    iii) That mechanisms to enhance career development should include contract researchers, and that the recommendations of the Research Careers Initiative should be embraced.

- There is a fairly widespread view that academic staff are inclined to concentrate excessively on research, to the exclusion of teaching in particular, but also of other more discretionary activities such as editorships of journals and activities for the public good. There is also concern that the RAE causes concentration on academic research to the exclusion of collaboration with industry and industrially relevant research, though we believe that the changes to the 2001 RAE will go a long way towards addressing this concern. We believe the RAE should take a broad and inclusive view of research, recognising the different dimensions of excellence and changes in the nature of research, but not extend to activities beyond its remit – it should remain focused on research. We therefore recommend that research assessment criteria should appropriately reflect the diversity of activities indicative of quality, including applicable research and 'secondary' research activities such as editorship and systematic reviews. However, we strongly support the evolution of the Higher Education Reach-out to Business and the Community Fund (HEROBC), to provide rewards for knowledge transfer activities, as we believe that alternative rewards should be created for different activities rather than seeking to address all HE issues through the mechanism of the RAE.

- There is a widespread view that the RAE encourages research orthodoxy and discourages interdisciplinary and multidisciplinary research. It is essential that panels should not give the impression that some types of research are valued over others. We noted evidence that interdisciplinary and multidisciplinary research fares no better or worse than other research in the RAE. However, the amount of interdisciplinary and multidisciplinary research being carried out is increasing substantially, and we welcome the steps taken to improve their assessment in the 2001 RAE. In addition, we recommend that the Funding Councils should continue to keep under review the number and structure of units of assessment (UoAs) and introduce new panels when this is justified by the changing research landscape.

- The UK research base is highly, and increasingly, collaborative, so collaboration cannot be discouraged by the RAE. However, there is a perception in some quarters that collaborative research is not dealt with well. Although we do not recommend directly assessing and funding research networks, we do recommend that the HEFCE should consider how it can assure researchers that collaborative activities will be properly assessed.

- In addition, we recommend that consideration should be given to recognising broadly-based areas of activity (that is, including more than one UoA) assessed as evidencing quality in both depth and breadth as 'Centres of Excellence'. Such recognition should be a necessary, but not sufficient, condition for the provision of additional funding to these centres to support collaboration with others outside the UK HE system.

- Finally, we considered alternative approaches to the assessment of research, but do not believe that any of these can entirely replace a research assessment exercise based upon peer review. We therefore recommend that the RAE, with a number of significant modifications, should continue to be the mechanism to support the policy of selectivity.

Background

1. The 1980s saw an explicit acknowledgement of the role of science in a national strategy for economic competitiveness, against a background of declining resources for HE. In response, the University Grants Committee (UGC) developed an explicit and formalised assessment process – the Research Assessment Exercise (RAE) – to replace the less transparent policy of selectivity it exercised through its subject-based committees. The mechanism was developed with two objectives in mind:

- to identify what research was being funded
- to identify what outputs were being generated by this funding, that is, what research was being published.

2. In this sense the RAE is a classic quality assurance process, with a hierarchy of enhancement, standards, assessment and audit.

3. The value and evaluation debate has become of increasing concern to all developed countries as they seek to compete in an increasingly global research environment.

4. This report assesses the continuing need for, and fitness for purpose of, the RAE, and discusses possible alternatives. In undertaking this work, we have sought particularly to find evidence to test, and then support or reject, views on the stated effects of the RAE.

5. Membership of the sub-group is shown at Annex A and the terms of reference are shown at Annex B.

What is research assessment for and to what purposes has it been put ?

6. The principal role of the RAE is to provide an explicit basis for the Funding Councils to allocate funding selectively to support existing or developing excellence, but it also provides accountability for research funding. Research undertaken as part of this review has also confirmed that it has increased the quality of research being undertaken.

7. RAE ratings are sometimes used by research students seeking to identify the best research training opportunities, but this practice is not encouraged as the ratings may be a poor proxy for the quality of research training provided. Most higher education institutions (HEIs) make use of RAE ratings for public relations purposes and to promote themselves to the outside world. The study we commissioned from PREST to support our work has found that post-1992 institutions find RAE ratings a valuable way to highlight departments that have achieved research excellence.

8. In discussion with the Research Councils, PREST heard that departmental RAE ratings ares not considered when discussing the merits of an individual proposal. However, it was acknowledged that it was used as a proxy for general resourcing and infrastructure provision and as such was taken into account in assessing the deliverability of projects. Some Research Councils also acknowledged that RAE ratings were taken into account when assessing the research training environment. In particular, the Economic and Social Research Council (ESRC) noted that RAE ratings were considered when assessing its long-term support for ESRC Research Centres.

9. A number of R&D led industrial companies also acknowledged using RAE ratings as a guide to the selection of academic research collaborators and recruits. However, typically, the RAE is used as one indicator among several. As might be expected, it was a factor in identifying potential collaborators, but it did not carry the same weight when the company had existing links with a researcher or research group.

Has research assessment encouraged an improvement in research quality?

Research activity

Average RAE ratings have increased progressively with each exercise, for instance, between the 1992 and 1996 exercises the number of 5 rated departments increased from 348 to 573, the proportion of staff in 5 rated departments increased from 22 to 31 per cent, and the average rating increased from 3.4 to 3.7. However, as this comparison is self-referential, it cannot be said to demonstrate definitively an increase in quality, nor the role of the RAE in promoting such an increase.

11. However, analysis undertaken by the Higher Education Policy Unit (HEPU) at the University of Leeds as part of the research review indicates that the RAE has brought about

a significant improvement in general research quality. Evidence from HEPU interviews with senior management on the balance of activity and volume of outputs, supported by quantitative data, indicates less wastage, more purpose, and increased support for the most promising ideas. The effect of these changes, at least in part driven by the RAE, is most evident in the performance of the top sixth of institutions (a proxy for UK peak performance): their research has significantly increased in impact (the peak is higher) and they make up a greater proportion of UK publications (the peak is broader).

12. Analysis has also revealed a shift in grant funding, numbers of students and postdoctoral research associates, and relative outputs compared to funding indicating that quality and productivity have increased in what were formerly the most productive areas of the sector (though of course the production of papers is not the only, or a definitive, measure of productivity or quality). The UK average has shifted less, probably because of the increasing competitiveness of the global research effort and the increasing tendency for researchers from all countries to seek to publish in anglophone journals.

13. It is of course impossible to attribute this change directly to the RAE. However, the timing of the effect is compelling evidence. Changes in research will have two phases, relating to inputs and outputs. As one would expect if the RAE was driving change, there is a change in the input measures in 1986 but change in the outputs some time later.

14. It may also be argued that the effects of the RAE are simply reinforcing a 'big is beautiful view of science. It appears from work undertaken by HEPU for the group exploring the role of selectivity and the characteristics of excellence that critical mass is significantly correlated with quality in most science-based subjects, particularly in the experimental sciences. However, we have noted the conclusions of the selectivity sub-group that there are many institutions with small departments functioning at an international level. In mathematics, for instance, it is clear that there is no scale effect.

15. In addition, HEPU have provided evidence of increased relative output, greater growth in national and international collaboration in the UK than in other countries, increasing interaction within industry, and greater correlation of industrial and public funding. This suggests that UK researchers are much more likely to communicate the outputs of their research.

16. This is not to suggest that there is any room for complacency in efforts to improve the performance of the research base, and there are areas where further significant improvement is required. However, it does indicate that there has been a marked change in research activity and quality since the mid-1980s as a result of the quality assurance mechanisms that were introduced. In particular, peak performance has improved and broadened relative to the world average.

Research training

17. Research Council reports, United Kingdom Council on Graduate Education studies, and interviews undertaken as part of the HEPU study indicate that the RAE has also acted as an incentive for HEIs to improve research training, as institutions have sought to develop their future research potential.

18. Research training has improved as a result of institutions investing more financial and management resource, and adopting a more systematic approach to identifying researchers' training needs and the ways in which these can be met. This has produced much better PhD completion rates, and improved the quality and quantity of papers published by research students.

19. We recognise there is no way of defining the relative influence of RAE and Research Council driven changes on research training. It seems likely that the Research Councils had more specific influence where their funding was a significant component, whereas the RAE supported a generic system change.

Has research assessment encouraged an improvement in research management?

20. Previous evidence suggesting that institutions had become more strategic in response to the RAE was mostly anecdotal. The HEPU study has produced evidence to demonstrate that there has been a real change, in particular:

- major changes in attitude and strategy, underpinning significant moves towards more conscious and active management of the research environment

- improved management of the research environment (though research itself is still comparatively 'unmanaged') in both the social sciences and the natural sciences

- development of structures that ensure better accountability and can support a more strategic approach to the development of research policy (for example establishing research committees at both institutional and departmental or school level)

- large institutions have increased their share of research students and dedicated research workers more than they have increased their share of funds, suggesting that research management has promoted improved efficiency

- institutions with a generally high quality research base are focusing management attention on, and making financial investments in, 4-rated departments in an attempt to raise them to 5/5* status

- institutions with a number of poorly-rated departments often seek to increase quality by merging less successful departments with those that were highly-rated, in the expectation that exposure to the new culture and management will raise the overall level of the merged entity – thereby creating viable and dynamic units capable of securing appropriate on-going levels of research funding. This has had the added benefit of creating broadly-based interdisciplinary and multidisciplinary units.

What approaches are other countries adopting to research assessment?

A study by SPRU (Science and Technology Policy Research, University of Sussex) commissioned as part of the review, found that countries generally adopt one of three approaches to determining university funding (the country-by-country analysis is summarised in Annex C). They vary considerably in the extent to which research funding is informed by performance.

a. *Research funds allocated, at least in part, on the basis of some sort of research evaluation.*

In addition to the UK, countries such as Australia, Poland – and China in respect of the Hong Kong-based universities – are determining research allocations based on research performance. In the UK and Hong Kong, the approach is based on informed peer review undertaken for the purposes of determining the allocations, while in Australia and Poland existing performance indicators are used.

b. *Formulaic allocation using research performance indicators, but also including parameters related to the size of the teaching and learning activity.*

Research funds are allocated with funding for teaching and learning as part of the general institutional funding. Countries such as Germany, Italy, Sweden and Norway follow this kind of approach. Finland and Denmark allocate the largest share of general research funds on the basis of the amount of teaching and learning activity, but a small portion of core research funding is allocated on the basis of the teaching and research performance of the university.

c. *Negotiation between institutions and the Government department responsible for university funding.*

In Austria this model is applied without undertaking any form of research evaluation, while in France information gleaned from an evaluation of university teaching, research and management activities is taken into account during the negotiation.

21. There is also a fourth category in which research assessment is carried out but it does not inform funding decisions. In The Netherlands, evaluation of research performance is used to improve the research quality of the university. Funds are allocated using a formula based on student numbers.

22. The report from SPRU notes that, although only a small number of the countries studied are currently using performance-based approaches to university research funding, most are either in the process of implementing some form of performance-based allocation, or are

considering doing so. The benefits and problems associated with the different approaches to research assessment in comparator countries are shown in Annex C.

What can be learnt from the history of research assessment and successive RAE cycles?

23. PREST have found that one of the reasons that the RAE has brought about positive changes in behaviour is that it is generally held to be sound. The ratings produced are widely accepted by the research community as benchmarks of quality against a universally understood 'gold standard', while its processes are seen as robust, credible, and transparent.

24. That said, we note that for some time commentators have expressed considerable concern about the distorting effect the RAE has on academic activity (many of these concerns were raised after the very first exercise). It is alleged to:

- encourage short-termism
- discourage interdisciplinary and multidisciplinary research
- discourage applied research
- reinforce research orthodoxy
- devalue teaching and other learned activities, such as editorship
- encourage premature publication.

25. We also note that some researchers, particularly in the arts and humanities but also in some practice-based subjects, consider the RAE to reflect a notion of research that is irrelevant to their subject.

26. It is also held to have undesirable management effects such as encouraging an academic transfer market, discouraging the employment of young researchers, and creating stress.

27. In addition, there have been, and continue to be, concerns about the assessment process – that the criteria for assessment are not clear or consistent or necessarily appropriate across disciplines, and that the exercise places too much emphasis on retrospective quantitative performance indicators.

28. We noted the concern, expressed by several professional organisations to PREST, that the lack of feedback available from RAE panels may act as a barrier to self-improvement. It was acknowledged that there would be considerably more feedback following the 2001 RAE. However, we recognise that the Funding Councils are mindful of the danger of reinforcing research orthodoxy if research assessment becomes too formative.

29. It is also clear from the PREST study that a belief continues to persist that the panels will use simple impact factor analysis in the 2001 RAE. We welcome the efforts of the Funding

Councils to ensure that the assessment of published output reflects the inherent quality of the work, and is not simply inferred from the reputation of the journal.

30. In general we were concerned that respondents' perceptions of the RAE did not seem to reflect the extent to which the RAE 2001 would be different from previous exercises.

The effect of the RAE on the behaviour of organisations and individuals

31. Responses to the "Call for Evidence" issued as part of this review indicated that, in general, the RAE is well tolerated by the sector. It is seen as having increased the quality of research within institutions, the results are seen as largely credible, the ability to benchmark performance is seen as a clear benefit, and the cost and effort involved are compared favourably with the Teaching Quality Assessment (TQA) and the preparation of bids for special funding initiatives. However, it is clear that in some cases stated support for the RAE is actually support for the continuance of the dual support system.

32. Pre-1992 institutions stressed the need for a mechanism to support selectivity based upon excellence, in order to maintain world class research capability in the UK. Post-1992 institutions are somewhat less content with the operation of the policy of selectivity and the mechanism used to support it, but nonetheless appreciated the access it has given them to research funds.

Staff movement and retirement

33. The PREST study found a widespread belief that there is an active staff transfer market, and that the RAE is at least partially responsible for it. Given that the promotion of a transfer market is one of the most often cited negatives of the RAE, we have devoted considerable attention to exploring the effect of the RAE on staff movement.

34. We reviewed a detailed analysis of RAE-related staff movement undertaken by the HEFCE in February 1998, using information from the 1996 RAE database. It is obviously extremely difficult to disaggregate any promotion of staff movement due to the RAE from that which one would expect in a growing and dynamic HE sector. However, one can look relatively easily at the level of staff movement over time and see if there is any correlation with the timing of the RAE. It seems clear to us that while there is an increase in the staff moving in the year immediately before the RAE, this was less than 500 staff out of the total recorded research active staff population at the time of 50,430. On this basis, the RAE could be said to be associated with, at most, the movement of 1 per cent of the research active staff population during a period when the sector grew by 25 per cent.

35. We also reviewed an analysis undertaken by the HEFCE which looked at the movement of staff returned to the RAE compared with other staff, using HESA data. It may have been the case, for instance, that although the overall level of staff movement was not high, the overall figures disguised significant patterns such as the movement of different types of staff, or movement between particular types of institution. In fact, it is clear to us from this

analysis that the percentage of staff moving to another HEI is significantly higher for non-RAE returned staff than for RAE returned staff for the years 1994-95, 1995-96 and 1996-97. Also, the percentage of research active staff moving to another HEI in the year after the RAE was significantly higher than in the two years before the 1996 RAE. This appears to confound the argument that the RAE is the main driver of staff movement in the system. It was also clear that there was no significant pattern in the movement of staff between different types or sizes of institution.

36. We also re-examined a Royal Society analysis of the movement of its research fellows, which it was suggested demonstrated an increasing level of movement. However, the sample size is too small, and the overall level of movement too low, to draw any firm conclusions from it.

37. As part of our commitment to developing a thorough understanding of the effect of research policy on institutional behaviour and how individuals respond to this, we commissioned a further analysis of staff movement. In particular we wished to explore the suggestion that in the run up to an RAE institutions suffer disproportionate loss of research leaders who are key to the delivery of the institutions research strategy. This analysis demonstrated that for the three years 1995-96, 1996-97 and 1997-98 research-intensive departments had a lower or equivalent ratio of people recruited from outside the institution moving into senior posts (professors and senior lecturers) than the sector average. Again, this seems to us to confound the notion that the RAE drives undesirable levels of staff movement.

38. The analysis did show that in 1998-99 there was a small increase in the level of recruitment from outside institutions: however this amounted to only 304 staff. This may be attributable to the imminence of the 2001 RAE. However, it is just as likely that it is the result of investment funded by additional monies following success in the 1996 RAE. The changes in funding would have occurred in the academic year 1997-98, taking account of time for funds to filter down to academic departments, and the time lag inherent in the recruitment process, it would be 1998-99 by the time institutions were actually in a position to recruit staff. It would be expected that high performing research areas would use part of their additional funding to invest strategically in increased volume and research leadership.

39. These analyses do suggest that the RAE may have some limited effect on the timing of recruitment, although it is likely that this recruitment would have happened anyway. This may explain the persistence of anecdote about a 'transfer market', particularly since these individuals will have a high profile within and outside institutions and be a key concern of senior managers.

40. However, we believe that the main driver of staff movement is the increasing emphasis on the development of effective research strategies in an increasingly competitive arena, which is a global phenomenon, though the RAE may bring this more sharply into focus.

41. This view is confirmed by the USA experience where there are a number of studies showing that, particularly in some areas, there is substantial staff movement.

42. In addition, interviewees have suggested that staff mobility was higher in the 1970s than it became in the 1980s.

43. Analysis undertaken by PREST also suggests that the scale of any movement produced by the RAE is not large. Their key findings are as follows:

- the proportion of staff moving is generally low, ranging from 1.6 per cent to 2.6 per cent per annum, although there is an increase of 0.9 per cent in the year immediately preceding the 1996 RAE

- the proportion of movement accounted for by professors increased very slightly immediately prior to the 1996 RAE (from 5.6 per cent to 5.96 per cent) and fell immediately after to 5.73 per cent in 1996-97. It has subsequently risen again in 1998-99, but at 1.4 per cent of the professorial population the level of movement is low in absolute terms and lower than it was in 1996

- the number of staff moving out of the HEI sector is small and shows no clear pattern

- the percentage of staff retiring or moving out of active employment rose in the year following the 1996 RAE from 1.84 per cent to 3.30 per cent.

44. There were concerns expressed following the 1996 RAE that the academic population was ageing. Subsequent analysis by the HEFCE showed that this concern was generally misplaced, as the proportion of academic staff aged between 25 to 34 was only just below the percentage for the working population as a whole (30 per cent compared to 31 per cent). However, the proportion of older staff was marginally higher than that in the population as a whole, which may explain why there was a small increase in retirements after the 1996 exercise. We note that the UK situation is dramatically different from that in France where PREST analysis indicates 50 per cent of current academic staff members are due to retire between 1998 and 2005.

45. The picture of a generally static population is confirmed when the level of movement in HE is compared with that in R&D intensive industrial sectors. The level of HE staff turnover for the years 1996-97, 1997-98 and 1998-99 is 6 per cent. This is at the bottom of the normal range for a research-intensive company, which is generally 5 to 12 per cent. However, the rate of movement in such industries for people who have an advanced scientific qualification – who may be considered to most closely equate with academic staff – is often higher than 12 per cent. Rates approaching 20 per cent per annum are not unusual in specific areas where particular skills are at a premium.

46. On this basis, we concluded that the rate of movement in HE is lower than is consistent with the development of a dynamic sector. It is likely that competitiveness in research would produce such movements anyway, with or without the RAE, though their timing might be different. This view is supported by data from the USA which show an active level of staff movement, in the absence of a national assessment exercise. Though of course this could be related to the ability to attract funding.

We are also inclined to agree with the conclusions of PREST that the movement of staff is often effective in placing the best researchers in the well resourced and managed departments. We believe that special funding initiatives such as the Joint Infrastructure Fund (JIF) can be far more disruptive than the RAE in promoting the movement of key members of globally competitive teams between institutions. We recognise that there are concerns that the imminence of the RAE is cited as a lever by staff to increase their salary. We have little sympathy with this view: staff are rarely motivated simply by personal reward and there are many other factors more likely to be significant in increasing salaries, including the fact that in a people business good people will always attract a premium.

48. However, we conclude that, whatever the evidence, the belief in the existence of an unhealthy level of movement will prove difficult to dispel entirely, as research networks are pervasive and even a relatively small number of staff moves can create an impression of an active and de-stabilising transfer market. Also, with the RAE on a four or five year cycle it is difficult to dissociate any movement from the RAE that has just gone or the one that is coming. We do, though, note the widespread approval of the RAE 2001 one-year rule whereby staff who move in the year prior to the RAE may be cited by both the exporting and importing institutions.

49. Nevertheless, we have two particular concerns: that talented younger researchers may be tempted to take senior posts in institutions that cannot provide them with challenging research or teaching opportunities; and that flows within the system are not necessarily balanced. In particular they can be inhibited in one direction by labour market inflexibilities, notably property prices in the South-East of England.

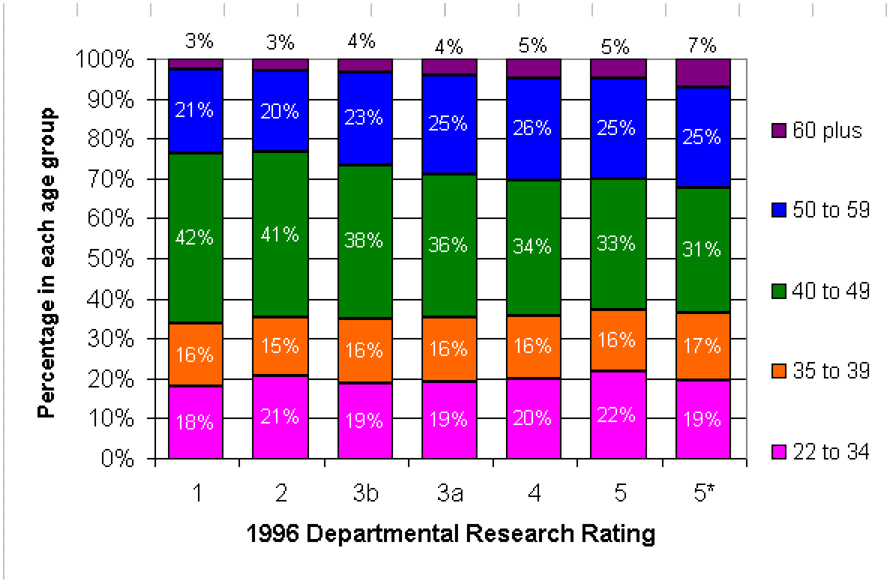
Treatment of new or young researchers

50. Interviews with senior institutional managers evidenced some concern about the effect of the RAE on new researchers. In particular it was suggested that the RAE is raising the typical age of recruitment from mid-20s to early 30s as a result of a requirement for appointees to have four publications, especially in the latter half of the RAE cycle.

51. However, the age of staff submitted to the 1996 RAE, shown in Table 1 below, shows no relationship between the age profile of departments and RAE grades, which would be expected if such a phenomenon was having a significant effect. We are therefore sceptical about the significance of the suggested effect on the sector as a whole.

Figure 1: Age profile of Researchers by Departmental Research Ratings

Extracts from the 1996 Research Assessment Exercise Database



52. To explore this important issue further, we commissioned a direct analysis of the age profile of staff recruited to research-intensive departments compared with the overall age profile of staff recruited. The data, presented in Table 1 below, show no evidence of discrimination in the recruitment of younger staff. In fact, research-intensive departments recruit a slightly higher than average ratio of younger staff. The same pattern was found at the broad subject level.

Table 1: Ratio of proportion of young researchers recruited to research intensive departments to proportion recruited to the sector overall

	1996-97	1997-98	1998-99
Young recruits in research (Yr)	748	730	786
Total recruits in research (Tr)	3088	3076	3646
All young recruits (Y)	1224	1314	1317
Total recruits (Y)	5720	6267	7163
Ratio [Yr/Tr] / [Y/T]	1.13	1.13	1.17

53. Since the need to have four publications was often cited as the constraint on recruitment, we asked for an analysis of the associated publications data. This showed that, on average, researchers over 29 submitted 3.8 publications to the 1996 RAE, whereas younger researchers submitted 3.5. As the average is lower this would seem to be evidence that institutions are willing to recruit and submit staff who have fewer than four publications.

54. As part of their work for the review, PREST have also undertaken an analysis of the age of recruitment using HESA data. They found a small increase in the age at which staff are recruited. However, they suggest that a number of factors other than the RAE could be affecting the age at which staff obtain their first established position:

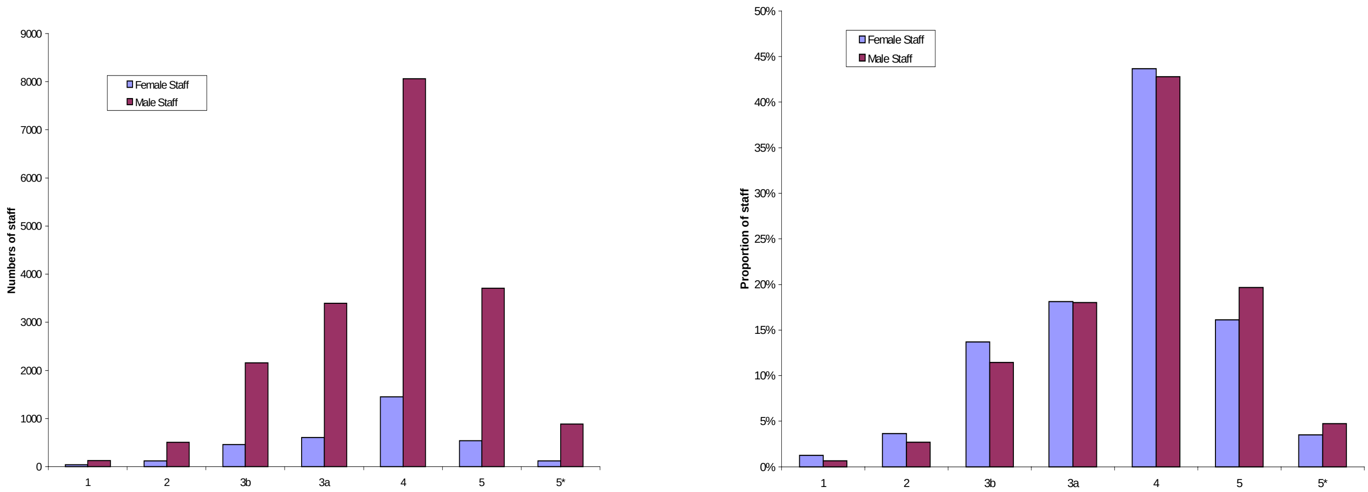
- the growth in the number of postdoctoral contract researchers from which pool of talent appointments are made
- the availability of research fellowships funded by charities or other organisations with significantly better remuneration packages than those offered by lectureships
- the cost to institutions of early-retirement packages.

Diversity and gender issues

55. Extensive interviews with senior institutional managers and researchers conducted by PREST have suggested that the RAE reflects, rather than causes, the more general problems which women face in pursuing an academic career (it was also noted that many of the problems that women face pre-date the RAE).

56. The pattern of distribution of researchers by RAE grade is shown in Figure 2below. The levels are broadly similar for women and men, though the numbers of women are dramatically lower.

Figure 2: Distribution of male and female staff by RAE grade



57. However, this is an important and complex issue that demands thorough exploration. We therefore particularly welcome the fact that the sub-group looking at the development of research people are exploring this issue in considerable detail, including an in-depth analysis of data currently available and the collection of new data as appropriate.

58. The Funding Councils have already addressed equal opportunities issues to some extent for the 2001 RAE: women constitute 20 per centof panel members though they hold only 9 per cent of professorial positions.

59. We welcome the changes for the 2001 RAE which will see an enhanced facility to record the individual circumstances of staff where this is pertinent to the assessment process. Female researchers interviewed by PREST made it clear that the panels must take account of the effect of career breaks on continuity of research activity and research output. They also pointed out that women may be less able to establish an 'international' presence as they may be unable to travel as much as men to conferences, etc.

60. We welcome the fact that the forms submitted to the RAE 2001 dealing with institutional strategy will require institutions to address recruitment and staffing strategies and will be more structured, so panels will have a better basis on which to compare and assess submissions. However, we urge the Funding Councils to consider what more can be done for the 2001 RAE and beyond, and in particular to note the recommendations detailed below.

Conclusions

61. We believe any further exercise should see a further enhancement of the assessment of people. This should include better assessment of individuals who have had a career break, but also those who are engaged in long-term research, have recently embarked on an academic career and have not developed a research portfolio, or have undertaken research that was of high quality, but was risky and did not prove to be fruitful in terms of publications.

62. We recommend that a facility should be developed to permit a personal statement in respect of such individuals to be included in a submission where appropriate.

63. We also recommend that the current assessment of research culture should become a much more explicit and focused assessment of research training and career development. Separate ratings of the quality of research training and career development could be produced and published, either as part of the current assessment process or in a linked assessment process. This could draw on other sources of information about the quality of research training such as that produced by the Research Councils, major charitable funders and industry. An alternative or complementary approach would be to encourage institutions to adopt policies on the provision of research training and career development which could be assessed at an institutional level.

Peer review and panel behaviour

64. Given the central role of peer review in the assessment process, we asked PREST to review its strengths and weaknesses, and in particular to explore, including with RAE panel chairs, whether the process may pre-dispose to conservative judgements about the quality of research activity. It has been suggested that senior professionals are innately conservative and that assessment group dynamics may reinforce this tendency.

65. The most authoritative document on peer review remains the report of a working party commissioned by the Advisory Board for Research Councils in 1990 (Boden et al). Many of its main conclusions were specific to the use of peer review to inform the award of project grants. Those with relevance to the RAE process are shown below:

- peer review is ‘the only practicable method of assessment’ across all research areas
- quantitative measures can legitimately inform peer review judgements; they cannot however replace them
- the use of international referees should be encouraged throughout the system
- the peer review process should maximise transparency
- peer reviewers should continue to be drawn principally from specialists in the appropriate fields, but where experts from other disciplines or sectors have relevant expertise, efforts should be made to recruit them
- feedback should be provided as a matter of course
- steps should be taken to ensure consistency across similar assessment processes.

The report also gives the following list of ‘necessities’ for any system of peer review:

- the need to review practices continually bearing in mind the inherent fallibility of peer review
- the need to expose peer review practices thoroughly to the academic community, including providing information on peer review, and strategic and managerial, inputs to decision-making
- the need to choose peers well
- the need to provide a 'return' to the research community in the form of feedback from decisions
- the need to demonstrate fully the equivalence in intellectual standard across all modes of support
- the need to substantiate initial appraisal decisions by appropriate monitoring and evaluation.

66. There is a substantial amount of literature on peer review and we asked PREST to review this. PREST concluded that there is little evidence of gender, age or institutional bias. This confirms the findings of recent studies of their grant-giving practices by the Wellcome Trust and the Research Councils.

67. There remains a concern that the inherent conservatism of panels will work against truly original research. However, we believe that RAE panels are less disposed to conservatism than project funding committees because:

- of the range of information provided to inform judgements
- of the transparency of the process
- assessments are made, and recorded, against specific criteria
- assessments are made of body of work, from a number of people
- it is not a binary decision (to fund or not to fund)
- the funding outcome is not so direct or immediate as with project funding.

68. The requirement of the panels to demonstrate good judgement will be reinforced in the 2001 RAE by the use of international peer reviewers and umbrella panels in cognate areas.

Publication behaviour

69. We are inclined to give credence to the findings from the PREST survey of editors – that the RAE affects publishing behaviour, but not significantly. We note that there is no evidence of UK researchers increasing the number of publications or 'salami slicing' them with a consequent deleterious effect on quality. Instead, there appears to be a global trend for shorter and sharper articles.

70. However, we notethe view voiced by some editors that the RAE deters review articles and refereeing – partly due to a belief that insufficient credit is being given to these activities.

We believe that both these activities should be appropriately recognised in the RAE process. We recommend that the Funding Councils ensure that the diversity of HE research-related activities is properly incorporated in the quality evaluation process, and that they take responsibility for communicating to the academic communities that this will occur.

Recognition of collaboration

71. We noted the concern expressed in the Call for Evidence that the RAE does not sufficiently recognise collaboration. It is generally acknowledged that the networks of UK researchers are extremely well developed, and support considerable collaborative research activity. This is illustrated in Table 2below.

Table 2 Number of collaborative papers published by UK researchers per annum and percentage of the total numbers of papers published each year that were collaborative from 1981 to 1994.

Number of collaborative papers

Sector	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	Total
Education	5728	6050	6244	6598	7263	7537	8111	8341	8834	9763	10562	12183	12626	13814	123654
Medical	3016	3101	3269	3466	3887	3924	4116	4296	4674	4997	5102	5749	5942	6108	61647
Research Council	1315	1418	1573	1622	1685	1887	1961	1964	2035	2303	2348	2646	2703	2851	28311
Industry	899	979	1080	1045	1156	1173	1340	1426	1522	1705	1809	1976	2063	2125	20298
Government	432	404	442	453	485	527	520	565	568	685	709	772	815	833	8210
Non-profit	204	252	274	329	328	329	358	391	412	506	512	647	606	718	5866
Unknown	40	45	46	66	60	43	48	57	61	46	102	103	98	141	956
Total	8659	9088	9537	10059	10970	11375	12058	12567	13417	14732	15674	17904	18418	19814	184272

Percentage of Sector Papers

Sector	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	Total
Education	31	32	33	35	35	37	39	41	42	44	46	49	49	50	41
Medical	40	40	41	42	42	43	44	44	46	47	48	50	52	53	46
Research Council	37	39	40	44	42	44	50	52	53	56	58	61	61	65	50
Industry	35	38	40	40	42	44	50	49	51	54	57	58	62	63	50
Government	31	32	33	37	37	39	41	45	45	51	51	53	55	55	43
Non-profit	42	45	48	52	53	52	57	57	58	60	60	68	65	71	58
Unknown	37	32	34	46	44	41	52	52	54	51	63	58	63	65	51
Total	28	29	29	31	31	33	35	36	38	40	41	44	45	46	37

72. It is also clear from OECD data, shown in Table 3 below, that UK researchers collaborate extensively with others on an international basis, despite the physical separation from Europe. Co-publications by UK authors with researchers in other countries almost doubled between 1985 and 1995, the biggest increase of any European country except Spain.

Table 3: Percentage of internationally co-authored papers, for UK researchers compared with those in France, Germany Italy, Spain and Japan

	1985	**1990**	**1995**
France	20.0	25.7	33.3
Germany	18.4	25.4	32.2
Italy	20.0	26.8	33.0
Spain	13.5	23.5	29.0
Japan	7.1	9.8	14.4
UK	**14.0**	**18.8**	**26.2**

73. Given the level of growth in collaboration over the recent past, we do not believe that the RAE acts as a significant barrier. However, we recommend that the Funding Councils do more to ensure that researchers are confident that such work is properly assessed by the RAE.

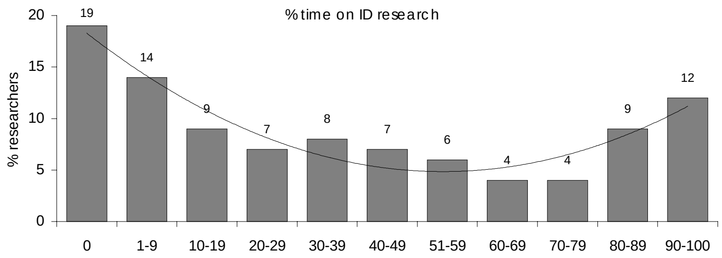
74. However, we do not believe that collaboration per se should be seen as a quality indicator. We noted the conclusion of the selectivity sub-group that collaboration may be an indicator of peer perceptions of excellence, because collaboration involves investment and costs and these are only tolerated where there is an expectation of benefit. Senior managers, in their interviews with HEPU, often described centres of excellence as those that attracted international visitors and employees, and said that industrial collaboration reflected perceptions of added value. The HEPU study found that in several institutions external collaboration was used as a criterion for promotion. However, we believe that to establish collaboration as a direct quality measure would give rise to collaborations seeking recognition under the system, rather than those seeking to enhance research.

Recognition of interdisciplinary and multidisciplinary research

75. Despite the widespread observations that interdisciplinary and multidisciplinary research are increasing within universities, it is extremely difficult to capture and measure this general trend. Perhaps one of the few robust sources of information is a study undertaken on behalf of the Funding Councils and Research Councils in response to the recommendations of the National Committee of Inquiry into Higher Education (Dearing Report). A questionnaire asked researchers to indicate the amount of time they spent engaged in interdisciplinary and multidisciplinary research. The responses, shown in Figure 3 below, clearly demonstrate that

almost all researchers consider themselves to be engaged in interdisciplinary and multidisciplinary research to some extent.

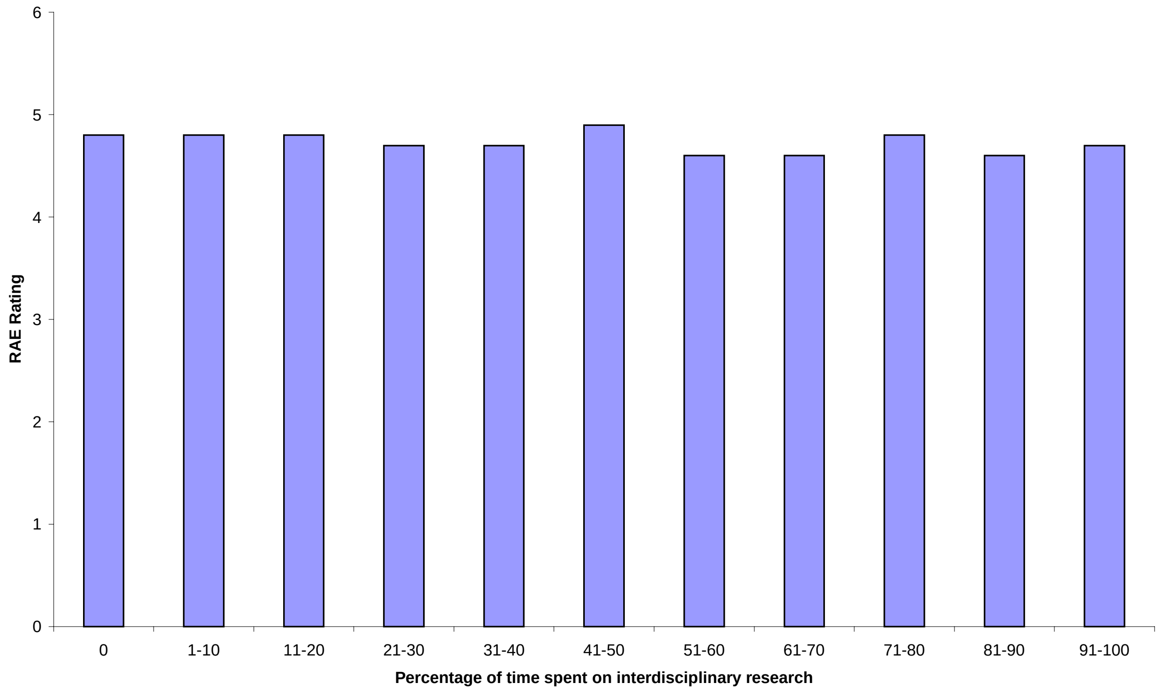
Figure 3: Percentage of time that researchers spend engaging in interdisciplinary and/or multidisciplinary research



76. These data are confirmed by evidence from the PREST survey of universities and researchers, undertaken as part of this review, which also showed that the overwhelming majority (97 per cent) of research staff collaborate with researchers from another discipline. Most respondents indicated that they believed that research would become more interdisciplinary and multidisciplinary in the future.

77. We noted the absence of any evidence that interdisciplinary and multidisciplinary research are inappropriately assessed by the RAE, despite the fact that the exercise takes a discipline-based approach. We found most compelling the analysis, reproduced in Figure 3 below, from the study of interdisciplinarity following the Dearing Report. This showed no relationship between the level of interdisciplinarity of researchers and the RAE rating for the submission of which they were a part.

Figure 4: Percentage of time researchers spent on interdisciplinary and/or multidisciplinary research versus the rating of the 1996 RAE submission of which they were a part.



78. However, we noted the continuing concern about the assessment of interdisciplinary and multidisciplinary research. We therefore welcome the changes for the 2001 RAE, which we believe will reassure researchers that interdisciplinary and multidisciplinary research will be properly addressed:

- panels to include interdisciplinary and multidisciplinary researchers
- assessment criteria specifically address how interdisciplinary and multidisciplinary research will be treated
- enhanced mechanisms for cross-referral of submissions
- monitoring of the treatment of interdisciplinary and multidisciplinary submissions.

79. We believe that the PREST study may reconcile the findings that the assessment of interdisciplinary and multidisciplinary research is not prejudiced with the continuing negative perception in the sector. The study suggests that the problems cited with interdisciplinary and multidisciplinary research are not general but relate to three distinct areas, and are a function of panel behaviour rather than the boundaries between UoAs:

- where specialised departments in a field are covered by a single panel – for example textiles, paper science, materials and corrosion all fall under 'materials'

- where a view develops in the sector that in order to get the highest ratings a submission must show a breadth of activity within a discipline

- when a large area involves more than one discipline, for example “development” is spread across economics, politics, and sociology.

82. It was notable that in the interviews PREST held with representatives from interdisciplinary research centres the RAE was not held to be a significant issue..Aligning UoAs with the research landscape

83. There is clearly a need to ensure that changes in the research landscape are appropriately reflected in the structure of the RAE. Some of the ways in which research is changing are documented in the next section.

84. The PREST study found that individuals, departments/faculties or institution managers believe that there two specific areas where there are structural problems that need to be addressed:

- lack of clarity about the most appropriate panel to which to send interdisciplinary and multidisciplinary submissions
- research areas that are subsumed within much larger UoAs.

85. We believe that there are a number of ways to address these issues, some of which have been implemented for the 2001 RAE:

a. Restructuring existing units of assessment to create larger thematic assessment panels with the expertise to assess work crossing narrow disciplinary boundaries. UoAs 12 and 14 have been merged for the RAE 2001. Though one must be careful not to create the other kind of problem of small areas 'lost' within a much larger UoA, this is best done by employing the approach described in the next sub-paragraph.

b. Increased use of thematic sub-panels within large disciplinary UoAs. A series of thematic sub-panels have been created for the 2001 RAE, most notably under the three clinical UoAs.

c. Expanding the role of the umbrella panels that have been developed for the RAE 2001, to ensure comparability across cognate groups and to facilitate arrangements for cross-referral between panels.

d. Introducing new ‘thematic’ units of assessment, alongside the existing familiar ‘disciplinary’ units. Such an approach would mirror that used in Foresight, which has a mix of thematic and sectoral panels.

e. Establishing a regular and systematic process to review UoAs to ensure that new and emerging areas will be properly assessed.

86. As mentioned above, there has been a significant move towards a thematic assessment in the 2001 RAE (shown at Annex D). There may be benefits in adopting more widely the

approach used for medicine in the 2001 RAE, which will see a number of thematic sub-panels established beneath a main panel. For instance, there would appear to be an argument for restructuring the engineering panels, the language panels, and possibly, the biomedicine/preclinical panels, by merging them and forming thematic sub-panels beneath them.

87. Table 4 below summaries some of the overlaps of themes in Foresight, Research Council programmes and from a bibliometric analysis undertaken by HEPU, which might also serve as a basis for reviewing the UoA structure. However, we are concerned about the opportunity for 'games-playing' and the lack of consistency of assessment that might accompany the development of a broadly based thematic approach to assessment.

88. We strongly recommend that the Funding Councils continue to keep under review the number and structure of UoAs, and introduce new panels when this is justified by the changing research landscape.

Table 4: Map of some key overlapping research themes identified within the Foresight process, from Research Council strategic priorities and from bibliometric analysis

	Foresight	**PREST – Research Councils**	**HEPU - bibliometrics**
Use of materials	Manufacturing **theme**, Materials sector	EPSRC programmes	Polymers and surface science; condensed matter; materials
Instrumentation	Manufacturing **theme**	EPSRC programmes NERC programmes	Instrumentation and measurement
Society and crime	Crime prevention **theme**	ESRC governance, regulation and accountability	Criminology, social studies, law, politics
Cognition	Ageing population **theme**	MRC cognitive decline	Neuroscience and behaviour
Health and ageing	Ageing population **theme**	ESRC life-span, lifestyles and health MRC health of the public	Social, health, disability, nursing
Health and disease	Healthcare sector	BBSRC bioinformatics MRC inflammation and immunobiology NERC microbial biodiversity	Bacteriology/microbiology
Environment and energy	Energy and natural resources sector	ESRC environment and sustainability NERC programmes	Environment
Food and crops	Food chain and crops sector	BBSRC technologies for functional genomics MRC challenge of post-genome research	Cell biology, nutrition and metabolism
IT and society	Information, communications and media sector	ESRC technology and people	AI, robotics and information

The changing nature of research and the future of research assessment

89. The PREST study found a widespread belief in the sector that research within universities was going to become more focused, but undertaken in interdisciplinary and multidisciplinary teams that increasingly involved international collaborators. The specific trends most frequently cited are listed below:

- continued growth in interdisciplinary and multidisciplinary research
- increased collaboration
  - internationally
  - in the use of expensive research equipment and infrastructure
- increased 'user' involvement
- further growth in applied research
- growing importance of research teams/groups vis-à-vis the department
- emergence of more virtual research teams and the use of remote data gathering
- a further drift towards short-term research both in terms of employment contracts and the type of research work undertaken.

90. There were differing views about whether the size of research groups would become an issue of greater significance. There was a view that, in the case of biosciences, larger teams would become more important, while in other subjects such as social sciences it may become less important as IT-mediated forms of collaboration increased. It was clear that a number of interviewees from the arts and humanities believed that the model of the single scholar working on his or her own was gradually dying out, in favour of working more collaboratively with other researchers. As a consequence of increased collaboration and joint working, virtually all respondents believed that multi-authorship would continue to grow.

Alternative approaches to the current assessment approach: what opportunities exist to reduce the costs and enhance the utility of research assessment?

91. We are committed to exploring radical alternatives to the RAE, which are detailed below.

<u>Formula funding based upon research income and/or bibliometric analysis</u>

92. Formula funding is the allocation of resources in direct proportion to a series of weighted indicators of activity, output or achievement. Normally it involves establishing a given quantity of funding for research and dividing according to shares driven by the formula. To this extent the allocation of quality-related (QR) funding by HEFCE is a type of formula funding.

93. However, what is proposed is the narrower concept of using existing indicators rather than panel judgements to drive the formula. The most relevant example of its use is to be found in Australia where a component of grant funding, the research quantum, is separately calculated from that provided to support teaching and research training; 4.5 per cent of the operating grant will be provided in this way in 2000. As with QR funding in the UK, universities have considerable discretion as to how they spend this money.

94. The formula for distributing quantum funds has evolved over time. Currently the elements are research income (weighted 60 per cent), research student load (weighted 30 per cent) and research publications, including types of output not currently counted (weighted 10 per cent).

95. The quantum requires a data collection process similar to that used for the RAE, though excluding staff and strategy-related items, so its cost is similar to the RAE. Its main cost advantage lies in not having panels in operation, but this is a small proportion of the cost of the RAE.

96. It may be argued that it is more objective than panel judgements, effectively resting upon the decisions of specialised peers who referee publications and grants. In this respect it side-steps issues about the treatment of interdisciplinary and multidisciplinary research or emerging excellence (though such an assertion depends on how grant funding agencies and journals treat such work).

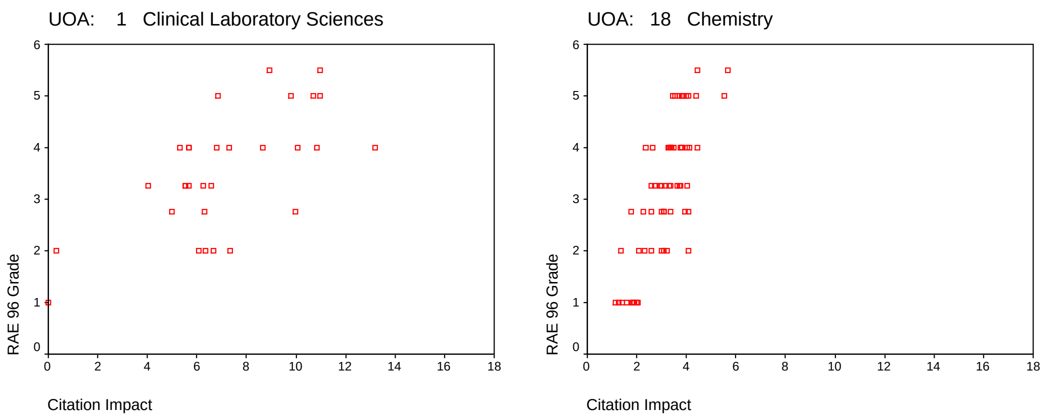
97. There is a concern that this approach rewards quantity of publication rather than quality, and that the high weighting for inputs (income) as opposed to outputs discriminates in favour of labour-intensive research (since labour costs are the main driver of grants). It is also weighted against research in the arts and humanities.

98. This approach is transparent and acknowledges that greater reliance might legitimately be placed on such indicators in some disciplines, for instance the sub-group noted that Glaxo Wellcome's response to the Call for Evidence cited a 94 per cent correlation between QR institutional allocations and Research Council funding for chemistry. However, there are

concerns that the system might become unstable if it became too self-referential. In looking at research income and bibliometric data the analysis could become more retrospective than at present, and might exclude any assessment of those elements held to be of increasing importance – strategy, research culture and research training.

99. We agree that any attempt to employ this simplistically across the spectrum of research activity would lack credibility. We were also concerned that such a mechanistic approach might reduce risk taking and the willingness of institutions to invest in emerging talent and areas. The inability of an algorithm based on one or more existing performance indicators to drive the assessment process has been demonstrated by PREST. Their study shows that, although the number of Category A staff entered and the amount of Research Council and charity income per member of research active staff in the department were, in general, statistically significantly correlated with RAE rating, in no instance was it predictive in a meaningful sense – the margin of error was at least plus or minus 1 grade. This is illustrated by Figure 4 below which plot citation impact against RAE rating. Even for chemistry – the discipline in which, uniquely, RAE ratings are very strongly correlated with other performance indicators – there is still significant variance; in clinical laboratory sciences the correlation is very poor.

Figure 5: Correlation between citation impact and the RAE rating of research units submitted to the 1996 RAE in the Clinical Laboratory Sciences and Chemistry UoAs.



Source: HEPU, University of Leeds

<u>Competitive bidding for buildings/equipment costs and/or staff costs</u>
100. The likely costs of such a bidding process would be a real concern, as would the ability of such an approach to support the appropriate level of flexibility in institutional decision making. Institutions might find their high priority bids did not find favour with the decision-making committee, while bids of lesser importance to institutions were approved.

101. JIF has proved workable as a mechanism for allocating infrastructure funding, and could be extended. However, it was recognised that it only funded capital items, not people, so an additional mechanism would be required to meet pay costs if institutions were not bidding for their entire block grants. There is also a concern about the extent to which the

provision of capital items, let alone staff, was suited to a large scale competitive peer review bidding process. This worked well for short-term investments, such as project grants, and for providing catch-up investment. However, it was unclear whether, if employed more widely, such an approach was sophisticated enough to take into account past and future performance in a way that ensured the best decisions were made for the long-term vitality of the research base.

Reputational assessment

102. At its simplest, reputational assessment can be defined as obtaining ratings of the perceived quality of an activity by consulting an informed group. Reputational assessments in an academic context are most common in North America and are usually applied to graduate programmes. Reputations are normally explored among peers (ie those providing programmes in the same field) but the concept is also sometimes extended to alumni and employers. Rankings or ratings may be sought on several criteria and there is some variability in the level of evidence which is offered to assessors.

103. The National Research Council in the USA carried out two major exercises to assess the validity of reputational assessment, one in 1982 and a second in 1995. The studies encompassed over 2,500 graduate programmes in 32 disciplines. The four measures examined by the assessment were:

- scholarly quality of programme faculty
- effectiveness of the programme in educating research scholars/scientists
- improvement of programme quality in the past five years,
- evaluators' familiarity with the work of the programme faculty.

104. At the outset five main criticisms of reputational rankings were recognised:

a. Such studies reflect perceptions that may be several years out of date and do not take into account recent changes.
b. The ratings of individual programmes are likely to be influenced by the overall reputation of the university – a 'halo effect'.c. A disproportionately large fraction of the evaluators are graduates or faculty members of the largest programmes, which may bias survey results.
d. It may not be possible to differentiate among many of the lesser known programmes in which relatively few of the faculty members have established reputations in research.
e. The process reinforces a narrow model of graduate education.

105. The emerging judgements were regarded by the authors as reflecting a strong consensus and correlating well with other measures (though reputational assessment could be seen as a holistic or summary variable). To this extent reputational assessment represents a large-scale broadening of the peer review process, though this is at the

expense of the depth of analysis and comparability of approach which may be achieved within a panel.

106. However, two key differences from the RAE are relevant. Firstly, reputational assessment does not exist to allocate resources, though its consequences almost certainly have major resource implications for those concerned. If the RAE function of very direct resource allocation was attempted through reputational assessment there would be a significant opportunity for game-playing and collusion, with much less chance of detecting it than in the present system. The second key difference is that reputational assessment is largely confined to graduate education. In this sense it represents a summary of a process which takes place naturally, in that academic staff will advise their graduating students as to which schools will offer them the best graduate education. There is not the same underpinning process in research, where quality judgements require a higher level of expertise and specialisation.

107. It would therefore seem to us that reputational assessment is not a viable alternative assessment mechanism to drive funding, as it is open to manipulation and is subject to considerable potential distortions from geography, established relationships, and the public profile of researchers or departments. However, it did have some merits in widening the assessment base, and may be useful input to a funding mechanism by providing more contextual information.

<u>Maintain present dual support system but increase overheads payable on project funding: transfer of funds to the Research Councils</u>

108. The sub-group noted that a dual support transfer of the type that had occurred in 1992 could enable full overheads on Research Council projects to be paid.

109. However, such a change would still require a mechanism for the remaining funds to be distributed and for the provision of infrastructure support to charitably and industrially funded research, both of which are growing in significance. It was also clear that although the 1992 transfer had been accompanied by policy statements that the transferred funding would not be used to finance growth in the volume of research, this had in fact occurred. One of the key issues for policy makers was to establish a better project-infrastucture balance, and past experience suggested that a transfer of this type might exacerbate the problem rather than help to solve it.

110. A dual support transfer would also see a further move to defining and directing the research activity that HEIs undertake. While it was acknowledged that there was real benefit in the Research Councils providing earmarked funding in strategically important areas, HEIs had a good track record in identifying and participating in research in areas of emerging economic and social significance. QR funding played a vital role in enabling HEIs to identify and develop these activities before they were identified as needing support by project funders.

Combining teaching and research assessment

111. Such an approach generates intuitive support, as it is a holistic approach to assessing the activities within HEIs. However, TQA alone is a mammoth task, which is one reason why it is done on a rolling basis (see below). Linking such an exercise with the assessment of research would probably be unmanageable, from the point of view of both the assessors and the assessed. The RAE is a relatively well tolerated and cost-effective exercise, but this support could be rapidly eroded if it were linked to the TQA. We were also concerned about whether such an approach would be able to discriminate between good and poor quality submissions as effectively as the RAE.

Rolling assessment

112. Although the RAE makes significant management demands of HEIs every four or five years, as it is a 'snap-shot' exercise the arrangements are not permanent. We believe that if the RAE became a rolling assessment process, dedicated staff would need to be recruited and specific management information systems developed, all of which would increase rather than decrease the management burden. We are also inclined to the view that never being free from a large scale assessment of teaching or research would be unacceptable to institutions.

Highly-rated departments only participating in every other RAE

113. The RAE has demonstrably improved the research performance of institutions. However, there is an argument that the system has reached a level of stability at which departments that have demonstrated their ability to maintain quality at the highest levels over a period of time should be provided with longer planing and funding horizons.

114. We therefore considered an approach in which those who have achieved the highest ratings in successive RAEs are not required to participate in a subsequent cycle. This would focus the RAE on aspirational departments, and it would reduce the burden on those institutions who are performing at the highest levels.

115. We modelled such an approach and it appearsto be financial stable (it does not distort the allocations calculated in respect of institutions) and, based on assumptions relating to the cost of the 1996 exercise, it could reduce the cost of the RAE to institutions by 40 per cent. Also, providing longer planning and funding horizons could be expected to encourage researchers to engage in riskier, longer-term and more innovative research.

•

116. However, we are concerned about how such an approach could be put into practice in an even-handed way. 'Exempt' departments would need to have their funding allocations uplifted by a certain amount (by inflation and/or HEFCE QR changes and/or some other basis) to ensure their funding was not diluted by not participating in the RAE, or subjected to uncertainties due to the RAE outcomes of other departments or changes in selectivity. But,

safeguarding their interests in this way could be considered to disadvantage others, because they would be the subject of the vagaries of open competition.

117. Such a system would also raise serious concerns about extending for five years the funding of under-performing departments previously rated 5/5*, which would be at the expense of developing research excellence in both pre- and post-1992 universities. In addition, there would be possibilities for games-playing under such an approach, where a whole institution was not assessed. In particular, institutions might 'decant' staff into 5 and 5* rated departments (which were exempted from participation in the following RAE) to increase the recorded volume, and/or move ‘research stars’ out of exempted departments into 4-rated departments to improve their ratings.

118. There are measures that could be employed to identify or discourage such behaviour including:

- charting the movement of named individuals within institutions using HESA records
- obliging exempted departments to provide an RA2 return to the RAE
- auditing a sample of institutions
- requiring 'exempted' departments to make a 'supplementary submission' to contain only the details for any staff who have produced an increase in volume; this would be assessed by the panels, and would obtain a rating, ranging from 1 to 5*.

119. However, we believe such an approach is flawed and not even-handed in its treatment of different research units when the current RAE is predicated on open competition on a consistent and transparent basis.

120. We also considered whether 1 and 2 rated departments should be excluded from every other RAE, as the development track of such departments was long and therefore, as with highly performing units, the cost may be considered to be disproportionate to the benefit. However, we have taken the view that institutions must retain the right to decide for themselves whether participation in the RAE is in their interests, accepting that there are perfectly valid reasons for institutions to submit RAE returns in UoAs where they have little prospect of funding. These include:

- to clarify the institution’s strengths and weaknesses
- as an index of improvement or decline
- as a stimulus to strategic thinking within departments
- as an incentive for research active staff to maximise the quality of their outputs
- to establish to existing and potential staff that the department has a focus upon research and thereby to recruit and retain higher quality people.
- to benchmark research activity against wider standards.

Centres of excellence

121. Although we are disinclined to recommend exempting high (or) low performing units from the RAE, we do feel that there are some benefits to additional recognition of research activity that evidences quality in both depth and breadth (ie wider than in a single UoA) by recognising such as a 'Centre of Excellence'. For example, being recognised as a centre of excellence might require as a necessary but not sufficient condition, substantial links between high performing units in related UoAs.

122. This would not necessarily imply an increase in funding for those departments so recognised, but would provide a basis to focus support for world class research if appropriate. This might be on condition that they engage in additional activities to collaborate with others to share excellence, expertise and equipment – for example, with users, other academic units in the UK or overseas (within or across disciplines) – and possibly that they engage in activities to increase public understanding of science.

123. This would mirror similar initiatives in a number of countries (including Canada, the Netherlands, Japan, Australia, Belgium, the USA, Finland) which seek to identify and support centres of excellence.

124. We noted, in particular, the Finnish system in which centres of excellence are identified by international peer review. Themain criteria are:

- the national and international position of researchers
- the scientific significance, innovativeness and effectiveness of research
- quality, quantity and focus of scientific production
- patents
- national and international mobility of researchers
- number and level of foreign researchers in the centres.

125. We recommend that elements of the Finnish approach be adopted in the UK, though we recognise that the criteria identified above are heavily skewed towards sciences.

126. A more general approach, reflecting the differences between disciplines, would need to provide much greater flexibility in the performance measures employed for each discipline, depending on the relevance of, or particular nature of, the different dimensions of excellence:

- production of knowledge at the leading edge
- value to, and impact upon, users and society
- training and sustainability.

127. Such an approach would also accord with the key features of the European centres of excellence model that was discussed by the EU Commission at the Lisbon summit in March 2000. EU research ministers agreed that there was a need to facilitate the development of European centres of excellence, and that these would be identified as centres with.

- a ‘critical mass' of high level scientists and researchers
- a well-identified structure
- capability to integrate connected fields and to associate complementary skills
- capability to maintain a high rate of exchange of qualified human resources
- a dynamic role in the surrounding innovation system (adding value to knowledge)
- high levels of international visibility and scientific and/or industrial connectivity
- reasonable stability of funding and operating conditions over time (the basis for investing in people and building partnerships)
- sources of finance which are not dependent over time on public funding.

Other alternative approaches

126. Pros and cons of number of alternative approaches were explored, summaries are detailed at Annex E:

- Royal Academy of Engineering approach
- fund self-defined research networks on a competitive bidding basis
- freezing allocations after the 2001 exercise
- self-assessment
- assessment based on research productivity
- assessment based on economic impact.

127. We are inclined to the view that the other alternatives to the current assessment approach, as they are currently conceived, are unable to form a robust basis for funding allocations. However, we recommend that the assessment process embrace the principles of the recent Royal Academy of Engineering report, which recognised five dimensions of excellence. The emphasis on increasing the recognition of applicable research, strategy, and vitality and sustainability as dimensions of research excellence (discussed at Annex F) clearly resonate with our own thinking.

What alternative indicators could be employed

128. It is clear that as the research landscape changes, and in particular as new research areas emerge, the diversity of this activity needs to be properly assessed within the RAE.

129. We therefore explored what alternative indicators have been employed or could be employed to capture the diversity of excellence. We first commissioned a review of the

additional indicators cited in forms RA5 and RA6 from the 1996 RAE. This clearly demonstrated that many institutions were putting forward patents, innovations and spin-off developments, consultancy, and industry links as indicators of quality. In addition, a large number of references were made to indicators such as:

- **awards, prizes and fellowships.** Numerous returns in all subjects described a range of awards, prizes and fellowships but did so in a relatively unstructured way

- **journal editorships and editorial board membership.** This was another commonly described attribute of esteem found across all disciplines

- **visiting positions elsewhere.** Another commonly cited indictor across all UoAs, but particularly strongly in chemistry, sociology, management, and history

- **Professional body activities.** These were widely mentioned in almost all returns.

130. It is clear from this analysis that there are many indicators other than bibliometric and income measures that are currently being considered as part of the assessment process. However, we believe that incorporating these indicators into the assessment process in a more structured way would have considerable benefits.

131. We commissioned an analysis of responses to the question in the recently issued Call for Evidence that asked: 'What measures should be used to assess the performance of the research base?' and looked for a new indicator. The list of relevant responses is shown at Annex G. This yielded some interesting suggestions.

132. We recommend that as part of its consultation exercise through the summer the Funding Councils test the enthusiasm of the researchers and stakeholders for some of the measures suggested in responses to the Call for Evidence and in the Royal Academy of Engineering report.

How can research assessment better engage with research users?

133. The Funding Councils, jointly with the CBI and the DTI, have sought to more fully engage user representatives in the 2001 RAE. The guidance to panel members places great emphasis on the need to develop structures and processes that will ensure the appropriate assessment of user relevant research. As a result the user participation rate has tripled from that in the 1996 RAE, but the level of engagement with different industrial sectors was variable and there has been particular difficulty in engaging withsmall and medium enterprises (SMEs). These problems arise because of the time commitment involved rather than because of any unwillingness to participate on the part of business, or resistance from the panels to their inclusion.

134. In focusing upon the need for user representation on main panels it is easy to gloss over the magnitude of the time commitment involved. We are convinced that this is the problem where there are insufficient users coming forward. Only large companies which employ HE liaison officers may be able to consider releasing staff to be full panel members.

135. Where there are such constraints, user sub-panels play a vital role in enabling the peer review process to be widened, and, as is clear from Annex D, the 2001 RAE will see more use of user sub-panels than in any previous RAE.

136. However, we fully agree with the HEFCE/CBI Task Group on user involvement in the RAE that user sub-panels should not be seen as a substitute for user representation on full panels; we emphasise that they perform a separate and valuable function. For future exercises, we recommend that the emphasis on recruiting users to full panel membership is retained, but that flexible arrangements are in place reduce the burden on all user members, and to enable those willing to make a more limited commitment to participate. There is no reason why such flexibility should apply only to those considering submissions: there may be room for user representatives to take a more active role in designing panel criteria and procedures without taking on the full responsibility for reading and reviewing submissions.

137. We recognised that, although business has some generic concerns, the needs of large scale corporate research funders and users were not necessarily the same as those of SMEs: research training issues were likely to be the same, but research 'output' issues might well be different. Also, SMEs are likely to be less able to release staff than large industrial companies. We therefore recommend that the Funding Councils consider whether it is necessary develop specific mechanisms to enable SMEs, particularly technologically focused small businesses, to engage directly with the assessment process.

Is the cost of the RAE disproportionate to the changes in institutional and departmental funding it produces?

138. The RAE is the largest single research assessment exercise in the world involving over 55,000 academics, 3,000 departments and 191 institutions. An exercise of this magnitude inevitably raises cost benefit issues.

139. It has been suggested that the cost of the RAE is disproportionate to the benefit as it does not bring about significant changes in funding. Although this is true at the sectoral level (ie the aggregate funding for pre-1992 universities, post-1992 universities, and HE colleges) there is significant movement at the institutional level. Analysis commissioned as part of the review has demonstrated that percentage changes in funding at the institutional level can be large.

140. The study by HEPU has shown that institutional resource allocation models generally follow the HEFCE formula. We have therefore repeated this analysis at the level of individual submissions (that is, at the level of allocation to individual subjects in each institution) to get

an indication of the level of funding movement brought about by the RAE at departmental level.

141. We have excluded capital funding, and again looked at the effect on pre- and post-1996 RAE allocations at the UoA level for each HEI of:

- QR volume
- RAE ratings relative to sector
- rate of funding
- QR allocations.

142. The overall 'funding churn' following the 1996 RAE was some 55 per centof the funding available. The effect of each of the individual factors is shown below:

- funds attributable to change in rating only = £114 million (19 per cent of 1996-97 baseline)
- funds attributable to volume changes only = £112 million (19 per cent of 1996-97 baseline)
- funds attributable to change in pot size only = £62 million (10 per cent of 1996-97 baseline)
- funds attributable only to submissions no longer funded after the 1996 RAE = £31 million (5 per cent of the 1996-97 baseline).

143. Clearly some of the changes will be in opposite directions for a given submission. However, the net effect is still large, with 30 per cent of research funding moving. Even when one removes the effect of factors associated with, but not the result of, the RAE, the 20 per cent movement in total funding is still large, plus there is the 5 per cent change attributable to submissions that were no longer funded.

Changes in ratings associated with successive RAE cycles

144. The changes in funding identified above are clearly associated with a changing and dynamic sector. To understand this better we have looked at the movement of ratings associated with successive RAEs. Tables 7 and 8 show considerable changes in ratings associated with the 1992 and 1996 RAEs. Although it has often been argued that increases in the ratings reflect better presentation of research activity rather than real improvement, the HEPU study for the selectivity sub-group demonstrated that changes in RAE ratings are associated with real changes in the quality of the research to which submissions relate.

Table 5: RAE grade movement from 1989 to 1992

1992 Submissions	1992 Rating						
1989 Rating	**0**	**1**	**2**	**3**	**4**	**5**	**Total**
0		90	80	106	47	36	359
1	41	13	45	49	10	1	159
2	37	5	107	189	58	8	404
3	31		46	284	176	48	585
4	10		1	72	181	86	350
5	8		1	6	44	143	202
Grand total	127	108	280	706	516	322	2059

Improved rating in 1992	1029	50.0%
Same rating in 1992	728	35.4%
Decreased rating in 1992	175	8.5%
Dropped out after 1989	127	6.2%

Table 6: RAE grade movement from 1992 to 1996

1996 Submissions	1996 Rating							
1992 Rating	**0**	**1**	**2**	**3**	**4**	**5**	**5.5**	**Total**
0		126	207	131	30	14	7	515
1	180	78	84	60	2		1	405
2	87	28	130	290	44	4		583
3	64	2	36	370	271	54	5	802
4	13			79	254	162	22	530
5	6			4	43	150	120	323
Grand Total	350	234	457	934	644	384	155	**3158**

Improved rating in 1996	1634	51.7%
Same rating in 1996	982	31.1%
Decreased rating in 1996	192	6.1%
Dropped out after 1992	350	11.1%

What are the true costs of the RAE and are they justified by the benefits it brings?

145. The cost of the RAE is not small. The 1996 exercise was estimated to have cost institutions £27.3 million, in staff time and material costs; this value was calculated by aggregating individual returns from institutions. A more in-depth analysis, including opportunity costs, has been conducted in one institution. Extrapolation of these findings to the whole sector suggests an upper limit for the cost of the RAE of £37.5 million. However, this still represents only 0.8 per cent of the total funds whose allocation it will directly determine. Notwithstanding the above, it is debatable whether the RAEis expensive: although the stated cost is large, it represents less than 1 per cent of the funds to be allocated, which compares very favourably with other evaluation mechanisms.

146. Indications from the PREST studies suggest that the RAE is considered by staff to be much less burdensome than teaching assessment, and does not make undue demands.

147. We believe that for the most part institutions should be developing management information systems for other purposes, on which they can draw to provide information to research assessment exercises more cost-effectively in future.

148. Although the RAE is presently cost effective, it would become less so if the amount of funding to be disbursed is diminished, or if reallocation between exercises tends to zero.

Conclusions

<u>Supporting world class research</u>

149. We believe that research activity should be assessed in a way that interferes as little as possible with the way in which researchers naturally work, so that quality is promoted but activity is not distorted. Alternatives to the RAE, or a modified RAE, must be able to deal effectively with the growth in interdisciplinary and multidisciplinary research, the development of new forms of research output, and assessment of future potential and strategy.

150. We recommend moving to a part thematic assessment framework in order better to support high quality research.

151. Discussion of the possible alternative assessment frameworks listed in Annex E suggests that the likely distortions outweigh any potential benefits of any of the proposed alternatives. However, we recommend some significant changes to the assessment process.

Recognising the characteristics of excellence

152. We recommend that the assessment process should more explicitly recognise the different characteristics of excellence such as:

- production of knowledge at the leading edge
- value to, and impact on, users and society
- training
- sustainability.

153. The 2001 RAE will see a move in this direction, with the increased emphasis on the assessment of research culture and strategy.

154. However, we believe it is desirable to further formalise and standardise the assessment of these strategic and cultural aspects of excellence, and that the RAE is sufficiently flexible for such a move to be accommodated.

155. This should not be interpreted as a belief that innovation can be planned: the HEPU study found that although the RAE has improved management of the research environment, research itself has remained comparatively ‘unmanaged’. Our suggestion is that the assessment process should review whether institutions and departments have institutional management systems in place to recognise, support and develop the research effort. Such an approach would safeguard institutional autonomy by demonstrating the ability of institutions with excellent research to sustain and develop their various, self-defined, priority areas and their general research effort.

156. We also recommend that the Funding Councils should ensure that panels share the increasing recognition that excellence is often inseparable from applicability. In particular, we note the findings of the HEPU study commissioned as part of the review that shows close correlation in the UK between industrial and public funding of research, that industrial income per head generally increases as public funding rises, and that there is in most disciplines a positive correlation between bibliometric impact and industrial funding. There is, therefore, no evidence of a separation between excellence and innovation.

Recognising disciplinary differences

157. We recognise that in establishing umbrella panels the Funding Councils have developed a mechanism that allows the characteristics of excellence in particular disciplines to be assessed appropriately, while ensuring comparability across cognate areas. We recommend that further moves to recognise differences between disciplines would create greater flexibility in the assessment process, to the benefit of the exercise and the institutions being assessed.

Recognising institutional missions

158. We concluded that it was entirely appropriate for the RAE to continue to prompt those research-led institutions capable of maintaining broadly-based world class research to do so. It was also appropriate for assessment mechanisms to encourage teaching-led institutions to focus their research effort in areas where they can compete effectively.

<u>Encouraging applicable research</u>

Clarity in responsibilities

159. We noted the importance of clarity in the dual support arrangements (and welcome the fact that the HEFCE is establishing a joint group with OST to look specifically at this). We also concluded that there should be a clearer distinction between the purposes of QR and HEROBC funding, with assessment mechanisms appropriate to the purposes of each. There is a need for funding that supports local and regional needs and developments. If such

support is not to be provided by HEROBC, further thought needs to be given to how it can be delivered.

160. However, we believe that it is essential that applicable research should be funded through QR: separating out the funding stream would make applicable research seem to be of lower merit and quality. This might lead to the creation of a two-tier system where some institutions would only do teaching, topped-up with monies provided for applicable, third-leg research. We believe that the Funding Councils should ensure that the RAE appropriately assesses applicable research.

Increased engagement with users

161. We noted the concern that applicable research is undervalued in the current assessment system and these views have been repeated by active researchers in their interviews with PREST. PREST found that there were mixed views about the treatment of applicable research, and some panels were perceived as being good and fair in treating applied research, the Engineering Panel being cited as one exemplar of good practice. There was an acceptance that changes in the regulations for RAE 2001 recognised that applicable research could be as good quality in terms of research excellence as non-applicable research. However, we concluded that the response to the Call for Evidence indicates that many researchers and institutional managers have yet to appreciate the significance of the changes in guidelines and working methods for panels for the 2001 exercise. In summary, these are:

- more user representation
- explicit statements that research would not be differentiated by type or purpose
- all forms of research output can be submitted
- statements of neutrality in working methods, ie, that no hierarchy in research inputs or outputs would be assumed
- procedures to enable the assessment of confidential research undertaken in collaboration with commercial partners.

Eliminating perceptions of prejudice

162. There was some scepticism that the panels would act in accordance with their statements on working methods, and concern that the assessment of confidential work would require the release of these papers for the RAE from the industrial collaborator, which could result in complicated negotiations. We recognise the particular concerns of researchers in vocational areas such as medicine and architecture who feel that their research is not valued as highly as research in other areas.

163. We recommend that the Funding Councils put in place mechanisms to demonstrate to the research community that the changes for the 2001 RAE will result in applicable research being properly assessed. We also recommend that after the exercise is complete they establish a joint group with the CBI and other stakeholders to review whether, based on the

2001 experience and the planned changes arising out of the recommendations in this report, yet further changes are necessary for any subsequent exercise.

Recognising/supporting collaboration and networks

164. It is generally acknowledged that the networks of UK researchers, both nationally and internationally, are extremely well developed, and support considerable collaborative research activity. We wish to recognise and support the development of research networks and collaboration, but are concerned that if a specific funding stream was developed that explicitly sought to foster the development of these links, it might have unintended consequences that outweighed the benefits.

165. We therefore recommend that the Funding Councils actively discuss with institutional managers the benefits of making collaborative submissions, and the working of the arrangements that have been developed for the 2001 exercise.

166. Collaboration can be a characteristics of research excellence. In addition, we therefore recommend that, where appropriate, panels give credit to the formal and informal collaborative links between researchers in HE, and between those in HE and industry, government, and other organisations engaged in research.

167. We are not inclined to suggest developing a specific funding scheme to support interdisciplinary and multidisciplinary research. Rather we feel that the leaders of institutions are best placed to extract synergy out of the various sources of income they receive, and to make strategic investments for the future, so should have considerable discretion over how they apply it., This would in many cases include providing funding to back the work of interdisciplinary or multidisciplinary units. Although much value could be obtained from multidisciplinary working, such research varies greatly in quality; one of the problems associated with attempts to promote this work directly through a centrally funded initiative would be the likely distortions produced.

168. Instead, we recommend establishing structures to ensure that interdisciplinary and multidisciplinary research are properly assessed. This should include greater use of thematic sub-panels, where appropriate, and continuing to keep under review the number and structure of UoAs, introducing new panels when this is justified by the changing research landscape.

Development of research people and production of PhDs

169. One of the principal outputs of the research base is a supply of trained researchers. These individuals may go on to pursue academic careers, but increasingly they are required by society generally, and industry in particular. However, the supply and development of such individuals, including postgraduate and postdoctoral researchers, is not currently addressed either by the current RAE or by TQA.

Explicitness and encouragement

170. RAE panels do consider factors related to postgraduate training as part of the quality assessment process. This has acted as an incentive for institutions to invest financial and management resources in research training, since universities need to develop their research potential. Panels have been encouraged to place more emphasis on assessing the research culture in the 2001 exercise. However, we are concerned that without changes to the RAE, panels might place disproportionate emphasis on the assessment of the quality of research outputs and simply expect that the contribution to the volume measure from postgraduate trainees will provide sufficient reward to institutions.

171. Research training is a key indicator of sustainability and therefore warrants more explicit recognition in the assessment process. We therefore recommend that research training and career development be the subject of a separate, but linked, assessment process. In respect of postgraduate researchers the panels might be able to draw on information held by the Research Councils. In considering this change we recommend that any proposed new approach does not provide an incentive to grow volume at the expense of infrastructure.

Appropriate recognition

172. HESA data for 1996-97 suggests that some 34 per cent of all academic staff are on fixed-term contracts of employment. More specifically, some 17 per cent of research active staff are on fixed-term employment contracts. It can be argued that currently the contribution of these individuals, who form a flexible but highly committed part of the academic workforce is not adequately recognised in the assessment process.

173. The Research Careers Initiative (RCI) has developed metrics to assess the quality of training and development, and found that the Research Councils where supportive in principle of its objectives.

174. We recommend that the separate assessment of research training and career development we have proposed should encompass contract researchers, in order that the contribution of this group is properly appreciated.

175. However, increased support for career development, involving departments developing explicit staff development strategies, might have unintended consequences. For instance, if institutions were obliged to employ people on a longer-term basis it might simply reduce their willingness to offer temporary positions to academic staff, and therefore the opportunities that these positions provided. This might lead to many good researchers leaving the profession. We therefore suggest that the implications of any proposed changes are considered in some detail, and implemented thoughtfully.

Recognising diversity

176. We noted PREST's finding that changes in the RAE procedures for the 2001 exercise were seen as positive in promoting opportunities for women and young researchers. However, we recommend that, where it allows a more appropriate assessment of the quality of research in which an individual has been engaged, they should be able to include a personal statement within the unit's RAE submission.

177. This would include individuals who have had a career break, but also those who are engaged in long-term research, have recently embarked on an academic career and have not developed a research portfolio, or have undertaken research that was of high quality, but was risky and did not prove to be fruitful in terms of publications.

Widening research

The response of post-1992 universities to the Call for Evidence indicated some dissatisfaction with the assessment of emerging excellence by the RAE. Although we interpreted this as dissatisfaction with funding rather than assessment, we were committed to exploring the best approach to assessing emerging disciplines. Greater recognition of the differences between disciplines would permit greater flexibility in developing assessment mechanisms appropriate for emerging areas. This was logically prior to, but linked with, developing thematic units of assessment, increasing the importance of the people dimension in the assessment of research, and increasing emphasis on research strategy.

Annex A

Membership of the sub-group

[PUT THESE IN A WORD TABLE, AS THIS CAUSES FAR FEWER PROBLEMS THAN TABS OR COLUMNS]

Professor Sir Gareth Roberts (Chair)		Vice Chancellor, University of Sheffield
Professor Peter Lachmann		Chairman, National Academies Advisory Group
Professor Julia Higgins		Dept of Chemical Engineering, Imperial College of Science, Technology and Medicine
Professor Stephen Newstead		Department of Psychology, University of Plymouth
Professor John Laver		Professor of Phonetics, The University of Edinburgh
Dr Colin Lucas		Vice-Chancellor, University of Oxford
Dr David Smith		Chief Executive, Whatman plc
Professor Phil Ruffles		Director of Technology, Rolls-Royce plc
Mr Bahram Bekhradnia		Director of Policy, HEFCE
Dr David Pilsbury		Head of Research Policy, HEFCE

Annex B

Terms of Reference

To consider the role quality assurance and evaluation mechanisms might play in underpinning the allocation of HEFCE research funding.

In considering this issue the group will wish to address the following questions:

<u>How has the research assessment function developed in the UK?</u>

- What is research assessment for and to what purposes has it been put?
- What can be learnt from the history of research assessment and successive RAE cycles?
- What has proved workable and what has not?

<u>What have the implications of research assessment been for the careers of researchers?</u>

<u>What have been the effects of research assessment on the research base?</u>

- Has research assessment provided quality assurance?
- Has research assessment encouraged an improvement in research quality?
- Has research assessment encouraged an improvement in research management?

<u>What approaches are other countries adopting to research assessment?</u>

- What are the benefits and problems associated with the different approaches to research assessment in comparator countries?
- How do comparator countries view the current UK approach to research assessment?

<u>Is research assessment necessary to deliver the research funding policies of the HEFCE?</u>

<u>How should the research assessment process change to ensure the health of the research base in the future?</u>

- Has research assessment run its course and fulfilled its original purposes?
- Should research assessment continue to be based on peer review?
- Should research assessment be more integrated with other assessment process – for example teaching quality assessment?
- What form of evaluation is most appropriate as the nature of research changes?
- How can research assessment better engage with research users?
- How might research assessment respond to a change in the nature and purpose of HEFCE funding?
- Should the assessment process differentiate between funders, the type of funding, and the type of activity?

- How can assessment best balance the need for consistency and comparability across disciplines, given that some disciplines have significantly different structures to others?

What criteria would best address the future funding needs of the research base?

- What alternative indicators could be usefully employed?
  - survey-based reputational assessment?
  - student-based assessment?
  - self assessment?
  - assessment based on commercial applicability of outputs?
  - assessment of research productivity?
- How should the assessment criteria be weighted?
- Should research assessment explicitly acknowledge strategically important research such as that related to Foresight priorities or regional needs?
- What is the optimal balance between quantitative and qualitative indicators?
- What is the right balance between retrospective and prospective assessment?

Are the costs of research assessment justified by the benefits?

- What is the true cost of the RAE?
- Is the cost disproportionate to the changes in institutional and departmental funding it produces?
- Is the cost disproportionate to the benefits it brings?
- What opportunities exist to reduce the costs?
- What opportunities exist to enhance the utility?

Annex C

Research evaluation practices in other countries

No	Country	University research evaluation for allocation of core research Fund	Other national evaluation of university research
1	**UK**	Ex-post informed peer review (RAE)	
2	**Hong Kong**	Ex-post informed peer review (RAE)	
3	**Australia**	Ex-post quantitative evaluation based on primary performance indicators	
4	**Poland**	Peer review + ex-post informed peer evaluation	
5	**The Slovak Republic**	Student numbers moderated by ex-post informed peer review	
6	**Denmark**	Student numbers/ Performance based	Evaluation of specific research areas
7	**Finland**	Student numbers / Ex-post quantitative evaluation	Evaluation of specific research areas
8	**New Zealand**	Staff and student numbers *	
9	**Germany**	Staff and student numbers	
10	**Italy**	Staff and student numbers	
11	**Norway**	Staff and student numbers	Evaluation of specific research areas
12	**Sweden**	Staff and student numbers	Evaluation of specific research areas
13	**Hungary**	Staff and student numbers	
14	**Austria**	Negotiation	Evaluation of specific research areas
15	**France**	Negotiation	CNE & CNRS Evaluation
16	**The Netherlands**	Staff and student numbers	VSNU research evaluation
17	**USA**	Staff and student numbers	Graduate-Research evaluation
18	**Canada**	Staff and student numbers	

What are the benefits and problems associated with the different approaches to research assessment in comparator countries?

Assessment based on existing performance indicators

In determining the allocation of funds to universities, Australia makes use of the research quantum which is based on input data (grants and contracts) and output indicators (publication counts and higher degrees). Such an evaluation system is cheaper to operate than the UK Research Assessment Exercise but is distorted by the inherent differences in research costs across and within research areas. Itdoes not work for funding the arts and humanities. There are also concerns about the extent to which such an approach supports diversity and innovation, as it tends to simply reinforce the decisions of grant-giving organisations, and citations reflect historic activity.

Funding based on institutional size

In several of the countries analysed as part of the report, research funding is allocated to universities as part of the general institutional funding which is based on size (numbers of students and staff). In such countries the universities also received research funds for specific projects from research council-type funders.

In such a system, no account taken of research performance and this is held to undermine innovation and to be responsible for a decline in competitiveness. For instance, in Germany there has been no comprehensive assessment of university research, although the state of Lower Saxony recently started to assess certain fields in its universities, and other Länder governments are beginning to allocate a proportion of research funding on a performance-related basis. In addition, the Freie University in Berlin has implemented its own evaluation system, the results of which are used in determining the internal distribution of funds.

In Norway and Sweden, universities receive a block grant based on their size with no attempt made to assess research performance. The same has been true for Italy but the National University Evaluation Council has recently proposed that funding be linked to, among other things, research results.

Small proportion of research funds allocated on basis of research performance

In Finland, universities receive most of their general research support as part of their institutional funding which is determined by university size. However, a small proportion of this funding is based on their teaching and research performance (eg numbers of PhD degrees). Likewise, in Denmark, since 1995 a small element of university funding has been linked to levels of external research income and numbers of PhDs produced. Universities in New Zealand have previously received research funding through the general institutional support (based on student numbers), but in 2000 there will be a switch to contested funding allocated on the basis of *ex ante* evaluation (carried out by peer review).

The benefits seen to derive from an evaluated approach to funding are likely to result in the proportion of research funds allocated to universities in this way increasing over time.

Funding based on negotiation with Government

In Austria and France university budgets are determined primarily through a process of negotiation with the relevant ministry. In the case of Austria, there is no strict relation between university size and funding in arriving at decisions on funding, although this is taken into account in the negotiation process. The lack of any transparent basis in decision making is proving problematic as the institutions face increasing local and global competition. However, no research assessment is involved as the academic community is strongly opposed to such assessments, and there has been no strong push in this direction from ministers. Evaluations of two fields have been carried out but the results had no formal consequences.

In France, the negotiation of university funding levels involves not only the Ministry of Education , Research and Technology and the universities but also the main research council, CNRS. Unlike Austria, however, in France these negotiations take into account the evaluations carried out of individual universities by the National Evaluation Committee. These evaluations focus on teaching and management as well as research, and they are based on peer review along with various indicators.

Research assessment not linked to funding

There is extensive evaluation of university research in The Netherlands but it is not used to inform funding decisions. Instead, the aim of such evaluation is to develop research strategies and to be used in decision-making processes. It is based on informed peer review carried out by panels (as in the UK) but with greater reliance on foreign peers; there are also some site visits. Further differences compared with the UK are that different fields are evaluated at different times and the evaluations focus on four aspects of research performance – scientific quality, scientific productivity, scientific relevance and long-term viability.

In the United States, the majority of explicit research support is obtained in the form of specific research grants and contracts, though there is also indirect support from tuition fees and endowment funding. There are periodic evaluations of doctoral programmes in universities, though these seek to assess the research standing of departments as well as the quality of their doctoral training. The approach is based on peer review but in this case it takes the form of an extensive questionnaire survey rather than a small panel (as in the UK and The Netherlands, for example). Use is also made of various performance indicators. The results are employed by doctoral students (and perhaps by academic staff) in choosing which department to join, and by universities in internal management and strategy development.

Annex D

List of sub-panels in the 2001 RAE

Panel description	**Sub-panel Names**
Clinical laboratory subjects ;Hospital-based clinical subjects	Neurosciences
	Cardiovascular Medicine
	Cancer Studies
	Infection and Immunology
Community based clinical subjects	Epidemiology, Public Health Research and Health Services Research
	Primary Care
	Psychiatry
Pre-clinical studies Anatomy Physiology & Pharmocology (constituted as a single panel)	Pre-Clinical Medicine
	Anatomy
	Physiology
	Pharmacology
Psychology	Users' Sub-panel
Agriculture, food science and technology and veterinary science (constituted as a single panel)	Agriculture
	Food Science and Technology
	Veterinary Science
Applied mathematics	Users' Sub-panel
Statistics and operational research	Users' Sub-panel
Computer science	Information Systems (See UoA 61)
Geography	Users' Sub-panel
	Development Studies
Law	Scottish Law
Politics and international studies	Users' Sub-panel
	Science Policy
Sociology	Users' Sub-panel
Sociology	Women's Studies
Middle Eastern and African studies	African Studies
	Arabic, Persian, Turkish, Middle Eastern, and Islamic Studies
	Hebrew, Jewish, and Ancient Middle Eastern Studies
Linguistics	Phonetics
Library and information management	Information Systems (See UoA 25)
Communication, cultural and media studies	Practice-Based Research
Education	Continuing Education

Annex E

Alternative assessment mechanisms and other potential changes, and their strengths and weaknesses

Alternative	**Strengths**	**Weaknesses**
Use of formula based on research income and/or publications citations	• Potentially cheap to administer • Transparent • Viability shown in Australia • Grant income based on mix of prospective and non-specific retrospective (track record) judgement	• Requires complex validation/checking which may raise cost • Reduces plurality through dependence upon Research Council decisions • Works best in Science, probably not at all in Arts and Humanities • Citations are USA-centric and may divert research from UK interests • Citations information may not relate to current activities • May distort activity as researchers seek to boost performance in chosen indicators
Competitive bidding for block funding (c.f. JIF model)	• Very selective • Accountability built in to funding contract • Model exists in funding for JIF and commercialised national labs (eg [SPELL OUT ACRONYM]NPL) • Judgement based on mix of prospective and non-specific retrospective (track record) judgement	• High transaction costs in preparing and assessing proposals • Can cause wide fluctuations and instability in institutional income • Contracts may limit scope for innovation and academic freedom
Reputational assessment	• Broadens base of peer review to whole of field • Relatively cheap to administer • Feasible in any subject	• Highly susceptible to prejudice as no specific examination of evidence involved • Open to game playing and collusion • Nominations only likely for most visible so little discrimination lower down table • Reputation may persist long after its causes have gone away • No instance of use in significant resource allocation
Freeze on present ratings	• Minimum cost option	• No incentive to improve for new entrants or revamped units • Could induce complacency in highly rated

		departments • No accountability

Partial further dual support transfer	• More explicit targeting of resources to successful research groups	• Reduced scope for HEIs to deploy funds strategically and to support new areas and emerging excellence • Leaves same transaction costs on either side of divide
Combined teaching and research assessment	• Recognises universities' holistic function • Single administrative effort	• Different criteria applied • Greater tension with interdisciplinary and multidisciplinary research • RAE much lighter demand on administration than Teaching Quality Assessment • TQA a rolling programme
Rolling assessments	• Spreads central administrative burden • Could facilitate emergence of distinct processes for science and humanities • Could smooth funding changes	• Makes judgement of cross-disciplinary work harder if only part being assessed • Administrative burden always present in institutions and may require additional staff
Partial RAE, excluding 5/5*-rated (and possibly 1 and 2 rated) departments	• Reduces burden on those who have demonstrated ability to sustain high performance levels over the long term • Reduces cost	• Need to establish equality of 'new' 5's • Need to ensure 5* funding not diluted by grade-drift • Games playing – loading staff into departments not subject to assessment
Royal Academy of Engineering approach	• Recognises the different dimensions of excellence	• Difficult to implemented as a funding basis as currently conceived, because of problems aggregating, or weighting, the different dimensions of excellence

Annex F

Royal Academy of Engineering measures and characteristics of research

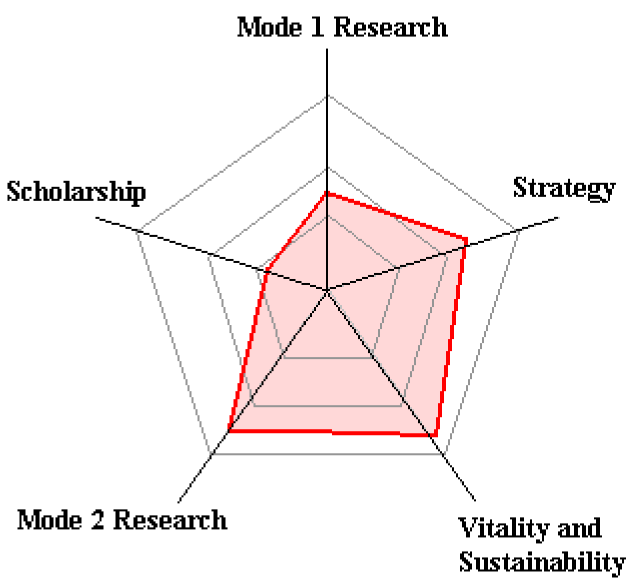
	Determines Impact on wealth creation and quality of life				
Characteristics / Measures	MODE 1 Research	MODE 2 Research	Scholarship	Vitality and Sustainability	Strategy
Publications	✓✓✓	✓	✓✓		✓✓✓
Peer Recognition	✓✓✓	✓✓✓	✓✓		✓✓✓
Consultation		✓✓✓	✓✓		✓✓✓
Partnerships and Industrial Support		✓✓✓		✓✓✓	✓✓✓
Patents and Practical Research Outputs		✓✓✓			✓✓✓
Independence Indicators			✓✓✓		✓✓✓
Core and Support Capability Indicators				✓✓✓	✓✓✓
Involvement with Society Indicators				✓✓✓	✓✓✓
Strategic Programme and Resource Planning Indicators	✓✓✓	✓✓✓	✓✓✓	✓✓✓	✓✓✓

✓✓✓ = Highly appropriate

✓✓ = Fairly appropriate

✓ = *Possibly appropriate*

The Royal Academy of Engineering report suggests that different research dimensions could be expressed as a research 'footprint', with the shape indicating performance along the different dimensions of excellence.



Royal Academy of Engineering measures and components of research performance

Measure	Components
Publications	Academic publications, conference publications, trade publications
Peer recognition	Professional recognition, e.g. FREng., FRS., international prizes, awards, honours
Consultation	Advisory committee membership, advisory services To Goverment/large business / small business / charities / pressure groups, standard setting panel membership
Evidence of partnerships, Industrial support And partnering activities	Large collaborative programme (number of, funding) e.g. LINK, consultancy reports, activities and processes, smaller collaborative programmes number and funding (e.g. Tcs, case), cluster membership, conference chairs, conference participation, personnel exchanges, workshop hostings, external-open seminars
Patents and Practical research outputs	Patents, licences, novel processes, products and instrumentation, spin offs/start-ups, researchers
Independence indicators	Prizes for non-basic, non-private funded research, proportion of research not funded by industry Proportion of research not included in foresight recommendations
Core and support capability indicators	Number of trained researchers and PhD students, staff experience, willingness to collaborate, teaching and presentation skills (qualifications, training) percentage research staff positions permanent, staff and student training and career development activities (number, quality and relevance) and employability, infrastructure quality and relevance, technical support staff number and qualifications, number of administrative support staff and qualifications/experience.
Involvement with society indicators	Number of staff/students engaged in public understanding activities, number of media articles/broadcasts related to research, number of staff with media relations responsibilities, staff with school relations responsibilities, activities and events held in connection with public understanding/school activities.
Strategic programme and resource planning indicators	Research strategy relevant to chosen mission and purpose, evidence of efficient budget management, quality of strategic management and planning evidence.

Annex G

Responses to the Call for Evidence - new and enhanced indicators

Outcomes in music include published work of internationally recognised quality, compositions and performances
Output measured by quality should be the first measure. Commercial exploitation could be used as a measure
The RAE gives little value to academic outreach – additional indicators such as number of citations or copies sold, editions and translations produced or use of work as general readers, textbooks or language-teaching texts would provide guidance on utilisation
For science and technology, commercial achievements could be weighted higher as an indicator of the performance of the academic base
The issue of industrial collaboration is currently included in the textual parts of the RAE submissions. It should be given a higher level of recognition. It is increasingly important to see that outputs of applied research are given recognition
The performance of the research base must be judged not only in terms of scientific criteria, but also in relation to the public interest; for instance the substantive or potential contribution to human welfare
The performance has been evaluated via the RAE concentrating on published outcomes. It is formula driven and does not take account of subject differences. There should be a broader evaluation to take account of research impact on quality of life
Some major aspects of research are never assessed at all, because the concern is with outcomes not processes. Much vital research never leads to an outcome
By the total quality of submissions to the research councils
A broader definition of research would allow a broader range of evidence to be considered, such as impact on quality of life, uptake of findings by industry, etc
The assessment process should adopt a wider definition of research, should provide for development as well as retrospective justification, should assess and fund ongoing research and collaboration
Quality should be assessed through the RAE while productivity should be assessed through PhDs awarded, publications, non-government research income
In engineering several indicators could be used – level of UK patent activity, active links with companies, start-up of new companies, and investment of international companies in UK research. Journal publications and citations are not the best tool to judge
Public funding of research as %GDP, output of trained researchers, research papers, support from research grants, commercialisation of projects, patents, licensing deals, economic growth and rate of new company formations
RAE takes no account of value for money or the time scale of the research. The current four publication rule and dependence on input factors of journals is very formula driven. There should be a broader evaluation of the quality of the research base
The present RAE methodology concentrates on quantity not quality – for 2001 it is four papers in referred journals. Citation statistics and reviews of books and papers, external recognition of staff and number of successful PhDs may be a better way
Peer assessment of quality of outputs, development of theoretical concepts, application, value for money, commercial viability and relevance to users could all be included. Differences among subject areas are

important
By the quality of outcomes: novel inventions/developments, patents, publications, and the exploitation of research
Results patents, processes, profits, publications, PhDs
The evidence that can be used is well documented. It would be valuable to add systematically an additional index which deals with exploitation (feasibility and actuality) of products from research
Technology transfer, spin outs, alternative forms of income generation by HEIs
Input data is fairly clear, on the output side there is some sophisticated work on citations which show the UK's comparative advantage in eg psychology
Publications and other outputs, citation analysis, evidence from users of research; movement of high-rated researchers. Research expenditure is important indirect evidence as are the number of PhD and other PG students
The way in which evidence- based health care has influenced patient care and patient outcomes is one example but there is still a long way to go in the production of high quality research evidence relevant to a huge range of health problems
The National Audit Office should conduct VFM audits of all funding allocated on selectivity grounds
One tangible assessment of the performance of the research base is delivery into industrial research programmes. Citation indices do not quantify the appropriateness of research only its extent - and then only approximately