

Annex A

Steering group: membership and terms of reference

Membership

Sir Gareth Roberts (Chair)	Wolfson College, Oxford
Sir Leszek Borysiewicz	Imperial College
Professor Vicki Bruce	University of Edinburgh
Professor David Eastwood	University of East Anglia
Professor Roderick Floud	London Metropolitan University
Professor Georgina Follett	Dundee University
Professor Alan Jackson	University of Southampton
Dr John Kemp	EVOTEC NeuroSciences GmbH
Professor Fabian Monds	Invest Northern Ireland
Sir Paul Nurse	Cancer Research UK
Professor Teresa Rees	University of Cardiff
Mr Phil Ruffles	Rolls Royce plc
Sir David Watson	University of Brighton
<u>Observers</u>	
Patricia Ambrose	Standing Conference of Principals
Gill Davenport	Scottish Higher Education Funding Council
Ann Hughes	Higher Education Funding Council for Wales
David McAuley	Department for Employment and Learning, Northern Ireland
Rama Thirunamachandran	Higher Education Funding Council for England (HEFCE)
Linda Thrift	Universities UK
<u>Secretariat</u>	
Vanessa Conte	HEFCE
Will Naylor	HEFCE
Tom Sastry	HEFCE
Dr Sian Thomas	HEFCE (until January 2003)

Terms of reference

1. The review will investigate different approaches to the definition and evaluation of research quality, drawing on the lessons both of the 2001 RAE and of other models of research assessment²⁶ and will advise on the future of research quality evaluation.
2. The output will be a number of models of research assessment and a short covering report to be presented to the chairmen and chief executives of the funding councils (including the Department for Employment and Learning, Northern Ireland). The report will either identify one preferred option or indicate the circumstances under which particular models would be most appropriate.

²⁶ The term 'assessment' is used here in its broadest sense to refer to any activity undertaken with the aim of providing information, assurance or feedback on the quality of research and associated activities and processes.

Annex B

Policy environment

Higher Education Funding Council for Wales (HEFCW)

1. HEFCW's policy for research operates within a broad framework provided by the Welsh Assembly Government's strategy for higher education, *Reaching Higher*, published in 2002, and the Council's own statement of research policy objectives. This was developed in consultation with the sector in 2000 and is consistent with the Assembly's wider strategy. The Council's policy is, essentially, to promote excellence, resilience, diversity, dynamism, sustainability and relevance in the research base.

2. The research base in Wales extends over all but three of the units of assessment used in the 2001 RAE, the exceptions being veterinary science, Russian and Asian studies. Twelve of the 13 institutions in the sector made submissions to the 2001 RAE. Approximately half of these can be described as research led; the remainder have research expertise in particular areas. Welsh institutions submitted a total of 2,289 FTE Category A research active staff to the 2001 RAE. Of these, 1,724 FTE were in units of assessment rated 4, 5 and 5*, an increase from 1,206 in 1996.

3. The quality of research in Wales has improved substantially since the 1992 RAE, and the outcomes of the 2001 RAE confirmed that it now stands comparison with that in the rest of the UK. In 2001, the overall weighted average rating for Wales was 5.2, only slightly below the figure of 5.4 for the UK as a whole. The rate of improvement in Wales over performance in the 1996 RAE, at 19%, was slightly above that for the UK, 18%.

4. Subjects in which there are particular research strengths in Wales include Celtic studies, civil engineering, environmental sciences, other studies and professions allied to medicine, psychology, and town and country planning, all of which achieved weighted average ratings which were above those for the UK as a whole.

5. Physical infrastructure for research in Wales has benefited significantly from investment from the various capital funding initiatives which have operated over the last few years. The sector received £35.4 million through the first round of the Science Research Investment Fund (SRIF), and will receive a further £46.7 million through the second round. Prior to SRIF, the HEFCW provided £22 million over two years through its capital grant for research infrastructure. Further support has come from HEFCW's Structural Change Fund, and the sector also secured a total of £20 million from the Joint Infrastructure Fund (JIF). Areas which have particularly benefited from large scale investment through these initiatives include biological sciences, clinical medicine, ocean sciences and optometry.

6. Two developments of particular significance are in research capacity and reconfiguration. HEFCW is seeking to build research capacity in areas of strategic importance to the economy, culture and society of Wales, and has established a fund for that purpose. Areas funded to date include nanotechnology, biosciences, public policy on health, and social sciences. Reconfiguration and collaboration between institutions is central to the Welsh Assembly Government's strategy for higher education. HEFCW has established a Reconfiguration and Collaboration Fund to take this

policy forward and expects to see significant reshaping and strengthening of the research base in Wales over the next few years.

Scottish Higher Education Funding Council (SHEFC)

7. SHEFC's funding through Main Quality Research Grant contributes towards the costs of providing a long-term, high quality research base that enables researchers to explore new ideas, and respond to opportunities and the research needs of the research councils and others.

8. SHEFC has a new Strategic Research Development Grant (SRDG), the aim of which is to fund research that supports national strategies intended to meet economic, social, healthcare, educational or other priority needs in Scotland and the UK. The SRDG scheme will provide SHEFC with a mechanism by which it can take a strategic view of the requirements of the research base in Scotland, and one that transcends institutional boundaries. Through the scheme SHEFC will seek to grow or strengthen areas of the research base that are of strategic importance to Scotland. It will also use the grant to create opportunities to bring together existing and complementary strengths to promote further improvements in quality and capability in strategically important areas of research for Scotland and the UK.

9. To inform research policy formulation, to learn of HEIs' research aspirations, and to promote the economies of scale, networking and appropriate research collaboration, SHEFC engages in bilateral dialogues with each HEI. In general terms, research in Scottish HEIs is very strong. Approximately 50% of researchers in Scottish HEIs work in RAE 5 or 5* graded departments. The average grade of all assessed departments in Scotland is higher than the UK average. Within the sciences, Scotland performed particularly well in biological and medical sciences, electrical and electronic engineering and computer science. Specific expertise exists in areas such as e-science, stem cell research and opto-electronics, areas of research strength that are also of strategic economic importance. A recent survey has shown that biotechnology and information and communications technology are the top two industry sectors targeted for university industry co-operation. Other areas of importance in which Scotland performs better than the UK average include geography, sports-related subjects, and drama, dance and the performing arts.

10. SHEFC works regularly with the Office of Science and Technology and the UK research councils on matters of common interest, including the implementation of the UK science strategy, 'Investing in Innovation: A Strategy for Science, Engineering and Technology' which was launched in July 2002. SHEFC also shares views on research policy and strategic priorities in discussions with individual research councils, the Arts and Humanities Research Board and major charity funders.

11. SHEFC also works with key stakeholders within Scotland on specifically regional policy. For example, Scottish Enterprise has announced the implementation of three Intermediary Technology Institutes (ITIs) focusing on energy, the life sciences and communications technology/digital media. These will be a new form of bridging institution between industry and the science base. SHEFC has been working with Scottish Enterprise to ensure that there is optimal fit between the ITI models and the Scottish research base so that the ITIs have the best chance to succeed. It also has regular dialogue with Universities Scotland, the Scottish Science Advisory Committee, and all Scottish Executive Departments that are significant users of the research base.

Higher Education Funding Council for England (HEFCE)

12. HEFCE's draft Strategic Plan 2003-08 has been developed within the broad policy framework of the Government's White Paper 'The future of higher education'. The plan envisages higher education in the future being significantly different in its structure, organisation and delivery, where institutions are both more diverse and increasingly interconnected. The plan is intended to provide universities and colleges with a secure and practical framework throughout and beyond the planning period.

13. Enhancing excellence in research is one of HEFCE's strategic aims. A dynamic, world-class research sector is vital to the health of universities and crucial to economic growth and social cohesion. HEFCE will allocate funding to develop and sustain excellent research across all subject areas. This funding will continue to be highly selective. All higher education institutions will be able to undertake research, selectively and funded from a variety of sources. HEFCE will work closely with other funders to develop a sustainable research base, and will encourage much greater collaboration between institutions.

14. Following publication of the Government's White Paper, and receipt of its grant letter from the DfES on 22 January, HEFCE has made further changes to the way funds for research are allocated.

15. Total recurrent funding for research in 2003-4 is £1,042 million, an increase of £102 million (10.9 per cent) over 2002-3. In distributing this funding the HEFCE Board decided:

- a. to maintain the average unit of resource for 5*-rated departments in real terms compared with 2001-2
- b. to provide supplementary research funding of £20 million to departments that achieved a rating of 5* in both the 1996 and 2001 RAEs (allocated for 2003-4 only)
- c. to restore in real terms the average unit of resource for 5-rated departments to 2001-2 levels
- d. to allocate any remaining quality-related research funding (QR) to 4-rated departments
- e. to discontinue mainstream QR funding for 3a-rated departments, but to establish a capability fund, amounting to £20 million, to support research in specific, emerging subject areas where the research base is not as strong as in more established subjects
- f. to provide £2 million for veterinary research, through a joint initiative with the Department for Environment, Food and Rural Affairs (Defra).

16. The HEFCE Board will review the basis for distributing research funding for 2004-5 and beyond, to take account of the Government's priorities signalled in 'The future of higher education'. The grant settlement suggests that there is unlikely to be any real terms increase in total recurrent research funding in 2004-5.

Department for Employment and Learning, Northern Ireland (DEL)

17. The Department for Employment and Learning (DEL) fulfils the role of a higher education funding council in Northern Ireland.

18. Research carried out by the universities in Northern Ireland – Queen's University, Belfast (QUB) and the Ulster University (UU) – plays a disproportionately large role in the economy when compared to the rest of the UK: 30% of all research undertaken in NI is carried out by the universities. This is mainly due to the preponderance of small to medium-sized enterprises (SMEs), which are traditionally unwilling to invest in research. It is estimated that only 10 large firms undertake 60% of the remaining research. The position of DEL, therefore, as the public body providing the majority of funds to the university research departments, is central to the success or otherwise of research activity in Northern Ireland.

19. University research funding is an area where some responsibilities are devolved and others reserved by the UK Government. DEL provides recurrent grant (quality-related research or QR) for infrastructural needs (permanent academic staff, premises, libraries and central computing costs) – as well as for the costs of training postgraduate researchers and conducting a certain amount of 'blue skies' research. It also provides capital funds. The Research Councils, using OST funds, provide for direct project costs and contribute to indirect project costs.

20. QR is calculated with reference to the most recent Research Assessment Exercise (RAE). Both Northern Ireland universities improved their performance compared with previous years: over 70% of research departments are achieving attainable levels of national excellence (a score of 4 or above), compared to 37% in 1996. Top Northern Ireland performers include mechanical, aeronautical and manufacturing engineering at QUB (5*), and biomedical sciences and Celtic studies at UU (both also 5*); UU submitted to 28 UoAs, QUB to 40.

21. For the academic year 2002-3, recurrent research funding for the NI universities stood at £29.65 million (£17.78 million to QUB, £11.87 million to UU). The baseline for 2003-4 has been increased to £35.25 million. The capital baseline for 2002-3 stood at £8.25 million, available for all capital requirements including research.

22. In addition to QR funding, DEL also provides funds for specific initiatives, set up to improve the quality of the research infrastructure, and to initiate new capital projects, such as the Support Programme for University Research. The first phase of this will invest a total of £43.6 million over the period 2001-5, the second phase will invest £50 million over 2003-7. Both phases are public-private partnerships on a pound for pound basis, with DEL contributing 50% of the funding.

23. DEL also administers funding originating from OST in the form of the Science Research Investment Fund (SRIF), running from 2002 to 2004. For Northern Ireland, the share of the OST element of this stream amounts to £7 million, and the universities contribute £2.3 million, which will enable four projects to be built.

Annex C

A guide to the 2001 Research Assessment Exercise

Overview

1. The RAE operates through a process of peer review by experts of high standing covering all subjects. Judgements are made using the professional skills, expertise and experience of the experts; it is not a mechanistic process. All research assessed is allocated to one of 68 'units of assessment' which are discipline-based. For each unit of assessment there is a panel of between nine and 18 experts, mostly from the academic community but with some industrial or commercial members.
2. Every higher education institution in the UK may make a submission to as many of the units of assessment as they choose. Such submissions consist of information about the academic unit being assessed, with details of up to four publications and other research outputs for each member of research active staff. The assessment panels award a rating on a scale of 1 to 5*, according to how much of the work is judged to reach national or international levels of excellence (see paragraph 15 below).

Units of assessment

3. There are 68 units of assessment in the 2001 RAE. Each unit covers a broad subject area: for example, mechanical, aeronautical and manufacturing engineering are included within one unit; drama, dance and performing arts are all included in another. The units of assessment have been identified in consultation with the higher education sector, and continue to evolve to reflect changes in the pattern of research in institutions.

Assessment panels

4. There are 60 assessment panels; usually there is one panel for each unit of assessment but a few units of assessment have joint panels. The panel chairs were nominated by members of the 1996 RAE panels and appointed jointly by the four funding bodies. Panel members are nominated by a wide range of organisations including research associations, learned societies, professional bodies and those representing industrial, business and other users of research. Panel members are then selected by the funding bodies, on the advice of the panel chair, based on their research experience and standing in the research community, so as to ensure coverage of the subject concerned. The funding bodies seek to reflect the profile of nominations received in terms of geographical coverage, gender and type of institution. The chair and members of each panel participate as individuals, rather than representatives of a particular group or interest. The names of the panel chairs and members are published.
5. Nearly half of the panels have established sub-panels; these include people who are not members of the main panel. The sub-panels advise on assessment of research in particular sub-areas within the subject. Panels may also draw on the advice of specialists covering specific areas of expertise outside the panel's experience. In addition, all panels consult with advisers based outside the UK, to confirm their application of the standard of international excellence which is the benchmark for the exercise.

Institution submissions

6. Each publicly funded university and higher education college in the UK is invited to submit information about their research activity for assessment. The information they supply provides the basis on which judgements are made. Submissions have to be in a standard format, which includes qualitative and quantitative information. Most of the information is provided electronically on specially written software.

7. The submissions are based around members of staff in each academic unit in which the institution is submitting. It is up to each institution to decide which subjects (and therefore which units of assessment) to submit to, and which members of staff to include in each submission.

8. For each member of research staff, up to four items of research output may be listed. All forms of research output (books, papers, journals, recordings, products) are treated equally; panels are concerned only with the quality of the research. Similarly, all research (whether applied, basic or strategic) is treated equally. In addition, the HEI must provide information in a number of different categories, as shown below.

Category	Description
Staff information	<ul style="list-style-type: none">• summaries of all academic staff• details of research-active staff• research support staff and research assistants
Research output	<ul style="list-style-type: none">• up to four items of research output for each researcher
Textual description	<ul style="list-style-type: none">• information about the research environment, structure and policies• strategies for research development• qualitative information on research performance and measures of esteem
Related data	<ul style="list-style-type: none">• amounts and sources of research funding• numbers of research students• number and sources of research studentships• numbers of research degrees awarded• indicators of peer esteem

Assessment period and census date

9. The 2001 RAE used a census date of 31 March 2001 to capture a 'snapshot' of all staff in post on that date. The census date determines the staff eligible to be included in the submissions of each institution. The RAE assesses the quality of research output by eligible staff over a period of seven years for arts and humanities subjects, and five years for all other subjects.

10. The deadline for institutions' submissions was 30 April 2001. The month between the census date and the submission date allows institutions time to finalise their submissions, although they are provided with the details of the submission requirements at least 12 months before each RAE.

11. The actual research outputs do not all have to be submitted, but must be made available to assessment panels on request.

Panels' judgements

12. The panels use their professional judgement to form a view of the overall quality of the research in each submission within their unit of assessment using all the evidence presented in the submission.

13. To assess submissions fairly and consistently within each unit of assessment, each panel draws up a statement describing its working methods and assessment criteria. These are published in advance of submissions being made. This statement shows which aspects of the submission the panel regards as most important, and areas that it wants institutions to comment on in their submissions. The differences in working methods and criteria between panels is a reflection of the need to recognise differences in the way research is conducted and published in the various disciplines.

14. Panels review all submissions, and read selectively from the research outputs cited. As the panels are concerned with quality, not quantity, information on the total number of publications produced is not requested. Panels do not visit institutions as part of their work.

Ratings

15. The subject panels use a standard scale to award a rating for each submission. Ratings range from 1 to 5* (five star), according to how much of the work is judged to reach national or international levels of excellence. The table below shows the definition of each rating.

The rating scale

Rating	Description
5*	Quality that equates to attainable levels of international excellence in more than half of the research activity submitted and attainable levels of national excellence in the remainder
5	Quality that equates to attainable levels of international excellence in up to half of the research activity submitted and to attainable levels of national excellence in virtually all of the remainder
4	Quality that equates to attainable levels of national excellence in virtually all of the research activity submitted, showing some evidence of international excellence
3a	Quality that equates to attainable levels of national excellence in over two-thirds of the research activity submitted, possibly showing evidence of international excellence
3b	Quality that equates to attainable levels of national excellence in more than half of the research activity submitted
2	Quality that equates to attainable levels of national excellence in up to half of the research activity submitted
1	Quality that equates to attainable levels of national excellence in none, or virtually none, of the research activity submitted

Funding allocations

16. Each of the four funding bodies uses the ratings to allocate research funding by formula to the institutions it funds. The formulae used by each funding body may vary, with the overlying principle of funding selectively – more funding for higher quality research.

17. The main way of measuring the volume of research is by the number of research active staffed submitted to the RAE for assessment.

The outcome

18. The ratings from the 2001 RAE were published in full in December 2001. In addition to the ratings, a general report on each unit of assessment was published, and feedback on each submission was sent to the individual institutions in confidence.

19. Extensive information about the RAE is available on the RAE web-site www.hero.ac.uk/rae. This includes all the publications describing the RAE.

Annex D

Operational review of the 2001 Research Assessment Exercise

Summary of a report by John Farrant, David Billing and Paul Temple of Universitas

Terms of reference and method

1. The objectives of this study, commissioned by the Research Assessment Review Team, are to:
 - a. Provide the Steering Group for the Research Assessment Review with a formative and summative evaluation of the 2001 RAE by SWOT analysis (strengths, weaknesses, opportunities and threats). This evaluation should be focused on the performance of the RAE against the overall objective of providing ratings of quality of research conducted in HEIs, rather than on the behavioural, funding and other impacts of the RAE on the sector.
 - b. Draw out lessons for the future of research assessment – highlighting those aspects of the exercise that worked successfully and could desirably be retained in any future exercise; and those which in practice proved problematic to manage, measure or otherwise prosecute successfully.
 - c. Take into account the comments received from key participants in the 2001 RAE, including panel chairs, members, secretaries and institutional contacts.
2. Our evidence comprised:
 - RAE documents published on its website
 - the replies to our questionnaires to all panel chairs (88% response rate) and to all institutional contacts (61%)
 - discussions with funding council officers and former members of the RAE team
 - submissions to the research assessment review making reference to RAE 2001
 - a relatively small amount of directly relevant material on paper files retained by HEFCE.

We are extremely grateful to all those who responded to the questionnaires and who discussed with us their experience of RAE 2001. Our conclusions are based predominantly on the perceptions of prime actors in the RAE, as recorded at the end of the assessment phase or as much as 15 months later, rather than on analysis of operational documentation.

General conclusions

3. RAE2001 was successful in its primary purpose: ratings of research quality were produced, to timetable, and were used by the funding bodies in determining grant for 2002-3. Furthermore, the ratings have commanded a large measure of confidence amongst the HEIs, the researchers assessed and the academic community more widely.
4. A timetable for the assessment phase even tighter than in 1996, the workload on key players (panel members and secretaries, and the RAE team), and a demand higher than predictable for the services supporting the panels' decision making, together risked a major disruption of the assessment phase, though none fortunately occurred. A significant contributory factor to the second and third of these factors was panels scrutinising research outputs more thoroughly than they had

proposed in their statements of working methods. The same degree of dedication and commitment which all those involved showed cannot be assumed for any similar further exercise. More staff (or funds to outsource services) would be required; and all inputs should be realistically costed and paid for.

5. One of the most important principles by which the RAE was governed is 'Consistency: assessments made through the RAE should be consistent especially across cognate areas and in the calibration of quality ratings against international standards of excellence.' This was the aim whose achievement institutional contacts judged to be least satisfactory. Several of the elements of the RAE process bearing on consistency were among those which attracted critical comments from panel chairs:

- the grade definitions
- cross-referrals between panels
- umbrella groups
- non-UK-based advisers.

To these may be added:

- the diversity in details in the criteria and working methods for panels in cognate areas
- the substantial variation between panels in the degree of scrutiny given to research outputs.

The issues of the extent to which consistency is necessary and of how it is to be achieved must rank high on the agenda in the design of any future RAE.

6. The 2001 RAE followed fairly closely the pattern of the 1992 and 1996 RAEs, so it is to be expected that most elements operated satisfactorily, as indeed they did. Those which operated less than satisfactorily or had less than satisfactory outcomes, in addition to those in the previous paragraph, and should command priority for close examination if they were to be incorporated in any future assessment exercise, are:

- networking of the RAE data application
- preparation of ad hoc analyses of data for panels
- electronic communication during the assessment process
- arrangements for panel meetings
- sourcing of research outputs for panel members
- panels' overview reports
- feedback to institutions.

The central machinery of the RAE

The Project Group

7. The Project Group of senior officers of each funding body embodied a considerable pool of knowledge and experience, and was probably the most significant element of administrative continuity between the 1996 and 2001 RAEs. It provided an essential link with the RAE's commissioning bodies. If the RAE team is to be constituted from scratch for each exercise and if the Group is the main element of continuity (but with a limited range of expertise, given its composition), then the Group should pay close attention, particularly in the early stages, to ensuring that all sources of sound advice are exploited and that proposals coming to it are well founded. It should be serviced outside of the RAE team.

The RAE team

8. The achievement of the RAE team was impressive by any standards, reflecting well both on the capabilities of team members and on the leadership they received. But staffing it largely by fixed-term appointments led to difficulties. The detailed knowledge built up during the 1996 exercise had been largely lost, and had to be recreated by the new team – which itself has now dispersed. The fixed-term contracts are likely to have increased the rate of staff turnover and to have contributed to the ragged run down in early 2002. HEFCE's Analytical Services Group was fortunate in being able to deploy the same staff for the data collection and analysis for 2001 as for 1996.

9. The view among panel chairs is that the RAE team was over-stretched and understaffed during the assessment phase, contributing to their criticisms of the handling of cross-referrals, the short suspension of the Research Outputs Sourcing System, the arrangements for meeting venues, etc. The volume of requests from panels was much greater than past experience suggested, and could have been predicted from their statements of working methods. It is highly probable that the team did indeed need more staff during the assessment phase, irrespective of turnover.

10. A rolling RAE (that is, a planned cycle of discipline-based reviews supported by a permanent secretariat) would greatly ease these administrative problems but would of course have to be adopted on policy grounds. If the RAE is continued on similar lines, continuity of staffing in its administration is still highly desirable. There are tasks for the quiet years: dealing with complaints and challenges, evaluating in detail individual processes (such as cross-referrals), reviewing the configuration of units of assessment, monitoring the burgeoning literature on the RAE, etc. But given the length of the cycle and the inevitability of staff turnover, staff continuity will not be sufficient to preserve the organisational memory from one exercise to the next. The systems for preserving and organising the knowledge base need to be far stronger. Maintenance of corporate memory of the assessment process from start to finish is, in any case, essential if the funding bodies are to defend gradings when challenged.

The panel secretaries

11. The recruitment of panel secretaries from HEIs as well as from the funding councils was a successful innovation. The great majority of panel chairs felt their panels were well served. But workload problems identified in 1996 – the impossibility of acting as a panel secretary during the assessment phase and doing the normal job, or most of it, simultaneously – were not resolved. Employers were not reimbursed on a full-cost basis.

12. The demands on the secretaries from the funding councils had some impact on other work, by delaying initiatives which might otherwise have been carried forward. But this impact was less than in 1996 because certain functions (such as booking accommodation, payment of expenses) were placed with the RAE team in 2001. Because the RAE team was housed in the HEFCE's headquarters at Northavon House, HEFCE corporate services supported it and suffered some disruption. Use of these services seems not to have been charged to the RAE.

The framework for assessment

Rating scale and grade definitions

13. Although the majority of panel chairs found the grading system readily workable, many were concerned about the need to interpret the grade definitions, and therefore about inconsistency of interpretations across panels. The main reasons given were the breadth of the grade bands, especially of grade 5 relative to grade 4, unclear borderlines, and 'international' and 'national' excellence being undefined – for which several panels laid down their own definitions.

Statements of criteria and working methods

14. Panel chairs in the great majority of fields found the generic criteria and working methods sufficiently broad as to allow each panel to shape its statement so that it was appropriate to its field; but there were concerns whether there was consistency with related fields. The published statements (RAE 5/99) indeed reveal many differences between panels in cognate areas, some less understandable than others.

15. Institutional RAE contacts were in the majority satisfied with the consultation on the draft statements and, while submissions were being prepared, found the statements unproblematic. The most frequent comments concerned ambiguity or lack of explicitness, and inconsistency across panels in cognate fields.

Preparation of submissions

Guidance on submissions

16. Contacts were generally complimentary about the clarity of, and long notice given by, the Guidance on Submissions (RAE 2/99). But more thorough review of wording taken over from 1992 and 1996 might have reduced the stream of clarifications through the Briefing Notes and the FAQs on the website. Clarifications were still appearing late in the preparation of submissions, and these were not consolidated into an updated (online) guidance document. The most frequently mentioned area of difficulty was the categorisation and definition of staff eligible for submission. The briefing seminars were widely appreciated.

17. Panel chairs considered that the funding bodies and the RAE team had done as much as reasonably possible to inform the sector about the RAE's workings. But myth, rumour and misinformation continued to circulate.

RAE data application

18. HEFCE's Analytical Services Group was contracted to prepare the software with which institutions prepared and submitted their submissions. Contacts reported problems where they networked the software for departments' use and with the upload facility. Though few rated the software, documentation and other support as poor or bad, a substantial minority were uncertain, leaving only half satisfied.

Standard audit and verification

19. Institutions found that the standard arrangements for audit and verification of submissions did not impose unreasonable burdens, though there were complaints about timing (during the holiday season) and the volume of paper. The RAE Manager's conclusion is that the data contained in RAE submissions were of a very high standard of accuracy. We have not attempted to form a view on whether the audit checks were sufficiently rigorous or whether further errors or deliberate false information may have gone undetected.

The panels

Panel membership

20. While broadly content with the procedures for the appointment of panel members, and recognising that nominations from professional bodies and subject associations were important to ensure credibility, panel chairs felt that these had too much weight and that some were implausible and unbalanced. The two issues most raised by contacts were that the HEIs were not invited to nominate; and that the process for selecting members from the nominations was not transparent and was likely to lead to ossification.

Representation of research users

21. Panel chairs were evenly divided on whether their panel's decision-taking was enhanced by access to expertise and advice of research users. Neither user sub-panels nor individual users as panel members came out as more effective. The latter had little effect on the final grades but were useful for credibility.

Sub-panels

22. The opportunity to constitute sub-panels was not widely used, the most extensive instance by far being the seven functioning under the three medical panels. Where used, the sub-panels worked well. For the future, sub-panels which span a range of panels might be adopted more extensively – perhaps to the point where they might provide an assessment matrix which was more sensitive to changing patterns of research activity than the main panels alone.

Workload

23. The general view of panel chairs was that the load on panel members was excessive – for inadequate remuneration and with loss of research time.

Panels' differential loading

24. Researchers submitted for assessment per panel member ranged from 8 to 196. The degree of scrutiny accorded to submissions tended to expand to utilise the capacity available. One panel with a lighter workload reported reading all outputs at least once and in most cases twice; some panels with heavier loads committed themselves to reading no more than a minimum of 10%. The resources available to each panel were broadly the same.

25. We have not noted any assertion that the grades awarded by the heavily loaded panels are any less reliable than the grades awarded by lightly loaded panels. If the grades awarded by the

heavily loaded panels command acceptance in their respective communities, less thorough scrutiny might be adopted in other units of assessment, with more differential allocation of resources in support of panels, or with some merging of panels supported by sub-panels.

Confidentiality and declarations of interest

26. Panel members were required to sign a confidentiality agreement. It is evident from some of the more serious challenges to gradings – and anecdotes in circulation – that the agreement was not universally observed.

The assessment process

Communications

27. On grounds of security, panel members were denied the submissions in any format other than paper, and were prohibited from using e-mail to distribute or discuss RAE data and panel business. The design of any future assessment exercise should enable those involved to use what are now the working tools of everyday life. The circumstances and requirements of the RAE for security and secrecy cannot be so exceptional that technical solutions are not readily available.

Timetable

28. The elapse time available to panels for assessment was shorter than in 1996 because two new steps were introduced for 2001: consultation with non-UK experts and discussion by umbrella groups. The volume of assessment activity was greater than in 1996, in terms of outputs read (and therefore requested), referrals to specialist advisers, cross-referrals between panels, preparing workbooks, and so on. Chairs' criticisms of assessment processes arose mainly from the massive peaking of activity during the summer of 2001, and from the consequent load on the panels and the RAE team. If the 2001 exercise were replicated, doubtless with some further complexities added, the same timetable would with high probability be unworkable. The risk to the exercise from staff resignation or absence – and also from IT system failure – was high.

Data collection and analysis

29. Panel chairs seem to have been broadly satisfied with the range of data collected in the forms RA1 to RA4 (covering staff details, research outputs, student details and research income), and the standard presentations of the data. But they criticised the limited ability to produce quickly ad hoc statistical analyses. With the data available to them only on paper, some panels re-keyed data and created their own databases.

Audit queries

30. Panel members raised over 300 queries about information in submissions, for investigation by the RAE team. A couple of chairs complained at the slow response to their queries and at the queries not being pursued hard enough.

Selection of staff for submission

31. Departments selecting fewer of their staff for returning in the RAE, compared with 1996, is one explanation of the increase in gradings seen in most units of assessment. A couple of panels felt that institutions should not have the option to omit staff. Selecting staff fails to recognise the team effort required to undertake research.

Use of RA5 and RA6

32. Panel chairs were split between those for whom the textual commentary in forms RA5 and RA6 was important or essential, and those for whom it was of little or marginal use. As HEIs have become better at writing 'hype', so the statements failed to discriminate. The large majority of chairs were confident that panels had been able to take appropriate account of 'staff circumstances' as set out in RA6. The major use of this information was for new and young researchers.

Panel workbooks

33. The panel workbooks were an innovation for 2001. Their drafting, a substantial task, was generally undertaken by the panel chair and secretary. Those that we examined were impressive for their clarity and apparent comprehensiveness. These are clearly important documents, which will be needed to rebut any legal challenge to the process.

Arrangements for panel meetings

34. Many panels reported that the quality of the accommodation, particularly in London, had been poor; that the centralised accommodation booking service was overloaded and often inflexible; and that support for panels in organising accommodation needed to be reconsidered. It seems that the maximum cost per head allowable under Civil Service rules was unrealistically low for London.

Research Output Sourcing System

35. HEIs were required to make available, on request, any item cited in their submissions. For items not readily available to a panel member, the RAE team established a central Research Outputs Sourcing System (ROSS). ROSS was suspended for the receipt of new requests for three days, at the height of the demand for items, because of the backlog of items received and awaiting redistribution to panel members. It may have been the victim of its own success, with some members using it in preference to consulting copies in their own institution's library. Contacts registered surprise at the high volume of requests, the high cost in staff time and postage, occasional losses or mislaying of items in handling at the RAE team's base in Bristol, and items returned damaged or not at all. Some panel chairs and contacts suggest that for another RAE institutions should deposit in one location a copy of each cited output, or that all outputs should be available electronically.

Specialist advisers

36. Some panel chairs registered concerns about the RAE team's efficiency in handling referrals to specialist advisers, about the variable quality of the advice received, about the time that advisers had to perform their tasks, and about their pay.

Cross-referrals

37. The system for cross-referrals between panels attracted many negative comments. The mandatory cross-referrals were relatively straightforward, though perhaps should not be permitted for single items of research output. The major difficulty lay with cross-referrals generated by panels. Because of the concerns about the security of electronic communication, the processes were entirely paper-based and became clogged up. Some panels by-passed the official procedure. There were then concerns about incorporating the advice received, given panels' heterogeneous ways of assessment.

38. If this element is likely to be a feature of a future RAE with a similar configuration of units of assessment, there should be an analysis of the flows of cross-referrals, as it may pinpoint some sub-optimal aspects in the configuration of units of assessment and in the composition of individual panels.

Non-UK-based advisers

39. Chairs' comments were mostly negative: the non-UK experts were not properly briefed, had too little understanding of the criteria and process, were often out-of-date about UK research, needed closer integration into the process (whereas previously they had been members of some panels), had too little time to do the job, were variable in the quality of their advice, were under-paid, and relied too much on their existing opinions about individuals' or departments' reputations. Nevertheless, in the overview reports, panels were pleased to note that the non-UK-based advisers validated the 5 and 5* gradings.

Umbrella groups

40. Umbrella groups were another innovation in 2001 which did not work effectively. Chairs' positive comments were on the lines of 'useful but of limited impact'. Where there were real discrepancies in panels' criteria, practices or grading profiles, the umbrella group process had no force. If their meetings had been more appropriately timed, they might have functioned better.

Interdisciplinary research

41. There has been continuous controversy around the RAE's capability to do justice to interdisciplinary research. Panel chairs were confident that the arrangements and criteria enabled the panel to give interdisciplinary research proper consideration, claiming that their field was interdisciplinary or was marked by interdisciplinarity or that the panel's membership encompassed interdisciplinary expertise. Only a small minority invoked cross-referrals or specialist advisers as the means by which the panel ensured proper consideration. This confidence in the equitable treatment of interdisciplinary work is not as widespread in the institutions: more contacts considered that interdisciplinary research could *not* be submitted appropriately, than that it could be.

42. After the 1996 RAE, the funding bodies commissioned a large study from Evaluation Associates Ltd on 'Interdisciplinary research and the Research Assessment Exercise' (March 1999, issued as RAE 1/99). We have not attempted to establish how each of its recommendations was implemented and then to assess the effectiveness of the measures taken. That this should be done is highly desirable if there is to be a further RAE of a similar character, as the assessment of interdisciplinary research continues to be a contested area.

Collaborative research

43. Over three-quarters of the panel chairs responding were of the view that the assessment mechanisms did *not* disadvantage research, of which evidence was submitted, which had been undertaken collaboratively with partners within and beyond higher education. Again, contacts were more divided, but those thinking that collaborative research could be submitted appropriately outnumbered those who thought not.

After the assessment

Overview reports

44. Half the chairs were uncertain whether the overview report had proved valuable to their subject community and beyond. Those more positive were still muted in their assessment. For the future, we suggest that either overview reports are dropped or more effort, with more licence, is put into them.

Feedback to institutions

45. The feedback reports to institutions attracted even less positive response from chairs. The feedback's potential utility was emasculated by fear of judicial review, by the lack of time and secretarial support for its preparation, and by insufficient clarity of purpose. Nor was there greater enthusiasm in the institutions: no other question elicited so low a percentage of positive responses from contacts, who found much of the feedback anodyne in the extreme. For the future, the implications of the Freedom of Information Act 2000 must be considered: there may be no exemption in the Act which would protect the workbooks from disclosure.

Complaints about the outcome

46. The number of ratings giving rise to written complaints was of the order of 60. About a dozen were pushed to the point where HEFCE felt obliged to seek legal advice, and one went for counsel's opinion in anticipation of judicial review. The ragged disbanding of the RAE team meant that there were inadequate arrangements at HEFCE for handling these complaints.

Matching the key principles: institutions' assessment

47. Contacts' assessment of the RAE's achievement against each of the key principles laid down at the outset is shown below:

RAE 2/99, paragraph 1.3	%Yes	%No	Yes/No
c Consistency	30	32	0.9
h Parity	28	25	1.1
g Neutrality	36	25	1.4
j Transparency	43	23	1.9
f Efficiency	49	17	2.9
e Credibility	55	10	5.5
d Continuity	54	8	6.8
a Peer review	66	9	7.3
b Clarity	53	6	8.8

The four principles where achievement was judged least satisfactory were:

Consistency. Assessments made through the RAE should be consistent especially across cognate areas and in the calibration of quality ratings against international standards of excellence.

Parity. The RAE is concerned only with assessing the quality of research of participating HEIs, regardless of its type, form or place of output.

Neutrality. The RAE exists to assess the quality of research in HEIs. It should carry out that function without distorting what it is measuring. In other words, the RAE should not encourage or discourage any particular type of activity or behaviour, other than providing a general stimulus to the improvement of research quality overall.

Transparency. The credibility of the RAE is reinforced by transparency about the process for making decisions. Except where there is a need to preserve confidentiality (for example in panels' discussions or when dealing with the names of nominees for panel membership or with the strategic research plans of institutions) all decisions and decision-making processes will be explained openly.

Annex E

Analysis of responses to the 'Invitation to contribute' to the review of research assessment

Background

1. One of our main concerns when we began the review was to make sure that our deliberations were thoroughly informed by the views of stakeholders, both within and outside the HE sector. To identify these views, we launched our work with an invitation to contribute on 27 September 2002. This was sent to a wide range of bodies, including higher education institutions, learned societies, major research charities, and companies with interests in research and development. We also published an open invitation to contribute on the review website, www.ra-review.ac.uk, which produced a large number of responses from individuals.
2. The invitation to contribute closed on 29 November 2002. We received 414 responses, which were divided into four categories:
 - a. Higher education institutions (HEIs).
 - b. Subject bodies, departments, faculties and learned societies.
 - c. Individuals responding on their own behalf or on behalf of small groups of individuals.
 - d. Stakeholders, including sub-sector groupings such as the Russell Group, and bodies outside the HE sector such as companies and charities.
3. To identify any marked preferences by subject or discipline, we divided categories 2b and 2c where possible into five sub-sections based on the umbrella units of assessment in the 2001 RAE: medical and biological sciences; physical sciences and engineering; social sciences; area studies and languages; and arts and humanities²⁷.
4. We invited responses to six different groups of questions. The first four groups asked them to identify a preferred mechanism for assessing research built from one or more of four components: expert or peer review, an algorithm based on metrics, self-assessment, and historical ratings. The fifth group invited comments on nine cross-cutting issues, including whether each subject should be assessed in the same way, and how research assessment could be designed to support equality of treatment for different groups of people in higher education. Finally, the sixth section simply invited respondents to comment on any other issues they thought we should address. The invitation to contribute remains accessible through the review website at www.ra-review.ac.uk
5. Each response was read in detail for qualitative information responding to the six groups of questions outlined above. The frequency of different types of responses to particular questions was also recorded, enabling us to make some quantitative comparisons. In the following paragraphs we present a summary of this analysis, taking each group of questions in turn. A full analysis appears on the website.

²⁷ In some cases it was not possible to determine a suitable umbrella group. Hence the total number of responses for subject bodies and individuals is higher than the sum of the constituent umbrella groups.

Expert review

6. There is overwhelming support for the continued use of expert review, organised around cognate areas of research, as the principal means of assessing UK research.

- a. Of those responses that make a clear statement on the matter, at least two-thirds in each category maintain that research assessment should be carried out principally by expert review.
- b. Support for expert review is particularly strong among HEIs and subject bodies. Among subject bodies, support is consistent across the five subject groups.
- c. A significant proportion of responses in all categories call for improvements in the consistency and transparency of the RAE expert review system, and in its treatment of interdisciplinary research.

7. According to its proponents, the efficacy of expert review has been demonstrated repeatedly by the success of the RAE in delivering results widely regarded as accurate by the community. This accuracy in turn inspires confidence among the academic community, who are thus far more likely to accept the results of the exercise. The following extract from the University of Manchester is typical:

'In contrast with a number of other forms of cross-institutional review and assessment in the higher education sector, the RAE has retained a good degree of support amongst the academic staff who are its subjects. The university considers this favourable situation to be, in large part, due to the strength and widespread acceptability of expert peer review as the RAE's core assessment methodology, and would therefore strongly recommend that it remain the key component of any future mechanism.'

8. Strong support for expert review, however, does not indicate that the process as practised by the RAE is regarded as ideal. Most responses in all categories that support the maintenance of expert review also propose changes in the workings of expert panels. These changes largely revolve around three issues: transparency (particularly in the selection of panel members and the weighting of assessment criteria); consistency; and improvements in the assessment of interdisciplinary research.

Algorithm based on metrics

9. Over half of all responses that express a clear preference agree that metrics should play a greater role in research assessment.

- a. Among HEIs, half of all responses (including those that do not express a clear preference) agree that metrics should be used to support the work of expert panels.
- b. Among stakeholders and subject bodies, of those making a clear statement, about half endorse the supporting use of metrics.

10. To most respondents, metrics alone are regarded as too crude a means to assess the quality of research (even in the hard sciences), and ill-suited to judge research culture and the strategy and vision required to attain research excellence. Furthermore, they argue that the sole use of metrics would:

- a. Distort UK research towards the counter-productive, short-term pursuit of largely irrelevant statistics.
- b. Preclude any prospective element in the assessment process.
- c. Favour established 'mono-disciplines' at the expense of emerging, innovative and/or interdisciplinary research.
- d. Only offer an illusion of objectivity, since many of the metrics proposed in the call for evidence are constructed through a series of subjective judgements.

11. Although there is very little support for the sole use of an algorithm to determine research quality, over half of responses expressing a clear preference agree that metrics should play a greater supporting role within research assessment. They felt this would reduce burden and costs and better inform the subjective judgements of panels. This is particularly the case among HEIs, where about half of all responses agree that metrics should be used to support the work of expert panels.

12. Unfortunately there is little consensus around precisely which metrics should be used to support research assessment. The data show modest support for bibliometrics, research income, expenditure/value for money and research student numbers, but most responses expend far more effort warning us off these and other measures. Respondents do agree that any metrics should be: appropriate for each subject area (with a range of different algorithms if required); thoroughly tested for the promotion of undesirable results; and made explicit to the community well in advance of the exercise.

Self-assessment

13. Among those responses making a clear statement, about half think that self-assessment should play a part in research assessment, while the other half oppose any extension.

- a. Support for self-assessment is relatively strongest among stakeholders.
- b. Among the other three categories, and across the five subject areas, there is an even split between advocates and opponents.

14. Few advocates argue that self-assessment is the perfect way to assess research. However, its apparent capacity to enable individual HEIs within a diverse sector to plan, pursue and manifest research quality according to local conditions is regarded as the best way to increase research quality and capacity. Self-assessment also attracts support from respondents who see it as a means to cut down the workload of expert review panels and thus the overall administrative burden of the RAE. However, this view is by no means universal. Many of the responses that oppose any extension of self-assessment whatsoever (roughly half of those that express a preference), argue that self-assessment would lead to an increase in administrative burden, since all institutional assessments would need to be carefully audited by expert panels in order to maintain confidence in the system. Other reservations include:

- a. The risk of research assessment becoming a more adversarial and disputed process.
- b. Potential preoccupation with the management of the research process rather than the academic merits and contribution of the research being generated.

- c. Possible descent into an assessment of the creative writing skills of self-assessors, rather than the quality of the research itself.

Historical ratings

15. Fewer than 1 in 20 responses endorse any use of historical ratings whatsoever. Of the remainder, about half oppose the use of historical ratings, while the rest do not make a clear statement on the matter.

16. There is almost no support for the use of historical data, except for a few responses that recommend historical ratings as a means to establish the extent to which strategic objectives had been met or value for money delivered over the assessment period. The vast majority of responses that express a preference tend to oppose historical ratings, as a recipe for complacency among research intensive HEIs and alienation among the rest. Many responses also point out that the retrospective expert review process operated by the RAE, coupled with the increasingly large funding gaps between different RAE grades, already serves to reinforce historical divisions in the sector.

Cross-cutting themes

17. There is significant support for a broader definition of research within research assessment, to encompass in particular applied research, research of relevance and utility, training of research students, and research that directly informs teaching.

Applied research

18. Roughly a quarter of HEIs, subject bodies and stakeholders agree that research assessment should be more representative of applied research. This support derives from a perception that the RAE has been far too ambiguous about the value of applied research, particularly that which relates to links with industry; and to professional practice in social work, clinical medicine and other community-based health and social research.

Research of relevance and utility

19. A third of HEIs agree that research of direct relevance and utility should attract more credit in research assessment. Support here is mainly predicated on the perception that the premium international excellence category in the RAE has discriminated against research that is focused on national, regional or local challenges, because by its definition this research fails to compare to international standards.

Research training

20. About a fifth of HEIs, subject bodies and stakeholders argue that the training of research students should be incorporated in research assessment to sustain and safeguard the UK's eminence in research.

Interface with teaching

21. Roughly a quarter of HEIs, subject bodies and stakeholders support broadening the parameters to embrace research that develops either the pedagogy or teaching subject matter in any given discipline. These respondents argue that the RAE has:

- a. Neglected, and thus devalued, pedagogical research by 'hiving it off to the Education panel for consideration, rather than assessing it within its parent subject panel.
- b. Encouraged more academics to focus on research at the expense of teaching quality (and the production of textbooks), by operating a rewards-based research assessment process in the absence of a parallel process for teaching.

This is perceived to have driven wedges between teaching and research, jeopardising the fulfilment of government policies in both areas.

22. Most respondents expressing a preference agree that all HEIs should continue to be assessed in the same way.

- a. About half of all responses from HEIs agree that the higher education funding bodies should operate the same assessment process for all institutions.
- b. Among subject bodies and stakeholders, of those that make a clear statement on this issue, a majority also support a single assessment process.

23. Most responses that express a preference agree that all HEIs should continue to be assessed in the same way. This is driven by a general perception that research assessment should have at its heart some absolute benchmarks equally applicable to all participating institutions, and that the results should be broadly consistent and comparable across the entire sector. However, there is also strong support within these responses for a single system to be sufficiently flexible to take account of institutional ethos, mission, resources and patterns of development in arriving at a final grade for any given unit of assessment. In other words, according to the University of Glasgow: 'All institutions should be assessed in the same process, but not necessarily in the same way.'

24. It is important to note that much of this support for a single assessment process is predicated on the assumption that institutions at the lower end of the scale have some expectation of funding.

25. There is a consensus that the assessment process should be flexible but comparable for different disciplines.

26. Discussion of whether each subject should be assessed in the same way reflects many of the concerns surrounding consistency versus flexibility apparent in the previous section. Most responses fall somewhere between these two views, coalescing around a compromise between the need to accommodate differences while ensuring that a given grade in one discipline is broadly comparable to the same grade in another.

27. There is support for the elimination of any incentive to tactically manipulate the proportion of staff submitted to research assessment.

- a. About a quarter of responses discuss the issue of staff submission to research assessment. Of these, two-thirds advocate the submission of all staff, mainly to eliminate the incentive for games-playing.
- b. Support for the submission of all staff is strongest among HEIs and subject bodies.

28. It is clear that the exclusion of particular members of staff in RAE submissions, aimed at securing the highest possible amount of funding, is regarded as the most unpleasant type of games-playing promoted by the exercise. Roughly two-thirds of responses making a clear statement on the matter argue that this phenomenon should be eliminated by enforcing the submission of all research active staff (with allowances made for disproportionately high teaching, administrative or clinical roles). This is also seen as required to vindicate the UK's claims to international excellence.

29. However, a significant minority of responses oppose the submission of all staff. Reasons given include:

- a. Submission of all staff will disguise pockets of excellent research within institutions that are primarily concerned with other missions.
- b. Since institutions have to live with the funding consequences, it should be left to institutions to decide how many staff to submit in order to maximise income.
- c. Submission of all staff will lead to an unmanageable administrative load.

30. There is modest support for research assessment to address equality issues for young researchers and women.

- a. About 1 in 7 responses agree that the RAE has mitigated against the appointment and submission of young researchers in UK HE.
- b. Around 1 in 15 state that the RAE has discriminated against the appointment and submission of female researchers.

31. About 1 in 7 responses agree that the RAE has mitigated against the appointment and submission of young researchers, because they are less likely than older colleagues to have assembled a strong list of published outputs. Most responses go little way beyond making the point and asserting that young researchers should be protected. There are a few suggestions for remedial action, such as including young researchers in the volume measure but not in determining the grade; and making the development of young researchers count in the calculation of a unit's final grade.

32. Around 1 in 15 responses state that the RAE has discriminated against the appointment and submission of female researchers, mainly because the demands of four published outputs are unrealistic for women if they have responsibility for young families. As is the case for young researchers, most comments on women in research assessment tend simply to assert the role of the RAE in exacerbating the problem, rather than suggest solutions. Suggestions for action include: balancing gender on the assessment panels; establishing a national code of practice on people and research, incorporating the treatment of women; and setting a longer cycle of assessment to accommodate those taking career breaks.

33. Only one submission mentions the issue of researchers from ethnic minorities at any length. This response argues that the RAE has in fact had a positive impact on the career prospects

of ethnic minority researchers, since it shifts decision making about research quality (which are of course linked to career progression) away from closed local communities to the national arena. Of course this low response rate in respect of ethnic minorities may simply be a function of the type of audience receiving the call for evidence (as indeed it may be for young researchers, women and other groups).

34. More generally, several responses comment that the RAE is discriminatory in that it does not reward institutions, disciplines and types of research that tend to engage a relatively high proportion of young, female or ethnic minority researchers.

Have we missed anything?

35. Of the other comments noted, most focus on funding issues generated by the 2001 RAE financial settlement.

36. There is strong support among HEIs in particular for the next round of research assessment to make the financial outcome of attaining a particular grade explicit before the exercise is run. This position is echoed by both research intensive HEIs and institutions without a strong record of research, as well as sector-wide bodies such as Universities UK. This reflects widespread disquiet about the ambiguous relationship between assessment and funding allocations.

37. Many of these responses also call for a longer or continuous grading scale with smaller funding gaps between each step, to ease some of the financial pressures associated with the exercise.

Annex F

Assessing research: the researchers' view

Summary of a report by Dr Steven Wooding and Dr Jonathan Grant of RAND Europe

Introduction

1. This summary presents the key findings from a series of nine facilitated workshops with academics and research managers, which examined how research in UK higher education institutions (HEIs) could be assessed.
2. The full report, 'Assessing research: the researchers' view', is split into two volumes and can be obtained from the websites of both the Higher Education Funding Council for England (www.hefce.ac.uk) and RAND Europe (www.randeurope.org).

Methodology

3. Facilitated workshops were used to stimulate broad and innovative thinking about research assessment, while at the same time providing a structure that would allow comparison between workshop groups. As with all qualitative research methodologies, the facilitated workshop provides a method of gaining an insight into attitudes and opinions; however, this insight is – by its very nature – not necessarily representative of the population sampled in a statistical sense. It is also possible that even the relatively flexible structure used may have constrained the views participants were able to express.
4. HEIs were approached by regional representatives of the UK higher education funding bodies, and asked to nominate participants for the workshops. The workshops were attended by 142 participants. These participants represented 60 out of the 172 HEIs, and 42 of the 68 subject-based units of assessment submitted in RAE2001. Around a quarter of the participants were administrative research managers. Possibly due to the recruiting methodology, senior academics were over-represented among the participants, relative to the more junior staff.
5. The workshop was structured around a number of tasks, with the views of participants collected as flip chart sheets produced by participants and used in feedback to the workshop group.
6. In the first task participants worked in pairs, to consider what characteristics mark out high quality research and what characteristics are important in research assessment systems. All the characteristics suggested were recorded, and then prioritised by the participants using a system of multi-voting. This prioritised list provided a context for the work in the remainder of the workshop.
7. In Task 2 small groups of participants were allocated two of the four approaches to research assessment laid out in the joint UK funding bodies 'Invitation to contribute': expert review, algorithms, historical ratings and self-assessment. The groups were then asked to suggest the strengths or weaknesses of their approaches, and the questions they would want answered if that system were to be implemented. This exercise revealed how the participants thought about each of the approaches, and made them aware of the range of possibilities for research assessment.

8. In the remaining two tasks, shuffled groups of participants were asked to design their ideal research assessment system, basing it on one of the approaches examined in Task 2. The participants were then asked to consider how their system could be implemented, what its weak points might be and how they hoped its use would change research culture in UK higher education.

Analysis

9. Our analysis sought to draw out recurring themes from across the workshops. For Tasks 1 and 2 this was done by grouping the participants' suggestions into related clusters. For Tasks 3 and 4 the 29 systems designed by participants were compared to extract common design elements.

Key findings

Peer review

10. The overwhelming majority of the academics and research managers who took part in this study felt that research should be assessed using a system based on peer review by subject-based panels – of the 29 systems designed, 25 were based on expert review. The participants also indicated that these panels should be informed by metrics and self-assessment, with some input from the users of research.

Transparency, stability and professionalism

11. There was a very strong desire for a system with clear rules, and transparent procedures, that were established at the outset and not modified during the assessment process. The appointment of panels and the selection of the criteria they used were thought to be critical areas for transparency. Participants in the study considered that the panels themselves should be professionalised and that there should be increased and earlier involvement of international members. They suggested that panel chairs from outside the subject area with more experience of facilitation should be used, and that these chairs might be paid.

Clarity of submission

12. Almost half the groups were unhappy with the flexibility and lack of clarity in the current assessment system over which staff should be submitted, and one-third of the groups felt that more staff should be submitted in future. In addition to reducing the scope for 'playing the system' it was felt that submission of more staff would improve the inclusiveness of the process. A few groups included other steps to make the process more inclusive and sensitive, both to researchers and to a lesser extent to institutions.

Unit breadth and interdisciplinary research

13. Almost half of the groups suggested that units of assessment should be broadened and reduced in number, with many hoping that this would help the assessment of interdisciplinary, multidisciplinary and transdisciplinary research. Other mechanisms for improving assessment of interdisciplinary research were suggested, including allowing panels to call on – or second – external expertise.

Frequency

14. Around half of the groups who addressed the issue of frequency recommended that the research assessment cycle should be extended. But in order to retain dynamism some suggested a light-touch 'interim' assessment should be added at the half-way point.

Agreement between disciplines

15. The most important characteristics of high quality research were seen as rigour, international recognition, originality, and the idea that the best research sets the agenda for new fields of investigation. There was general agreement about the importance of these characteristics by participants from different disciplines and academic roles – although absolute ranking varied.

16. There was also broad agreement across disciplines about the most important characteristics for a research assessment system. However, researchers from medicine, science and engineering placed a greater importance on peer review, while their colleagues in the arts, humanities and social sciences felt subject-related flexibility in the assessment system was more important.

Comparability and flexibility

17. Participants ranked both comparability of assessments between subjects and methodological sensitivity to the subject being assessed very highly when considering characteristics of research assessment systems. Despite this, when designing research assessment systems, almost half the groups suggested that panels should be given more autonomy in developing their criteria and assessment methods.

Acceptance of burden

18. Although most participants were keen to avoid a system that was onerous, they appreciated that any system capable of providing the necessary fairness would be relatively time and labour intensive. Given this realisation, it was felt that the system should provide more useful feedback to the participants to help them improve and develop their research.

Communication

19. Listening to participant discussions about the current research assessment system, it became clear that whatever new system is adopted, the funding councils will need to put in place programmes to engage the academics in the system's development and explain its final structure and processes.

Annex G

Changes in research assessment practices in other countries since 1999

A report by N von Tunzelmann and E Kraemer Mbula of SPRU (Science and Technology Policy Research) at the University of Sussex

Introduction

1. As requested by HEFCE, we are reporting on key changes in research assessment practices in a selected group of countries where we have prior evidence of substantive change since 1999, when SPRU's previous report surveyed the general picture²⁸. Within the time and budget available, we have not been able to produce an exhaustive account of all changes taking place in all countries regarding research assessment. However, we feel that the material contained in this report gives a fair picture of the extent and direction of change that has taken place over these three and a half years since the 1999 report.

2. For the present report we were also requested to obtain material where possible on several new issues currently on the UK RAE agenda; an overview of what we have found appears in the summary below. The full report, including details on specific countries, is on the web at www.ra-review.ac.uk under Research reports.

Summary overview

3. The 1999 report developed a methodology for considering various types of research assessment. Specifically, the results for individual countries were grouped into:

- a. Evaluation performer (national level and institutional level), and evaluation purposes (summative or formative).
- b. Evaluation criteria (quality, impact, etc) and methods.
- c. Evaluative remarks (strengths and weaknesses, impact on research and teaching).

4. Most attention was then paid to the first two points. For this present report on the changes since 1999 we have found few instances of change in regard to evaluation performer under paragraph 3a. We do however get a sense of a switch of emphasis from summative towards formative purposes of evaluation. Or more precisely, perhaps, it could be claimed that jurisdictions embarking on more formal research assessment exercises are giving greater weight to formative functions than they envisage as existing in the UK RAE (which is often cited as a 'benchmark'). Examples below include Ireland, and New Zealand. The chief justification for more formative measures appears to be giving a greater degree of individual 'ownership' of performance improvement, for example to make more distinctive choices of a strategy for achieving improvement.

5. A possible exception is the case of Taiwan, where a rather different situation seems to apply. Taiwanese studies allege that the country's industry is very strong by international standards, for example in patenting, but its academic publication record is a lot weaker (Luo, 2002). This is

²⁸ Resource Allocation and Research Performance: The Assessment of Research; Geuna et al. 1999

arguably the reverse of the UK situation. The great merit of the UK RAE is seen in Taiwan as its boosting of the competitive spirit in universities (ibid). We have received some rather contradictory messages from correspondents in Taiwan about what is actually happening there, and our judgements may thus be incorrect.

6. As in the 1999 report, we cannot claim to have a great deal of evidence to bring forward regarding paragraph 3c, evaluative remarks. Most countries still regard their systems as somewhat experimental. The most obvious inference we can make is that the nature of changes introduced is a likely reflection of dissatisfaction about previous systems (or indeed the previous absence of systems).

7. Our discussion in the present report therefore has more to say concerning paragraph 3b, the evaluation criteria and methods. A partial critique of existing methods was given in the 1999 report. The UK RAE is frequently criticised for not using a broader basis of bibliometric indicators. Objections to such broadening of the indicators in the UK are often made on what to us look somewhat specious grounds. But we have to admit that the only case we know of in which a more thoroughgoing adoption of formal bibliometric criteria has been made, namely Flanders in Belgium, does raise serious doubts about the practicality of implementing such methods in a country such as the UK. Again, the experiment is too recent to be properly assessed.

8. In the event, our present work is directed to an issue which falls under several of the above headings, which can be referred to as evaluation governance. In line with standard definitions of 'governance' (as comprehending structure, control and process), this includes the structures of decision-making (which relates to paragraph 3a on performers) but also processes (which relates more to paragraph 3b on methods).

9. The 1999 report drew attention to four categories of countries in regard to university research funding practices:

a. The first used a performance-based approach like that in the UK, using research evaluations to distribute funds at least in part. Poland, Slovakia, Hong Kong and Australia were included in this category.

b. The second comprised countries that used an indicator other than research evaluation (though sometimes with a small portion of this), generally student numbers or similar. This was the largest group in the 1999 report (Germany, Italy, the Nordic countries, Hungary and New Zealand).

c. Third was a small group of two countries (Austria and France) where funding allocations were described as open to 'negotiation'.

d. Fourth was a group of three, the Netherlands, the USA and Canada, where research assessment and funding were separated.

10. The countries included in the present study do not fully overlap with those examined in 1999 (for example we have not looked at Eastern Europe this time), but it seems reasonable to assert that there has been some shift towards the first category. The Netherlands and New Zealand have both made explicit moves in this direction, though in both cases during the course of our enquiry (in

January 2003 and December 2002 respectively). Other countries such as Taiwan, not considered in 1999, seem to be moving in the same direction. On the other hand, the North American countries have not really budged insofar as university research is concerned, and Italy too has seemingly not changed.

11. Some of the other key conclusions from our study are the following.

a. A general conclusion is that countries such as Ireland, France, Switzerland, Denmark, Japan and New Zealand are intending to rely to a greater degree on self-evaluation, subject to oversight panels. (The USA does the same in its benchmarking of federal research agencies.) This is intended mainly as a way of sharing the burden between centralised and decentralised agencies, and perhaps giving a more rounded picture. It is not transparently clear how much this would differ from the submissions made by universities as the first stage of the existing UK RAE, but certainly the countries concerned appear to feel it is indeed different. From this we infer a shift to a more formative role for research assessment in such countries. All countries introducing such schemes appear to be aware of the probable 'puffery' of self-evaluations, which the oversight panels are intended to bear down upon, but the same could be said for the existing UK system. Self-evaluations may need to be linked to more objective measures of performance benchmarking if they are to have any real standing (see below). In favour of self-evaluation is said to be the opportunity for making special anti-discriminatory cases, for example in relation to the treatment of women (see below).

b. While there is widespread interest in the UK RAE, intensive evaluation schemes are mainly limited to smaller countries or regions, and indeed some of our correspondents have stated to us that they would not be feasible in larger countries. This suggests that the UK has problems of intensive application in a comparatively large country, as of course has often been expressed in relation to the resource cost involved. This perhaps suggests that the practicability of a more devolved system (possibly associated with the recent emphasis on Regional Development Agencies) might be investigated. Taking this one stage further, we might suggest coupling 'intensive' evaluation with the previous point about self-evaluation, that is, getting universities to produce their own benchmarking using bibliometric or other methods.

c. We would also draw attention to the basic theoretical underpinnings of any research assessment, as noted in relation to Ireland. This overlaps with concerns expressed in some quarters that the RAE is too narrowly academic, though it goes rather beyond that to deeper issues of the 'contribution of basic research to the modern economy'.

d. Finally we could mention that most of the evaluations occur with longer gaps than has been the past practice of the UK RAE.

12. We were also asked to report on any evidence located with respect to four supplementary issues (again we have been unable to provide an exhaustive analysis):

- use of benchmarking practices
- special treatment of the social sciences and humanities
- special reference to the treatment of women
- explicit use of the UK RAE model and the UK debate.

Use of benchmarking practices

13. Benchmarking is attracting growing interest, though mostly independently of university research assessment. (See evidence below for Ireland and the USA, with the US adoption of the 'virtual congress' worthy of note.) However, benchmarking covers many different approaches at many different levels, and a more concerted study would be needed to draw any strong conclusions. We make the point above that, if benchmarking is to be adopted for assessing individual universities and colleges, it might be necessary as a practical matter to devolve the responsibility to them for furnishing the data, since all known benchmarks suffer from a variety of limitations (see the discussion of Flanders). There is also some concern about over-use of benchmarking as encouraging 'moral hazard', for example, an undue focus on attaining the specific benchmarked targets.

Special treatment of the social sciences and humanities

14. We found relatively little reference to this in research evaluations, though some countries have weighting systems (see evidence for Ireland, Flanders, Japan and New Zealand, also references to Hong Kong and Australia), while many express general concerns about perceived imbalances in the educational system or student uptake.

Special reference to the treatment of women

15. This area in a general sense has been recently overviewed by the European Commission, particularly through the Helsinki Group on women and science (1999). This group has produced national reports on the situation of women scientists in EU countries, and represents an important base to compare different experiences across countries. From these and other sources we include some remarks on individual countries. What we found less on were the specific allegations of implicit discrimination through the use of research assessment itself, although it does appear to be widely accepted that this occurs (see the note for New Zealand).

Explicit use of the UK RAE model and the UK debate

16. As stated above, there is clearly near-worldwide interest in the UK model, which itself has become a benchmark for research evaluation of higher education. However it has not yet been precisely replicated elsewhere, and most countries are probably holding a watching brief (see the evidence for Ireland, Australia and New Zealand). We have learnt recently of a proposed implementation of a UK-style RAE in Taiwan, although as implied above we are getting somewhat mixed messages about how far this has progressed, and indeed what its precise form will be.

Annex H

Additional costs implicit in the recommendations

Administration costs

1. The main report sets out the reforms we believe are necessary to run a subject-based peer review system similar to the RAE – our proposed Research Quality Assessment (RQA). We would not hesitate to make any of these recommendations were our preference to retain the RAE in its present form²⁹. There are, however, aspects of our recommendations which may necessitate additional activity. We touch on these below.

International review

2. The costs of international review as recommended above will be non-trivial. We must assume that flights, accommodation and expert time will have to be funded at commercial rates. On this basis, we assume that the cost would be not less than £5,000 per reviewer. If we assume 25 units of assessment each requiring 16 reviewers, this yields a cost of £2 million.

Competence assessment/institutional level assessment

3. We envisage that these would be undertaken by the funding councils using their existing mechanisms of audit and institutional monitoring³⁰. There are arguments for and against embedding the competence assessment within the RAE administration. On the downside, it would carry with it the risk of creating a new assessment bureaucracy which might suffer from ‘mission creep’. On the other hand, it is more compatible with the employment of a permanent RAE staff, which would guard against the discontinuities that have blighted previous exercises and their aftermath.

4. It is essential that these tasks are not devolved to the funding councils as a means of hiding the cost. If they are to be absorbed within existing functions, the additional costs to the funding councils must be properly costed and budgeted for.

Mid-period evaluation of the progress of emerging units

5. We have proposed that institutions be allowed to nominate a minimal number of emerging units, whose progress would be assessed at the mid-point of the assessment period. This would need to be co-ordinated centrally and would impose extra costs.

Collation of metrics

6. Responsibility for collation of metrics ought to sit within the RAE administration. A permanent statistical capacity would need to be embedded within the team or sourced from elsewhere. In the past the requirement for statistical expertise has been met from within HEFCE’s Analytical Services

²⁹ Although the role of the five chief moderators would be less demanding.

³⁰ It may well be that the appraisal of the institution’s arrangements for the ‘development of researchers’ are the subject of a separate process arising out of the implementation of the recommendations of the Joint Funding Councils’ Review of Research Degree Programmes.

Group. However, HEFCE's ability to provide this service has been dependent upon a level of staff continuity within that group which cannot be presumed for the future. What is more, our proposals will tend to increase the requirement for statistical support.

Panel workload

7. The principal determinant of panel workload is the amount of work panels are asked to review: the assessment load.
8. Moving from a quinquennial to a sexennial cycle would produce an instant 20% reduction in the assessment load.
9. The amount of work reviewed through RQA would depend upon the policies of the funding councils. If the councils provided sufficient incentives for large numbers of institutions to opt for one of the other forms of assessment (or if they obliged them to do so) the assessment load would fall substantially.
10. Over half (55%) of the staff in RAE2001 were in submissions graded 5 or 5*. A further 25% were in submissions graded 4. If we assume that the funding councils will wish to fund (and therefore assess) the equivalent of all the 5 and 5* rated submissions and half of the 4 rated submissions; and if we further assume that, because the results will not be known in advance, some of those who would fall outside these categories will also need to be assessed, it seems unlikely that reductions in the assessment load of greater than 20% would be achievable³¹.
11. Taking into account the saving created by the move to a six-year cycle we estimate that there would be a total reduction in the assessment load of around 35% to offset against the increased administrative costs described in the previous section.

Mid-period evaluation of the progress of emerging units

12. We have proposed that institutions be allowed to nominate a minimal number of emerging units, whose progress would be assessed at the mid-point of the assessment period.
13. In RAE2001, there were 2,598 submissions. Of these, 664 were graded 4. If half that number of departments were badged as emerging units, there would be 332 evaluations to undertake at the mid-point of the process. This would be a significant task.
14. We therefore propose that those identifying emerging units should set objective targets, to be met within three years. The RQA panels would have the task of certifying that achievement of these targets would represent progress in catching up with leading units. If they were not satisfied, or if the targets were not readily measurable, this would constitute grounds for the panels to refuse to identify

³¹ If these assumptions were revisited, greater reductions could be achieved, particularly given the move from grades to a quality profile. If for example the funding councils decided to target the research capacity identified as of international quality in the last RAE, only 10-50% of the capacity in 5 rated departments would qualify. Nevertheless, we consider it unlikely that they would wish to take such a radical course.

the unit as an emerging unit. The monitoring of progress against these targets should be an automatic process.

15. Nevertheless, panels would have to certify both that targets set by institutions were measurable, and that their attainment would indicate that the unit was catching up with the leaders in its field. This would be a major undertaking, complicating the role and increasing the workload of the panel.

Institutional workload

Tactics, strategy and games-playing

16. Institutional costs are the hardest to estimate. The formal data requirements of the old RAE were relatively modest, but the effort invested by institutions in planning and constructing their submissions was vast. This reflects two factors: the scope for tactical decision making, and the amount at stake in terms of reputation and funding. This review's proposals are relevant only to the first of these; the second will continue to ensure that the effort which institutions put into their submissions is disproportionate to the complexity of the task. It would be naive to suggest that either the institutions or the process are to blame for this.

17. For those institutions participating in RQA we assume, cautiously, that the burden would be similar to the RAE. Instead of taking decisions on who to submit as 'research active' at the level of the individual, we hope that institutions would take decisions at the level of the department or faculty – which would require fewer judgements and therefore less burden.

18. However, initial discussion with institutions suggest that they would go to considerable lengths to ensure they were able to submit all their strongest research staff to the RQA, even if this necessitated merging departments or 'hiding' staff in units which the institution considered weaker.

19. This need not be a problem so long as there remains an element of risk in the system. There must however be an effective means of verifying that the allocation of staff to sub-units of assessment by the institution reflects the way in which the institution organises its activities. If this does not happen, there is a risk that staff who should have been excluded from the Research Capacity Assessment (RCA) on the basis that their colleagues were submitted to the main RQA will attract funds to the institution.

20. There are three approaches the funding councils could take to this kind of games-playing:

- a. attempt to enforce the requirement to submit 80% of staff through rigorous audit
- b. accept a degree of games-playing whilst attempting to identify the most blatant examples
- c. abandon the requirement to submit 80% of staff in a unit to RQA on the grounds that the regulation would give rise to unacceptable distortions of behaviour.

Research competences assessment

21. Three of the proposed strands of the research competences assessment concern items on which the funding councils would need reassurance, whether through a process linked to research

assessment or through some other process. Given the levels of public investment in research, it is inconceivable that the councils would not expect institutions to have credible policies and processes to guarantee equal opportunities and the development of researchers, and to have a research strategy.

22. The key here will be to avoid duplication. Our strong preference would be for the research competences assessment to depend wherever possible upon processes already in existence or currently under development, including the funding councils' own routine monitoring processes. There would therefore be little assessment as such, merely a check to ensure that other processes had given the institution a clean bill of health.

23. The fourth strand – dissemination beyond the academic peer group – is an innovation. However, we do not consider that it is unreasonable to expect institutions to put together a short statement outlining their policies on dissemination, nor that a paper-based review of such statements would constitute an appreciable addition to the assessment burden.

Collation of performance indicators for RCA and mid-term volume monitoring

24. The RCA would rest principally upon discipline-specific performance indicators. It will be necessary to consult with subject communities before it is possible to identify these indicators. However, it is fair to assume that, if the data collection was undertaken using traditional paper-based methods the greater variety of data required would lead to a slightly increased burden (though we do not believe that formal data requirements are responsible for more than a fraction of the institutional costs of the RAE).

25. We are investigating the possibility of using a centrally maintained electronic data store to source the information required. Under such a scheme, institutions (or even individual academics) would be in a position to update their own entries on an ongoing basis. This information would, subject to audit, produce the data required for the performance indicators.

Institutional level assessment

26. There is always a risk that any process involving institutional level assessment escalates to the point where the activity on both sides is far greater than is justifiable given the aims of the process. For this reason, we suggest that institutional level assessment be undertaken within the context of the funding councils' normal interaction with their institutions. This would prevent institutions from anticipating a more intensive process than is actually envisaged.

Emerging units

27. It is probable that institutions will invest a large amount of time and effort in identifying their emerging units, outlining their plans for developing them, setting progress targets and ensuring those targets are met. We recommend that the funding councils be guided by the sector in determining whether the undoubted costs are balanced by the likely benefits.

Panel members

28. The estimated 35% reduction in the assessment load (see paragraph 11 of this annex) would have consequent benefits for institutions that employ panel members.

Annex I

Glossary of terms

Dual support	The system of public funding for research in the UK, in which funding is provided through two sources – for infrastructure by the UK higher education funding bodies and for projects by the research councils and the Arts and Humanities Research Board
FTE	Full-time equivalent
Funding bodies	The four UK higher education funding bodies: Higher Education Funding Council for England (HEFCE); Scottish Higher Education Funding Council (SHEFC); Higher Education Funding Council for Wales (HEFCW); and for Northern Ireland, the Department for Employment and Learning (DEL)
HEI	Higher education institution
PGR	Postgraduate research student
QR	Quality-related research funding: allocations are determined by the results of the RAE
RAE	Research Assessment Exercise
RCA	Research Capacity Assessment (proposal)
Research councils	There are six subject-specific research councils. They are funded by the Government through the Office of Science and Technology to support research in their own establishments and to fund research projects in universities. Research is also funded by the Arts and Humanities Research Board
RQA	Research Quality Assessment (proposal)
SCOP	Standing Conference of Principals, the representative body for principals and directors of colleges and institutes of higher education
Unit	Term used in this document to refer to the research activity represented by a submission to the RAE. In some cases, a unit may be a university department, although in many cases it will cross departmental boundaries
Universities UK	Representative body for heads of universities (formerly the Committee of Vice-Chancellors and Principals)