

UNIVERSITAS

Higher Education Management Consultants

PO Box 5060 Milton Keynes MK6 3ZX United Kingdom

OPERATIONAL REVIEW OF THE RESEARCH ASSESSMENT EXERCISE 2001

**REPORT TO THE JOINT FUNDING BODIES'
RESEARCH ASSESSMENT REVIEW**

April 2003

Consultants:

John Farrant (Project Manager)

David Billing

Paul Temple

CONTENTS

1. Conclusions and summary	1
1.1. Terms of reference and method	1
1.2. General conclusions	1
1.3. The central machinery of the RAE	2
1.4. The framework for assessment	4
1.5. Preparation of submissions	4
1.6. The Panels	5
1.7. The Assessment Process	6
1.8. After the assessment	8
2. Abbreviations	11
3. The Research Assessment Exercise 2001	12
4. Our approach	14
4.1. Terms of reference	14
4.2. The evidence assembled	14
4.3. The questionnaire to Panel Chairs	15
4.4. The questionnaire to institutional Contacts	18
5. The central machinery of the RAE	21
5.1. The Project Group	21
5.2. The RAE Team	22
5.3. The Panel Secretaries	23
5.4. Impact on the Funding Councils' Executives	24
5.5. Risk assessment	24
6. The framework for assessment	26
6.1. Units of Assessment	26
6.2. Rating scale and grade definitions	26
6.3. Statements of criteria and working methods	29
7. Preparation of submissions	32
7.1. Submissions by institutions	32
7.2. Guidance on submissions, etc.	32
7.3. Data requirements	33
7.4. RAE Data Application	33
7.5. Standard audit and verification	35
8. The Panels	36
8.1. Panel membership	36
8.2. Representation of research users	37

8.3. Sub-panels	37
8.4. Workload	38
8.5. Panels' differential loading	39
8.6. Confidentiality and declarations of interest	40
<i>9. The assessment process</i>	<i>41</i>
9.1. Communications	41
9.2. Timetable	41
9.3. Adequacy of the data collected	42
9.4. Presentation of data	43
9.5. Ad hoc statistical analyses	43
9.6. Audit queries	43
9.7. Selection of staff for submission	44
9.8. Use of RA5 and 6	44
9.9. Panel proceedings	46
9.10. Arrangements for panel meetings	46
9.11. Research Output Sourcing System	47
9.12. Specialist advisers	48
9.13. Cross-referrals	48
9.14. Non-UK-based advisers	50
9.15. Umbrella Groups	51
9.16. Interdisciplinary research	52
9.17. Collaborative research	54
9.18. Research into teaching and learning	55
<i>10. After the assessment</i>	<i>57</i>
10.1. Overview Reports	57
10.2. Feedback to institutions	57
10.3. Complaints about the outcome	58
10.4. Matching the key principles: institutions' assessment	58
<i>Annex 1 Email and questionnaire of 13 January 2003 to Panel Chairs</i>	<i>60</i>
<i>Annex 2 Email and questionnaire of 14 January 2003 to RAE Contacts</i>	<i>65</i>
<i>Annex 3 Summary and recommendations from Data collection for the Research Assessment Exercise 2001: A review by Ewan S. Page (2001)</i>	<i>70</i>

1. Conclusions and summary

1.1. Terms of reference and method

1. The objectives of this study, commissioned by the Research Assessment Review Team, are to:
 1. provide the Steering Group for the Research Assessment Review with a formative and summative evaluation of the 2001 RAE by SWOT analysis. This evaluation should be focussed 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;
 2. 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;
 3. 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 RAR 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.

1.2. General conclusions

3. The Research Assessment Exercise 2001 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. Three factors together risked a major disruption of the assessment phase, though fortunately none occurred:
 - a timetable for the assessment phase even tighter than in 1996,
 - the workload on key players (Panel members and Secretaries, RAE Team) and
 - a demand higher than predictable for the services supporting the panels' decision making.

A significant contributory factor to the second and third of these was that panels scrutinised 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 buy in 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 principle 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 the details of the criteria and working methods, amongst panels in cognate areas; and
- the substantial variation as 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. In addition to those in the previous paragraph, the following are those which operated less than satisfactorily or had less than satisfactory outcomes, and which should command priority for close examination if they are to be incorporated in any future assessment exercise:

- 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

1.3. The central machinery of the RAE

1.3.1. The Project Group

7. The Project Group of senior officers of each funding body embodied a considerable body 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.

1.3.2. 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. For both 1996 and 2001 exercises HEFCE's Analytical Services Group was contracted to undertake the data collection and analysis - and by contrast and with benefit deployed the same (permanent) staff into the key roles on both occasions.

9. The view among Panel Chairs is that the Team was over-stretched and understaffed during the assessment phase, which contributed 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 panels generated far greater volumes of requests for the Team's support than past experience suggested or than 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 (e.g. cross-referrals), reviewing the configuration of UoAs, 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.

1.3.3. 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 less than in 1996 because certain functions (e.g. booking accommodation, payment of expenses) were placed with the RAE Team in 2001. Because the RAE Team was housed 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.

1.4. The framework for assessment

1.4.1. Rating scale and grade definitions

13. Although the majority of Panel Chairs found the grading system readily workable, many were concerned at 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.

1.4.2. 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 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.

1.5. Preparation of submissions

1.5.1. Guidance on submissions, etc.

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 (on-line) Guidance document. The most frequently mentioned area of difficulty was the categorization 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.

1.5.2. 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 half satisfied.

1.5.3. 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 Manager's conclusion is that the data contained in RAE submissions was 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.

1.6. The Panels

1.6.1. 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.

1.6.2. 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.

1.6.3. 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.

1.6.4. 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.

1.6.5. 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 UoAs, with more differential allocation of resources in support of panels, or with some merging of panels supported by sub-panels.

1.6.6. 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.

1.7. The Assessment Process

1.7.1. Communications

27. On grounds of security, Panel members were denied the submissions in any format than paper and were prohibited from using email 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.

1.7.2. 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), of referrals to specialist advisers, of cross-referrals between panels, of work in preparing workbooks, etc. 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.

1.7.3. Data collection and analysis

29. Panel Chairs seem to have been broadly satisfied with the range of data collected in RA1 to 4 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 rekeyed data and created their own databases.

1.7.4. 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.

1.7.5. 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 UoAs. 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.

1.7.6. Use of RA5 and 6

32. Panel Chairs were split between those for whom RA5 and RA6 were important or essential etc, and those for whom they were 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.

1.7.7. 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.

1.7.8. Arrangements for panel meetings

34. Many panels reported that the quality of the accommodation for meetings, 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.

1.7.9. 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 Bristol, and items returned damaged or not at all. Some Panel Chairs and Contacts suggest that for another RAE institutions should deposited in one location a copy of each cited output, or that all outputs should be available electronically.

1.7.10. 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.

1.7.11. Cross-referrals

37. The system for cross-referrals between panels attracted many negative comments. The mandatory cross-referrals were relatively straight-forward, though perhaps should not have been 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 component is likely to be a part of a future RAE with a similar configuration of UoAs, there should be an analysis of the flows of cross-referrals, as it may pinpoint some sub-optimal aspects in the configuration of UoAs and in the composition of individual panels.

1.7.12. Non-UK-based advisers

39. Chairs' comments were mostly negative: the non-UK-based experts were not properly briefed, had too little understanding of the criteria and process, were often out-of-date on 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 individual or departmental reputations. Nevertheless, in the Overview Reports panels were pleased to note that the non-UK-based advisers validated the 5 and 5* gradings.

1.7.13. 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.

1.7.14. 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 panels 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, with more Contacts considering 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 *Ea* 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 a contested area.

1.7.15. 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.

1.8. After the assessment

1.8.1. 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.

1.8.2. 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.

1.8.3. 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 Northavon House for handling these complaints.

1.8.4. 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:

RAE 2/99, para. 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.

2. Abbreviations

ASG	Analytical Services Group at HEFCE
DENI	Department of Education for Northern Ireland
HEFCE	Higher Education Funding Council for England
HEFCW	Higher Education Funding Council for Wales
HEI	higher education institution
RAE	Research Assessment Exercise
RAR	Research Assessment Review
RDA	RAE Data Application
Rogers	J. M. Rogers, 'The Next Research Assessment Exercise', discussion paper for the Project Group, version 3, March 2001
ROSS	Research Output Sourcing System
SHEFC	Scottish Higher Education Funding Council
UoA	Unit of Assessment

3. The Research Assessment Exercise 2001

48. The primary purpose of the 2001 Research Assessment Exercise was to produce ratings of research quality which would be used by the higher education funding bodies in determining the main grant for research to the institutions they funded, with effect from the financial year 2002/3. The funding bodies may also use the exercise to inform policy development. The RAE was conducted jointly by the Higher Education Funding Council for England (HEFCE), the Scottish Higher Education Funding Council (SHEFC), the Higher Education Funding Council for Wales (HEFCW) and the Department of Education for Northern Ireland (DENI). The exercise was managed by the RAE Team, on behalf of the four funding bodies. The Team was based at the HEFCE offices in Bristol, giving rise to the erroneous perception that HEFCE conducted the RAE. All UK HEIs receiving funding from these bodies were eligible to participate.

49. The RAE 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. Few complaints about the ratings have been pressed beyond one letter. We say this at the outset because it is in the nature of the review which we have been asked to conduct, that we should be searching out those parts of the operation which might have worked better. Having vicariously experienced the processes of the RAE through the evidence we have reviewed, we applaud the commitment, professionalism and sheer hard work of those who enabled the universities and colleges to make submissions and who ensured that these were scrupulously assessed, above all the Panel members and their Chairs, the Panel Secretaries, the Contacts, the RAE Team and their colleagues in Northavon House. We are also grateful to them for sharing those experiences with us and responding so quickly, fully and interestingly to our enquiries.

50. Readers of this report are assumed to be familiar with the processes of the RAE which are not therefore described here. The fullest published account of the RAE is the 'Manager's Report - May 2002', to be found, along with all the RAE circulars and some other documents, on the RAE website which is now hosted by HERO: <www.hero.ac.uk/rae>. The timetable of the main events in the RAE was as follows:

July 1998	RAE 2001 announced.
November 1998	RAE Manager appointed. Nominations for panel membership invited. Panel Chairs elected.
March 1999	Nominations passed to chairs for recommendations for membership.
May 1999	Guidance on submissions published.
spring 1999	Panels constituted and draft criteria and working methods.
August 1999	Draft criteria and working methods published for consultation.
December 1999	Criteria and working methods published in final form. Membership of panels published.
31 March 2001	Census date for RAE 2001.
30 April 2001	Closing date for submissions.
May 2001	Panels receive submissions.

May – June 2001	Panels' first assessment meeting: Panels appoint rapporteurs for each submission and allocate reading of outputs between members. Panels agree which work to refer to sub-panels, to other panels, and to specialist advisers.
July 2001	Second assessment meeting: Panel members update on assessment work, calibrate marking standards and resolve any assessment issues. Some preliminary grading.
July - early September 2001	Panel members examine research outputs. Specialist advisers consulted. Sub-panel meetings. Cross-referral assessments completed and discussed between panels.
Late August - September 2001	Third assessment meeting (two or three day residential): Panels discuss submissions in detail and agree provisional grades. Submissions for referral to Non-UK-Based Advisers identified.
Mid September – Mid October 2001	Consultation with Non-UK-based Advisers. Further examination of submissions by panels as required. Workbook reports on reasons for grades awarded drafted.
Early October 2001	Umbrella Group meetings Consideration of profiles of provisional grades. Discussion of comparability of assessments, interdisciplinary research areas, and any assessment issues of general applicability.
Mid – Late October 2001	Final assessment meeting: Panels confirm final grades and workbook reports.
31 October 2001	Grades and reports received by RAE Manager.
December 2001	Grades published.
spring 2002	Panels' overview reports published, and feedback to institutions issued.

4. Our approach

4.1. Terms of reference

51. The Research Assessment Review Team's project specification of 22 November 2002, inviting our tender, set out the objectives of this study as to:

1. provide the Steering Group for the Research Assessment Review with a formative and summative evaluation of the 2001 RAE by SWOT analysis. This evaluation should be focussed 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;
2. 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;
3. take into account the comments received from key participants in the 2001 RAE, including Panel Chairs, members, secretaries and institutional Contacts.

We maintained close contact with the study's client, the Research Assessment Review Team, throughout the study, by telephone and e-mail and by an initial meeting in Oxford and two visits to Bristol when we also met HEFCE staff who had been involved in the RAE and read files.

4.2. The evidence assembled

52. The project specification envisaged that the study would be largely desk-based. The documentation immediately available comprised

- documents on the HEFCE, RAE and RAR websites, of particular importance being the general guidance to Panel members and to institutions, the generic criteria, (RAE 2/99), the individual panels' criteria and working methods (RAE 5/99) and the Overview Reports from all but one panel. These have enabled us to understand the intended processes, and to recognise the differences amongst the panels in their criteria and methods.
- 27 submissions to the RAR, identified by the RAR Team, which referred to the conduct of the 2001 RAE (cited as [RAR])
- a note by HEFCE's solicitors for the RAR
- two reports of feedback from HEFCE-based Panel Secretaries.

53. We decided with the RAR Team at an early stage that we should collect new evidence by means of questionnaires to the Panel Chairs and to the institutional RAE Contacts. We also had discussions with Funding Council officers and former members of the RAE Team

54. Shortly before and after sending out the questionnaires, we became aware of several steps which the RAE Team had taken to evaluate the RAE's operation. We understand that further work on evaluation was halted when the RAE Team dispersed and when it became apparent that the rise in grades in 2001 would necessitate a fundamental review of the methodology for research assessment (now in progress as the RAR). The steps, some of which had produced evidence which was made available to us, were:

A. a long paper by the RAE Manager on 'The Next RAE', dated March 2001, before the

- assessment phase but incorporating lessons from experience so far. This paper should be carried forward for future reference and is hereafter cited as ‘Rogers’;
- B. an invitation to panels at their last (October 2001) meeting to prepare not only the Overview Report for publication, but also comments on the Exercise’s operation. Some panels seem to have included those comments in their Overview Report; other recorded comments in their minutes and these we have extracted (but the records of panel meetings held by HEFCE are incomplete); other at least intended to submit comments through their chair or secretaries, but such comments have not reached their panel’s file nor, it seems, a separate file for them;
 - C. an intimation to panel members at their last meeting that individually they would be invited to comment; but no invitation was issued;
 - D. five ‘Operational Review Meetings’ held with Panel Chairs in May 2002, of which a consolidated set of minutes was prepared;
 - E. a report from Professor Ewan Page on data collection and analysis, commissioned by the Analytical Services Group This report should be carried forward for future reference.

55. We have drawn on the evidence in A, B, D and E. We were also given access to (paper) files registered in HEFCE’s central filing system, in series titled ‘RAE 2001’. These helped us to understand the panels’ working methods, but otherwise they proved limited in their coverage. We understand that most operational documents were retained only in electronic form on a HEFCE server, but that they have not been systematically listed and archived, and that no one currently at HEFCE is familiar with their arrangement.

56. We drafted the questionnaires in collaboration with the RAR Team who were particularly concerned that the questions should not appear, by inviting suggestions for improvements, to anticipate a future RAE of similar form. We aimed to ask fewer questions, each of which had a wide scope, rather than more questions, each narrower. Respondents could limit themselves to assigning a score in answer or could add comments.

57. We have not been asked to make recommendations on how elements of RAE 2001 might be modified for a future RAE. Rather, we have sought to establish how well each element functioned - largely on the basis of the stakeholders’ evaluation, rather than from the primary operational records. We have only in a couple of instances discussed the Chairs’ and Contacts’ criticisms with members of the RAE Team. Generally responses from Chairs to the questionnaire are more critical than the minutes of the Operational Review Meetings, while the comments in the Overview Reports are more emollient still. We have told Panel Chairs and Contacts responding to the questionnaires that they will not individually be identified. The rest of our evidence is already available to the RAR Team and the funding bodies (and much of it is or will be published on websites, e.g. the Overview Reports and submissions to the RAR), so, where it seems relevant, we have identified the source more specifically.

58. We have written the ‘Conclusions and summary’ for a wider audience than that to which the main text is directed, the administrators who will be charged with implementing whatever successor to RAE 2001 is adopted by the funding bodies. The main text provides a collation and distillation of the feedback from stakeholders in the Exercise.

4.3. The questionnaire to Panel Chairs

59. Following an introductory message from the RAR Team, we emailed the questionnaire to Panel Chairs on 13 January, with a reminder on 27 January. The response

was exceptionally high, with 53 (88%) of 60 chairs replying by 10 February, yet another evidence of the enormous commitment of those involved in the RAE. The questionnaire and the covering letter are given in Annex 1. The questions and the scores are summarised below.

60. Panel Chairs were asked to answer each question using the one of these two five-point scales appropriate to it:

5	Strongly yes	Excellent
4	Yes	Good
3	Uncertain	Adequate; neither good nor bad
2	No	Poor
1	Strongly no	Bad

Some respondents gave scores spanning two points, e.g. 3/4, and these are given below as e.g. 3.5. 'N' is the number of Chairs giving a score; the 'don't knows' and no replies are the difference between N and 53. 'C' gives the number of replies with comments amplifying the score.

No.	Question	Score	Scores as % of those responding to the question								N=	C	
			1	1.5	2	2.5	3	3.5	4	4.5			5
			<No								Yes		
											>		
1	Were the procedures for the nomination and appointment of Panel Chair and members appropriate?	-	-	8	-	19	-	45	-	28	53	31	
2	Did the processes involving the generic criteria and working methods (RAE 2/99 and the Guidance to Panel Chairs and Members), which shaped your Panel's statement, allow the Panel to make judgments appropriate to the field and consistent with related fields?	-	-	-	-	9	-	58	2	30	53	19	
3	Were the quantitative data, and their presentation, appropriate and sufficient?	2	-	6	-	23	-	55	-	15	53	26	
4	How valuable were RA5 and RA6 in reaching the Panel's judgements?	-	-	6	-	13	2	48	-	31	52	41	
5	Were you able to take appropriate account of 'staff circumstances' as set out in RA6?	-	-	6	-	11	-	49	2	32	53	24	
6	How appropriate, in aggregate, were the arrangements for sub-panels, specialist advice, cross-referrals and sourcing research	9	2	23	2	28	-	32	-	4	53	37	

	Score	Scores as % of those responding to the question									N=	C
		1	1.5	2	2.5	3	3.5	4	4.5	5		
	outputs?											
7	Panel secretariat: did the secretary and assistant secretary have enough time, training and support to give the level of service required?	10	-	6	-	10	-	38	-	37	52	40
8	The central RAE Team: assess the efficiency of the services it provided. Please comment on individual services (meeting arrangements, audit requests, sourcing outputs, ad hoc requests for statistics, etc.)	13	-	11	-	21	4	47	-	4	53	41
9	Did the funding bodies and the RAE Team do as much as reasonably possible to provide information to the sector about the workings of the RAE?	2	-	2	-	6	-	42	-	48	52	13
10	Was the workload on panel members manageable?	9	-	32	-	23	-	32	-	4	53	28
11	Did the generic grade definitions and the advice on their application make the grading system readily workable?	2	-	12	-	29	-	48	2	8	52	33
12	Were the mechanisms (beyond the grade definitions) to promote comparability and consistency across UoAs (Umbrella Groups, non-UK-based advisers) effective?	10	-	25	2	35	-	25	-	4	52	34
13	Was the panel's decision-taking enhanced by access to expertise and advice of research users?	11	-	25	-	20	-	16	-	27	44	23
14	Were the arrangements and criteria appropriate to enable the panel to give interdisciplinary research proper consideration?	-	-	2	-	14	-	50	-	34	50	21
15	Did the assessment mechanisms disadvantage research, of which evidence was submitted, which had been undertaken collaboratively with partners within and beyond higher education?	41	-	35	-	15	-	7	-	2	46	6

	Score	Scores as % of those responding to the question									N=
		1	1.5	2	2.5	3	3.5	4	4.5	5	
1	Were the arrangements for nominating Panel members appropriate?	1	-	18	-	32	-	47	-	2	92
2	Was the Guidance clear and timely, to you and colleagues?	-	-	3	-	17	2	62	-	16	106
3	Were the arrangements for consultation on the Panels' draft criteria and working methods appropriate?	-	-	6	-	23	-	63	-	9	102
4	Did the statements of criteria and working methods present any problems, while submissions were being prepared?	2	-	42	-	25	-	29	-	2	104
5	Was the supplementary guidance, in the Briefing Notes and website, clear and timely?	1	-	14	-	11	1	52	-	21	106
6	Were the briefing seminars helpful?	-	-	5	-	24	-	52	-	19	96
7	How do you rate the assistance the RAE Team gave in response to ad hoc enquiries?	-	-	3	-	20	1	40	-	36	103
8	Did the structure and processes of the RAE enable your institution to make effective submissions?	3	-	13	-	26	-	46	-	11	99
9	Did the RAE arrangements enable interdisciplinary research to be submitted appropriately?	10	-	26	-	38	-	24	-	1	87
10	Did the RAE arrangements enable collaborative research to be submitted appropriately?	7	-	17	-	42	-	33	-	1	86
		A		B		C	C/B/A	A/B	B/C	A/C	-
11	In which category do you place your institution's use of the RAE Data Application? Enter A, B or C in the box to the right.	48	-	38	-	8	1	3		2	101
12	Did you automatically populate any parts of the table from your own databases (as distinct from re-keying or manually copying and pasting)? Enter Yes or No.	-	-	59	-	-	-	41	-	-	96
13	Did you encounter technical problems in using the RDA?	-	-	31	-	6	1	59	-	3	98
14	Overall, how do you rate the performance of the RDA, including the documentation and other forms of support?	2	1	12	-	39	1	40	-	6	103
15	Was the process of data checking and verification carried out without imposing unreasonable demands?	6	-	14	-	15	-	54	-	12	103
16	How well did the Research Outputs Sourcing System function from your perspective?	8	-	30	-	28	1	31	-	2	93
17	Was the feedback from panels on your submissions helpful?	11	-	30	-	37	-	21	-	2	101

	Score	Scores as % of those responding to the question									N=
		1	1.5	2	2.5	3	3.5	4	4.5	5	
18	How do you rate achievement of each key principle, as described in RAE 2/99?										
a	Peer review	2	-	7	-	24	-	49	-	17	98
b	Clarity	1	-	5	-	41	-	43	-	10	100
c	Consistency	5	-	27	-	38	-	27	-	3	97
d	Continuity	3	-	5	-	38	-	45	-	9	93
e	Credibility	3	-	7	-	35	-	49	-	6	100
f	Efficiency	3	-	13	1	33	1	44	-	4	99
g	Neutrality	6	-	19	-	39	-	32	-	4	97
h	Parity	3	-	22	-	47	-	27	-	1	93
j	Transparency	8	-	14	1	34	-	41	-	2	98

64. The responses are mainly positive, except for questions 4 which was phrased negatively, and 11 and 12 which were factual rather than evaluative; however, questions 9 (interdisciplinary research), 13 (RDA technical - phrased negatively), 16 (ROSS), 17 (feedback) and 18c (consistency) revealed more dissatisfaction than satisfaction, although sometimes only marginally. The relatively high proportions of score 3 (Uncertain/Adequate) should be noted.

5. The central machinery of the RAE

5.1. The Project Group

65. A Project Group oversaw the development and conduct of the 2001 RAE. It comprised senior officers of each funding body, with the RAE Manager and other officers involved in research policy and assessment in attendance. The chair was held in turn by each funding body for a year at a time. The Group in fact had a continuing, though low-key, existence from the 1996 RAE and co-ordinated related work by the four bodies between the two exercises. It therefore embodied a considerable body of knowledge and experience, and was probably the most significant element of administrative continuity between the 1996 and 2001 RAEs. Its particular roles were to carry back to the funding bodies recommendations for approval, at certain key points in the exercise, to agree the budget which the bodies funded in established proportions for joint activities and to monitor progress against the Manager's project plan once it was established.

66. Once the RAE Manager was in post and had prepared his detailed project plan, the Group was essentially reactive to issues which the Manager brought to it, testing that proposals had been well prepared, and providing guidance, support and encouragement. Much therefore depended on the Manager and on the expert advice which he solicited in preparing proposals. By way of example, but a significant one, Panel Chairs' concerns about the grade definitions in the rating scale (reported below) suggest that insufficient attention was given to how the changes would actually operate. The Group approved the revised definitions within four months of the Manager taking up post, his consultations confined to the 1996 Chairs who were not eligible to serve in 2001. Similarly, fewer Briefing Notes to clarify the Guidance on Submissions might have been necessary if drafts had been put to a working group of HEI administrators with experience of previous exercises.

67. 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 in the early stages particularly to ensure that all sources of sound advice are exploited and that proposals are well founded. This point has added importance given the Manager's recommendation that all policy decisions relating to the assessment process should be made before Panel Chairs and members are appointed - to minimise the unproductive debate and bad feeling which evidently arose through too much being unsettled when panels started work (Rogers 5.8). The time from the Manager's appointment (a couple of months later than intended) to the publication of the Guidance on submissions was too short.

68. The Project Group should be serviced outside of the RAE Team, to ensure that meetings do not get squeezed out as the operational pressures build up on the Team.

69. We agreed with the RAR Team that we would not examine the budget. But we note at several points below where the budget did not meet the real cost of the services provided through it and where the volume of services bought may have too small, with the Exercise being under-resourced.

5.2. *The RAE Team*

70. Our surveys and our study of the files show that 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.
71. However, the decision to staff the RAE Team mainly by fixed-term appointments led to certain 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. Some key staff departed at crucial points of the Exercise. While 'golden handcuffs' may have been impracticable, the fixed-term contracts, without enhanced conditions of service over those of permanent staff, are likely to have increased the rate of staff turnover and to have contributed to the ragged run down in early 2002. Secondment from the funding bodies and HEIs would have been more secure, as occurred in effect, with 'back filling', in the Analytical Services Group where the same staff were deployed for 2001 and for 1996. We have already noted that the main, electronic, archive is not readily accessible and that awareness of the evaluative work started had been lost by the time we were commissioned.
72. It is hard to tell whether more experience and more continuity would have made sufficient difference, but several Panel Chairs (responding to question 8), referring to the assessment phase, described the Team as 'inexperienced', 'with significant and constant turnover of staff', 'at times overwhelmed', 'overloaded and less efficient towards the end of the summer', 'very efficient until the end when over-stretched', with a 'level of staff completely inadequate'. Furthermore, once the Panels started work, the Team Manager spent much time attending their meetings, leaving the Team in Bristol somewhat under-supported. These problems must have contributed to the criticisms of arrangements for meeting venues, the handling of cross-referrals, the short suspension of ROSS, etc., on which we report later. With little doubt there was a much greater volume of activity generated by the panels during the assessment phase than could have been (or was) predicted from previous experience: many more requests for research outputs and for cross-referrals. It is highly probable that the Team did indeed need more staff during the assessment phase.
73. Staffing is needed also after the results have been published. HEFCE's solicitors have drawn on their experience of advising on responses to challenges to gradings. They stress that where procedural issues are raised, a funding body should avoid having difficulty with defending potentially defensible actions, only because it cannot itself produce a clear picture of events supported by documentation. It is therefore important to retain corporate memory about the assessment process from start to finish (where finish is at the end of the period in which results are queried), by staff retention if possible and in any event by documentation of processes and decisions. Careful consideration should be given as to the appropriate point for destruction of documents, to strike the appropriate balance between data protection considerations and ensuring that the funding bodies are well equipped to respond to challenges. With the preparation of 'workbooks' the documentation of decisions in 2001 was much improved on 1996, but knowledgeable staff were not available during 2002 to handle the complaints and challenges.
74. 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 which are not within this report's scope to debate. If the

RAE is continued on similar lines to 2001, continuity of staffing in its administration is highly desirable. There are tasks for the quiet years: not only dealing with complaints and challenges, but also evaluating in detail individual processes (e.g. cross-referrals), reviewing the configuration of UoAs, monitoring the burgeoning literature on the RAE, etc. It is, of course, impossible to guarantee that the same staff will be present from one exercise to the next. That, however, is no reason not to replace staff if they depart at the end of one exercise. The issue is not merely one of staffing, though that is important. 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. Hence the systems for preserving and organising the knowledge base need to be far stronger.

5.3. *The Panel Secretaries*

75. For RAE 2001, Panel Secretaries were recruited for the first time from HEIs as well as being drawn from the staffs of the three Funding Councils. In the end, the 60 secretary and assistant secretary posts were divided evenly between HEI and Funding Council staff. Sector staff were recruited via a circular to HEIs, and each application, made with the support of the institution's administrative head, was reviewed by the RAE Manager. There was inevitably some wastage after appointments had been made, as staff changed jobs and so on during the course of the exercise, although this seemed to have a relatively minor impact on effectiveness.

76. A survey of HEFCE Panel Secretaries from 1996 showed that most would do the job again, but only if workload issues were addressed - in other words, they had found it impossible to act as a Panel Secretary and do their normal job, or most of it, simultaneously. This workload problem was not resolved during RAE 2001: several secretaries reported having to work for considerably longer than the 74 days spread over 1999-2001 (notionally 20, 8 and 46 days) originally contracted for. It has been asserted that the problem was compounded by institutions not reducing their staff's regular workload. We have not attempted to verify this, but it would be highly likely as the variable part-time commitment would make it difficult for the institution to make sensible alternative arrangements to deal with the individual's work. A fair proportion of the work was clerical; again we have not explored whether institutions provided support so that Panel Secretaries' time was used to best effect. The payment per day offered to institutions (and presumably also to the Funding Councils) was way below the full cost which the Councils under the financial memoranda require institutions at least to calculate for 'services rendered'. There may be grounds for remunerating panel members on a different basis, but we see no good reason for administrative services not being bought at realistic rates.

77. Some of this extra-contractual work was funded by the RAE as a result of claims by HEIs, but other extra amounts of work were absorbed by the staff and/or their institutions. We suspect that this unfunded work represents a further hidden cost, borne by the sector, of the operation of the RAE. Only the creation of full-time secretary posts, perhaps on secondment from HEIs, would ensure that the full costs of the exercise become apparent, and that workload issues can be properly addressed.

78. The HEI staff brought experience in servicing academic committees and saw these jobs as interesting and as potential career development opportunities. Those changing jobs generally took that role with them, a recognition of its value by institutions as well as individuals, for all the HEI-based Secretaries were well placed to provide expert guidance as

submissions were being prepared. They varied considerably in experience and seniority. This variation was put to use by appointing the more experienced staff to the Panels thought likely to be more problematic, mainly for the many submissions to be assessed. We consider that the use of staff from HEIs was generally successful, and made a positive contribution to the overall success of the exercise.

79. Training days were arranged for Secretaries, conducted by the RAE Manager. This training was well-conceived and executed. As well as imparting policy and technical information about the RAE, they performed an important role in establishing networks which enabled subsequent informal contacts and advice-giving.

80. Panel Chairs were asked whether the Secretary and Assistant Secretary had enough time, training and support to give the level of service required. The question was phrased so as not to invite judgements on the performance of individuals, but over a quarter chose to do so in complimentary terms, the words 'excellent' and 'outstanding' recurring. The 75% of 'Yes' or 'Strongly yes' was often qualified by comment on the enormous workload. Only 16% felt that they were inadequately served.

5.4. Impact on the Funding Councils' Executives

81. The three Funding Councils contributed Panel Secretaries roughly in proportion to their staffing and in that respect shared the load, as also in membership of the Project Group. HEFCW and SHEFC said that the demands on the Secretaries had some impact on other work, by delaying initiatives which might otherwise have been carried forward - but less than in 1996 because certain functions (e.g. booking accommodation, payment of expenses) were centralised in 2001. Other burdens fell exclusively on the HEFCE because the RAE Team was housed at Northavon House. It drew on HEFCE's support services, such as accommodation (at Centrepoint as well as Northavon House), human resource management, public relations, accounting and post room. The last may have been the most visible example, overwhelmed by the volume of mail generated by the Research Output Sourcing System, but the RAE also for example generated surges in claims for fees and expenses, with consequent delays and complaints. HEFCE also had to deal with the 'fallout', by answering correspondence and addressing challenges to the process. Use of these services seems not to have been charged to the RAE and are another instance of its real cost to the funding bodies being understated.

5.5. Risk assessment

82. KPMG carried out a risk assessment on behalf of the RAE during 2000, using the standard methodology of high/low impact set against high/low probability. KPMG based their work on discussions with the RAE Manager, a group of Panel Secretaries and (apparently) one Panel Chair.

83. Unsurprisingly from the experience of the previous RAE, this exercise identified the main risks as relating to the workload for the RAE Team, the Panels, and their Secretaries. These were clearly already the risks uppermost in the mind of the Project Group and the RAE Manager. However, the resource constraints on the Exercise meant that relatively little could be done to address them, other than by obtaining a wider recruitment pool and improved training.

84. The perceived need to commission an external risk assessment of a task which should by 2001 have become near-routine is a further indicator of the lack of organisational learning which the approach to RAE staffing has produced. The risks identified were already well appreciated within HEFCE, even if little could be done to remove them.

6. The framework for assessment

6.1. Units of Assessment

85. The 69 Units of Assessment (UoAs) from the 1996 RAE were retained for 2001, except that the separate UoA for Biochemistry (1996 UoA 12) was discontinued. We did not seek evidence on whether the mapping of UoAs was satisfactory or not. One Chair noted the need for any future exercise to look at the whole medical/biomedical field; another, of an area studies panel, commented on the artificiality of its scope, being an amalgam of several recognised, but small, fields of area study. However, the point we make later about the very high costs per researcher assessed by panels with few submissions builds a case for amalgamation into larger units served by sub-panels. This might have the benefit of more consistent assessments by benchmarking minority disciplines against a broader subject base - at least more effectively than the Umbrella Groups did. (See Rogers 4.21.)

86. Several institutions commented on the presumption behind the UoAs of traditional departmental structures, when their research groups were inter-disciplinary. Doubtless many other institutions could have said the same, as more and more research is organised through 'centres' transcending departmental boundaries. The 'map' for a retrospective RAE is bound to be out of date, may fail to pick up emerging areas of research and may militate against the development of research in areas without a strong tradition (as the Council for Industry and Higher Education and SHEFC note in RAR submissions). A largely unchanged configuration of UoAs has provided some valuable continuity of experience between exercises - and has allowed comparisons of research performance over time. Pressure for fragmentation should probably be resisted, but the way ahead may lie with continuous smallish adjustments and greater user of sub-panels to cater for emerging areas of research. (See Rogers 6.6-11 for a more extended discussion.)

6.2. Rating scale and grade definitions

87. The definition of 'research' was carried forward from the 1996 RAE unchanged. The RAE Manager reported that the terms 'scholarship' and 'teaching materials' were felt by some to be unclear (Rogers 4.9-13). About 20 panels offered guidance on submission of teaching materials (pointing in a minority of cases to the 'Other relevant details' field on RA2) and the European languages providing a Briefing Note; but only three Overview Reports (Law, German, Russian) mentioned them. The Music panel has sought further debate on the nature and assessment of practice-based research in advance of any future RAE.

88. The result of the revisions made to the grade definitions for 2001 may be summarised in terms of the proportions of the research submitted attaining:

Grade	International excellence	National excellence	Sub-national standard
5*	Over half	Remainder	[None]
5	Up to half	Remainder	[None]
4	Some (c.10%)	Virtually all	[<10%]
3a	Possibly some	Over two-thirds	[< one-third]
3b		Over half	[< half]
2		Up to half	[> half]
1		None	[All]

In advance of the assessment, the RAE Manager reckoned that these descriptors had proved easier to explain and interpret than 1996's, once two misunderstandings were removed, that the proportions were of staff submitted and that they were fixed formulaic parameters. He also considered that panels appreciated that work of international excellence did not have to demonstrate that by international dissemination, but less so did the academic community understand because the term 'national excellence' implied second best. He reported that many panels had difficulty in articulating the characteristics of international excellence (Rogers 4.22-27).

89. Panel Chairs were asked: 'Did the generic grade definitions and the advice on their application make the grading system readily workable?' The majority, 58%, agreed or strongly agreed, while 14% disagreed or strongly disagreed.

90. Typical comments from those who agreed were:

The fact that there was some overlap in panel membership between 1996 and 2001 helped to ensure understanding of the process, method and grading.

On the whole yes - I think helped by the fact that panels had to work in advance to develop their own criteria for quality so there was a shared understanding. There was far more concordance over grade assignment than I had predicted in advance

The two advantages were that (a) the system clearly emphasised the judgements to be those of *quality of output* and not other vague notions of national and international reputations, and (b) the quantitative measures to be achieved at different levels of excellence allowed us to be clear about how we made decisions in the difficult cases.

91. However, even those who agreed made constructive criticisms, and those who disagreed were more trenchant. The most frequent criticism was of the need to interpret the grade definitions, and therefore of inconsistency of interpretations across panels. The major reason given was the breadth of the grade bands, especially of grade 5 relative to grade 4; perhaps the threshold was too low. The one Chair scoring '1' said:

This was a major failing of the system. It is clear that different panels interpreted the scale differently and the consequences for funding of subjects are very worrying given the concentration of funding on certain institutions and departments graded 5 and 5*. The main fault lay in translating these into proportions of work: it is clear that the 4/5 divide could be interpreted very differently – a range from 11% to 49% could be interpreted as up to half being of international standard – and huge variations occurred in how panels read this division.

There was a suggestion for splitting the grade 4 band (as 3 and 5 are split), and an extreme suggestion from one panel was for 100 grades, i.e. using percentages.

92. Linked to the issue of breadth was the problem of distinguishing borderlines between grades:

As we discovered, the boundary between 4 and 5 was extremely unclear and even cognate panels operated the grading criteria quite differently. The problem was with the wording 'up to half' (grade 5 definition). Did this mean anything above 10% (as some thought) or did it mean between 30% and 50% (as others interpreted it). The difference in the number of grade 4s and 5s was significant as a result of how the phrase was interpreted.

Another panel found difficulty in distinguishing between grades 3a, 4 and 5.

93. These problems raised the issue of defining 'international' and 'national' excellence, and some panels did not find these terms meaningful. As the published panel criteria and working methods show (RAE 5/99), many panels made their own definitions. The engineering panels, for example, shared the same definitions :

The Panel's concept of international excellence in research is work that will be widely regarded as making an appreciable contribution to the knowledge base within the field, and will influence, or have the potential to influence, the global research and/or the practitioner communities. Such work will help set the international research agenda in the field, or contribute significantly to its development, through, *inter alia*, leadership, impact, publication and/or collaboration. Research of national excellence is work that is not of such widely recognised significance, but is still substantive and, where appropriate, relevant to policy and practice.

94. Many other panels went their own ways, for example UoA 5-8 Pre-clinical etc. related its definition to international journals :

The Panel will use as a guide to the identification of international excellence of outputs, its judgement of the standard of work published in internationally recognised quality journals in the field. The Panel expects that such work will have made a significant contribution to knowledge. Other outputs submitted will be matched to this standard as will information presented elsewhere in the submission. It will consider research that is worthwhile, timely and technically sound, but nevertheless of a lower standard, to be of national excellence.

UoA 10 Nursing refers to international work as 'highly regarded' by researchers in 'leading edge' research countries, while UoA 11 PAMs simply requires that researchers in such countries should be likely to access the work. UoAs 15-17 Agriculture etc. (19 Physics is similar) do not make global comparisons :

International: Highly innovative work, which provides significant new information or ideas resulting in the modification of existing theories, or understanding of mechanisms, or enabling significant new applications or policies.

National: Generates new information or ideas. Provides some understanding of mechanisms or includes some innovative developments of methods, techniques, deductions, policies or practical applications.

UoA 4 Clinical Dentistry had its own 5-point grading scale :

- A internationally outstanding – comparable to the best work in the field, whether conducted in the UK or overseas;
- B internationally significant – excellent;
- C nationally outstanding – good;
- D nationally significant – sound;
- E not nationally significant – making a limited contribution.

95. Three other issues raised in the Panel Chairs' replies:

- A small amount of 'sub-national' work in an otherwise good submission had a distorting effect, which, even with use of discretion, was difficult to avoid, given the grade definitions.
- Normalising outside advice to conform to the scale.
- Relating some of the quantitative data to the definitions, for example the number of research students.

96. Several issues came up again in Overview Reports: grade boundaries (between 4/5 and 5/5*), breadth of the grade 5, and inconsistent interpretations by different panels; although one panel pointed out that the umbrella Panel meeting helped it to review its grades. UoA 14 Biological Sciences felt that: 'the rating descriptors often do not fit well with the subdivision of activity judged international, national or sub-national.' Another panel, UoA 36 Law, reflected on the type of definition (above) which relates 'international' excellence to the nature of the journal :

Work of internationally-recognised excellence was found in a wide range of types of outputs and places, and in both sole and jointly authored works (the Panel adhered to its published criteria in allocating credit for joint pieces). First-rate articles were found in both well-known journals and relatively little-known ones. Conversely, not all the submitted pieces that had been published in 'prestigious' journals were judged to be of international excellence. These two points reinforced the Panel's view that it would not be safe to determine the quality of research outputs on the basis of the place in which they have been published or whether the journal was 'refereed'.

6.3. Statements of criteria and working methods

97. RAE Contacts in HEIs were asked whether the arrangements for consultation on the Panels' draft criteria and working methods were appropriate. 72% agreed or strongly agreed, while only 6% disagreed (the remainder neutral). A typical positive comment was: 'Obviously Panels were able to incorporate only a fraction of the comments received but in my experience there was a genuine attempt by the panels to engage with comments.' However, amongst those which made critical comments (even when scoring positively), the opposite view was put that little apparent notice was taken of HEIs' responses, and that 'critical issues (such as measures of research quality or international repute) were not subject to any significant consultation.' Two Contacts felt that the result was significant variation between panels. It was also suggested that the timing (the summer vacation) meant that some departments did not feel fully engaged with the process.

98. Panel Chairs were asked: 'Did the processes involving the generic criteria and working methods (RAE 2/99 and the Guidance to Panel Chairs and Members), which shaped your Panel's statement, allow the Panel to make judgements appropriate to the field and consistent with related fields?' The large majority, 90%, agreed or strongly agreed, while none disagreed or strongly disagreed.

99. The generic criteria were generally felt to be sufficiently broad to be adapted to a given field. Three Panel Chairs felt that they had used means of establishing criteria consistent with other panels, for example Umbrella Panels. However, the major comment was about inconsistency, on which eight Chairs commented, e.g. 'Not consistent interpretation across the social sciences.'; and 'Doubtful if correlation with related fields.' One panel fought and won permission to vary the RA6 specification, another used research groups rather than individual researchers as the unit to which to apply the grade definitions for outputs, and then had difficulties in so doing, one recounted divergence from other Humanities UoAs in the weighting between RA2 and other components, and a fourth was 'surprised that one panel in the social sciences group was allowed to write criteria which in effect made the RA2 the sole basis of decision'.

100. Although Chairs seem to have satisfied with the generic framework, HEFCE's fundamental review (00/37, paras 150, 151) proposes that panels should have greater freedom

to recognise ‘the different characteristics of excellence of different subjects’, even if that meant collecting different information in submissions for individual (or more likely groups of) panels. The Manager considered that panels needed to consider and articulate more fully the range of criteria by which they would assess the quality of work, exploring some of the more subtle aspects of excellence (Rogers 7.16).

101. The published panel criteria (RAE 5/99) reveal many differences between panels in cognate areas, some less understandable than others, for example some panels normalised all the quantitative data by dividing by the number of returned researchers (sometimes defined as full-time A/A*, sometimes not defined), whereas most did not. The engineering panels had a common complex weighting scheme for the types of evidence, which still allowed the individual panels to vary within common ranges.

102. The Overview Reports mainly say that the RAE 5/99 criteria and working methods were fully applied, that Panels read more of the outputs than originally intended, and that the non-UK advisers confirmed the 5 and 5* grades. Some panels did suggest problems and future changes (UoAs 29 Electrical and Electronic Engineering, 48 European Studies, and 59 History, made extensive comments). Otherwise, the major type of comment was that some HEIs had not followed the published criteria, perhaps due to misinterpretation or to speculation as to whether the panels really meant to apply them as written; for example whether putting in less than four outputs for category A staff would actually not be penalised (if explained). This had meant that many submissions had not optimised their evidence.

103. Consistency of criteria arose again in the Overview Reports, two panels giving examples of mechanisms to achieve it: Umbrella Panels, and the ‘workbooks’ used by panel members to record judgements against the published criteria. Two other panels undertook their own moderation exercise, in one case before the reading of outputs (by practising on outputs from cognate UoAs), and the other by each member sending random selections of cited works to be read by another panelist.

104. RAE Contacts in HEIs were asked: ‘Did the statements of criteria and working methods present any problems, while submissions were being prepared?’ More, 44%, felt there were no problems than that there were, 31%. Positive comments included: ‘I found them most useful as they highlighted differences in approach that might otherwise have been missed.’

105. The most common comment was about ambiguity or lack of explicitness in the statements: 16 Contacts made this point. HEIs wondered, as a result, how panels would operate in practice, and there was much speculation about this amongst researchers. Vague terms like ‘substantial’ and ‘significant’ were criticised, as was the lack of definition of ‘international’ and ‘national’ excellence. A specific problem was the need for more guidance on non-text and practice-based research.

106. The other main negative comment was of inconsistency across panels in the same cognate area; 12 Contacts made this point. ‘For example, some panels rated each publication, and then took an average, whilst others rated the individuals, based on their publications, and then took an average. This can have quite markedly different effects.’ Inconsistencies amongst panels possibly contributed to problems in the submission of interdisciplinary

research. A further example was on the treatment of young and new researchers, mentioned by four Contacts:

Evidence from panels is that they were not treated uniformly. There should be no prejudice against new academic staff, and either in the Guidance, or in the panels' statements of criteria and working methods, it should be made clear how this will be achieved (e.g. they will not be counted when calculating average number of research students per academic).

107. There was specific criticism of the appropriateness of UoA 13 Psychology's, criteria which included 'a critical mass of researchers':

Where institutions submit low numbers of staff for assessment, the Panel will consider whether there is sufficient 'critical mass' to sustain a research culture, and to achieve the stated research plans and objectives. Small submissions (where fewer than 10 staff are submitted as research-active) are encouraged to describe fully any aspects of their research environment within or beyond the submission which help sustain and support the research activities submitted for assessment.

Difficulties were encountered with the statements for UoA 64 Art and Design, and 68 Education (for both of which the requested additional information was a crucial part), and with practice-based research generally.

7. Preparation of submissions

7.1. Submissions by institutions

108. The standard structure of the submissions which eligible institutions were invited to make for assessment in a UoA comprised three main parts in its six forms:

RA1 and RA2: the researchers whose work in the defined period is put forward for assessment are named, with various characteristics (e.g., gender, year of birth, date of appointment if within the assessment period), and up to four 'research outputs' by which that work is to be judged;

RA3 and RA4: statistical data for each year of the assessment period, of research students and studentships and research income;

RA5 and RA6: textual commentary to cover: structure and environment, staffing policy, research strategy, evidence of esteem, individual staff circumstances, contributions by non-research-active staff, and additional information.

We sought the Chairs' views on the adequacy for their panel's work of certain components of the submission, and these are reported later.

7.2. Guidance on submissions, etc.

109. RAE Contacts were generally complimentary about the clarity of, and long notice given by, the Guidance on Submissions (RAE 2/99), with 78% agreeing or strongly agreeing. They had varying degrees of sympathy with the difficulty of getting the Guidance right first time for such a complex exercise and therefore with the need for the Briefing Notes and the FAQs on the website. The more critical regretted that some wording had been unchanged since 1992 and could easily have been clarified or tightened (e.g. by consulting experienced RAE Contacts), that clarifications were still appearing a month before the closing date, so that it became quite difficult to keep track of the prevailing rules, and that these were not consolidated into an updated (online) Guidance document. FAQs should not have been used for important clarifications which should have been published as Briefing Notes. And the Guidance, though available to all via the website, was not suitable for wide dissemination to academic staff, giving administrative staff the task of simplification for application to local circumstances. No one referred to the short guide to the RAE - which may have appeared too late to help in briefing for submissions.

110. Some Contacts found the email discussion list an important (but unofficial) source of clarification, but others felt the RAE Team could have been more active in scotching misinterpretations of the guidance raised there.

111. The most frequently mentioned area of difficulty was the categorization and definition of staff eligible for submission, particularly with supplementary guidance appearing as late as early April 2001.

112. The briefing seminars for Contacts were widely appreciated, with 19% rating them 'excellent' and 52% 'good', with only 5% returning a 'poor' verdict. The only negative comment for future reference is that occasionally promises were made which were overstated or incorrect, e.g. on the functionality of the software and when it would be available.

113. 90% Panel Chairs considered that the funding bodies and the RAE Team had done as much as reasonably possible to provide information to the sector about the workings of the RAE (Question 9). But several referred to the amazing power of myth and rumour and the continuing circulation of misinformation. The Manager and Chairs speaking to subject associations was valuable, and guidance to Chairs on communicating with their communities would be appreciated. The RAE Manager certainly believed that direct engagement with researchers and their organisations was the most successful (but costly) method of tackling the misconceptions. He was surprised that many researchers remained unaware of the email helplines and websites (Rogers 9.1-4). However, Contacts may take a different view as to how far responsibility for briefing researchers should be assumed by the RAE Team, rather than that being primarily institutions' responsibility.

114. To the overarching question, whether the structure and process enabled institutions to make effective submissions, 57% of Contacts responded 'yes' or 'strongly yes', with 16% responding 'no' or 'strongly no'. Amongst the latter, smaller and specialist institutions were over-represented.

7.3. Data requirements

115. The data required in RA3 and 4 impose a significant burden on institutions because they only partly align with data returned to HESA, the principal divergence being that the subject classification represented by the UoAs differs from those used by HESA for courses (to which students are coded) and for finance and staff. The proposed addition of an indicative RAE UoA field to the staff record is relatively simple (but is not included in the current revision: see HESA Staff Circulars 02/05 and 03/01). As research students are attributed to UoAs via their supervisors and as the financial statistics are aggregates of transactions by cost centre (e.g. embracing in some cases, hundreds of research grants), it is far from obvious what scope there is for further convergence without imposing substantial recurrent costs on institutions.

116. Institutions were invited to submit all forms of research output on RA2, and panels were required to assess them equitably. But, as some Contacts reported, the view persisted among researchers that the only form of output valued by panels was the article in a peer-reviewed journal. This gains credence from some panels' definitions of international excellence. Journal articles were by far the largest category of research output submitted and their proportion rose from 62% in 1996 to 70% in 2001 (report by RAE Manager to JISC Scholarly Communications Group [RAR]). When research outputs were not read, but 'reviewed' on the basis of their citations in RA2, the journals' refereeing and editorial standards as known to members must have augmented the information on RA2, more substantially than members' background knowledge about other means of publication, and must influenced their assessment.

117. An innovation in 2001 enabled a researcher who moved between institutions in the year before the census date to be returned as Category A* by both former and current institutions, with up to two publications. We have not collected any information by which to assess the impact of this new category.

7.4. RAE Data Application

118. HEFCE's Analytical Services Group (ASG) was contracted to prepare the software

with which institutions prepared and submitted their submissions and to undertake the data processing of the submissions once received; printing and dispatch to panels was contracted elsewhere. ASG had the considerable advantage of being able to deploy to this work two staff who had been in similar roles for the 1996 RAE. This must have enabled an early and effective start to development. Even so, the software development had to start before the Guidance on submissions was ready, whereas ideally the latter should have been completed first. In the summer of 2001 ASG commissioned Professor Ewan Page to review the data collection for RAE 2001. Professor Page had in 1997 undertaken a similar review of the data collection for RAE 1996. All but one of his 28 recommendations then were implemented for 2001. The one which was not was that: ‘The present [1996] software should be used as a basis for the next RAE and should not be greatly elaborated; particular attention should be given to the ease of use of functions needed by most users of the software.’ However he accepted that starting afresh was almost inevitable if advantage were to be taken of the enormous changes in equipment, control software and institutions’ capabilities. We have not retraced Professor Page’s steps, but have included at Annex 3 his summary and list of recommendations.

119. We did agree with ASG to include in our questionnaire to Contacts two questions on how institutions made use of the RAE 2001 Data Application (RDA) and two questions on the performance of the RDA. The first question asked whether:

A: the RA tables were completed centrally, after data had been collected using own tools	48%
B: the RDA was mounted centrally and networked to departments for completion	38%
C: the RDA was distributed for ‘stand-alone’ use by departments	8%
all possible combinations of A, B and C	6%

120. The second question asked whether institutions automatically populate any parts of the table from their own databases (as distinct from re-keying or manually copying and pasting):

Yes	41%
No	59%

Most of the comments are by those answering ‘yes’, to the effect that problems were encountered in using the upload facility. This was reinforced by the question, whether technical problems were encountered in using the RDA. Of those taking route A, 52% encountered technical problems, and of those taking route B, 74% (overall, 62% encountered problems). One Contact spoke for many:

The major problem that could not be resolved was that the software used Microsoft Access as its driver. The software worked effectively when run on free-standing PCs. However, when a central database held on a central filestore was created, and the pathways on the individual PCs were rerouted to point at this central database, there was a severe downturn in performance. The underlying problem here is that Microsoft Access does not provide a satisfactory driver for a multiple-user configuration. We had received no warning that there would be problems with this configuration.

Another Contact answered against four criteria:

Robustness – Access-based systems can’t cope with significant numbers of multiple users, so we split the database across different parts of the university, and then added them together.

Usability - the RDA was not tuned to different types of use (e.g. occasional versus frequent), and could be frustrating to use.

Reliability - the software had a number of stochastic tendencies, the statistic analyses couldn't be relied upon, and nor could the page length checks.

Scope – the software couldn't cope with all the likely data / circumstances, e.g. for large numbers of authors on Physics papers.

121. The answers to 'Overall, how do you rate the performance of the RDA, including the documentation and other forms of support?' were:

Excellent/Good	47%
Uncertain	39%
Poor/Bad	15%

and seem to reflect that the weaknesses of the software were partly redeemed by its being better than 1996's, and the high quality of the documentation and of the technical support available from ASG.

7.5. Standard audit and verification

122. The question to Contacts, 'Was the process of data checking and verification carried out without imposing unreasonable demands?' yielded 20% of 'No's'. But the comments suggest that perhaps half of these construed the question as embracing the requests for research outputs to be supplied for panel members, whereas it was intended to refer to the arrangements announced in *RAE 2001: Data Checking and Verification* (RAE 3/00, July 2000). The 'Yes's' were 66%. The main complaints concerned the timing, at the height of the holiday season, and the mountain of paper. A short report by the RAE Data Verification Officer appears at Annex B to the RAE Manager's Report; it states that 'as a result of the audit and information provided by institutions around 700 changes were made to the data originally supplied by institutions.' The Manager's conclusion is that the data contained in RAE submissions was 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.

8. The Panels

8.1. Panel membership

123. Panel Chairs were asked: ‘Were the procedures for the nomination and appointment of Panel Chair and members appropriate?’ The large majority agreed, 73%, compared with 8% who disagreed. However, only one of the 31 who made comments was entirely happy: ‘Very good – plenty of nominations for Panel members.’

124. The other comments all offered constructive criticism, of which the most frequent (12) was about nominations from professional bodies and subject associations, firstly that they were important to ensure credibility, but mostly that these nominations had too much weight and that some were implausible and unbalanced, for example: ‘Subject associations nominated people without regard to sub-disciplinary balance’ and ‘next time nominations should be justified e.g. by list of publications.’

125. Eight Chairs were concerned about spread and balance, and some felt that their better overview should justify them having more discretion, for example to nominate outwith the names submitted. Nominating bodies should be given more guidance, for example regarding balance by gender, qualifications, geography, subject (including minority interests), type of HEI. There were some difficulties in getting users of research, and one chair felt this was a ‘last minute’ process. Three chairs felt that their own appointment process could be more structured and transparent. The British Academy [RAR] considered that subject associations were not always the best way of soliciting opinion and further consultations should be held, including input from the user community. The RAE Manager noted that complaints about panel membership centred on the absence of members active in interdisciplinary work and/or new or emerging subjects (Rogers 5.3).

126. RAE Contacts were asked: ‘Were the arrangements for nominating Panel members appropriate?’ Considerably more agreed or strongly agreed, 49%, than disagreed or strongly disagreed, 19%. There were two major issues: that the HEIs were not invited to nominate (7 comments); and that while the process for *nominating* members was acceptable, that for *selecting* them from the nominations was not transparent (10) and likely to lead to ossification (7). Three also commented on the lack of transparency in selecting the Panel Chairs being a greater concern. Several RAR submissions also call for greater transparency and democracy in panel appointments, without suggesting how.

127. Four Contacts commented on the under-representation in panel membership of the post-1992 universities and the colleges (which has been attributed to the pre-1992 sector’s continued control of subject networks). The resulting poor subject balance of panels was a concern to four Contacts. Two commented on the non-UK experts, one suggesting that they should be full panel members (otherwise they can only ratify the grading process), the other suggesting that arrangements for requests for their nomination were inappropriate. A later deadline was also suggested for nominations to sub-panels, as the need for them sometimes emerged during panel formation.

128. No one has commented directly on the restrictions to length of service (no more than two consecutive RAEs unless appointed as Chair for the third), though a couple of Contacts

called for more 'new blood'. The RAE Manager (Rogers 5.1-2) noted that the actual proportion of new appointments was around two thirds, but that panels' starting point in drafting criteria was still how the panel worked last time: turnover of membership was not bringing new thinking. He also recommended that panel members should be required to attend a structured training course as a condition of appointment, as relying on Chairs and Secretaries to transmit briefing produced intermittent results (Rogers 5.7).

8.2. Representation of research users

129. Attempts to increase the involvement of researchers users in 2001 were only partly successful. About three-quarters of the panels involved users directly in assessment, six panels doing so by constituting a users' sub-panel. But recruiting users of the appropriate calibre was difficult: the body of nominations initially received was very deficient and several supplementary calls for nominations were required to secure sufficient names. Even then, Chairs frequently felt they had insufficient information on which to assess suitability (Rogers 5.4-5).

130. We asked Panel Chairs whether the panel's decision-taking was enhanced by access to expertise and advice of research users. Although more strongly agreed, 27%, than strongly disagreed, 11%, the balance of positive, 43% and negative, 36%, scores was closer. Positive comments included: 'substantial contribution', 'significant impact on decisions' and 'extremely helpful user sub-panel'; and negative: 'more a political inclusion matter than useful for research grading', 'users [on the sub-panel] just did not know enough about advanced level research to be able to comment'. Chairs distinguished between user panels (and one with hindsight would have liked one) which had a role, albeit complementary, and individual users whom it was doubted had much effect on the final grades, but were useful for credibility. Several mentioned that other commitments limited users' contributions, while the CBI [RAR] was told by some users that the time commitment, amounting to some 20 days, was unacceptable. One user thought that asking a user to act as an academic assessor was inappropriate and that users should contribute by commenting on the professional and policy relevance of the submissions.

131. It may be helpful to follow up in more detail those instances where Chairs were enthusiastic about the users' involvement, in order to identify 'good practice' for a future RAE. The RAE Manager has suggested that panels should have direct discussion with users during criteria setting and pointed to the value of the seminar held for panel members, with DTI and CBI, on industry, commerce and the RAE (Rogers 8.6; seminar report on the RAE website).

132. We have not collected evidence on the role of 'observers' and have no comment to offer.

8.3. Sub-panels

133. According to the published list, seven panels had one, and one (small) panel three sub-panels, to advise on a sub-field. Three of these handled a new or emerging field of study: Women's Studies; Science Policy; and Film, TV and Video Analysis, History and Practice. Six panels constituted sub-panels of research users (UoAs 13 Psychology, 23 Applied Maths, 24 Statistics, 35 Geography, 39 Politics, 42 Sociology). The Information Systems sub-panel was joint between UoAs 25 Computer Science and 61 Library & Information Management.

This RAE saw the introduction of thematically based sub-panel structures. The principal example were seven sub-panels which operated under the three Medical panels (UoA 1-3). The Development Studies Sub-panel was common to UoAs 35 Geography, 38 Economics, 37 Anthropology and 39 Politics. Six Chairs made positive comments about sub-panels working well, and the other was concerned about lack of administrative support in their appointment. Two Overview Reports (UoAs 1 Clinical Laboratory Sciences and 61 Library & Information Management) recognised the importance of joint sub-panels, to ensure consistency of treatment and standards. The Operational Review meetings with Chairs confirmed that sub-panels had worked well.

134. There was some support from Chairs for more sub-panels to span a range of panels (on the model of Development Studies). The RAE Manager went further: sub-panels combined effectively with the discipline-based main panels had the potential to provide an assessment matrix which is more sensitive to changing patterns of research activity than the main panels alone (Rogers 6.5).

8.4. Workload

135. Panel Chairs were asked whether the workload on panel members manageable. 41% disagreed or strongly disagreed, but 36% agreed or strongly agreed. The positive comments were along the lines, 'just about able to cope', but 12 thought it was much too heavy, using phrases like 'unreasonable and unrepeatable', 'outright exploitation', 'unsustainable' and 'too heavy, near revolt'. One chair felt that 'members had unrealistic expectations of what they could read.' Industrial members 'crumbled a little' and in another panel they 'thought the fee preposterous.' The inadequacy of remuneration was mentioned by four other chairs. The loss of research time by the panelists led one to question the cost/benefit of the Exercise. Chairs themselves were particularly afflicted by the workload, one had to 'live and breathe RAE far too much for far too long', another thought his 'health is marginally but permanently damaged', and a third felt he had 'lost the best part of a year's research'.

136. Factors exacerbating the workload problems were said to be :

- too much in too short time (5)
- timing very bad: clash with exams
- sourcing delays (3),
- use of email not permitted
- preparing workbooks an unnecessarily onerous chore (2)
- difficulties when another panel referred 100+ items
- significant reduction if category C staff excluded from RA2
- many universities allowed no relief to panelists
- involvement of non-UK based advisors close to the end of the exercise
- the amount of secretarial work

137. The two comments in Overview Reports (UoA 11 Other Studies and Professions Allied to Medicine; UoA 59 History) were more measured: 'while the workload for Panel members was high all agreed that participation in the exercise had been a positive and rewarding experience' and 'the timescale available for inspection of the outputs was short, and it was with considerable difficulty that the Panel managed to complete the workload to its satisfaction.'

138. A couple of Chairs and the RAE Manager (Rogers 5.9) registered concern at the modest fees paid to panel members (particularly user representatives), £150 per day for attendance at meetings plus £850 for preparation. These rates may have originated by analogy with those paid for external examining and similar tasks which academics undertake in the common weal and without relief from their obligations at their home institution - an 'honorarium' retained by the academic rather than reimbursement to the employer for time foregone. Gross salary cost alone for a professor (as costed e.g. for a consultancy) would be around £275, with overheads in addition (typically at least 70%). The assumption of about six days for preparation is probably a substantial understatement of actual for all but the lightest-loaded panels. For a future RAE, the basis for and rates of remuneration should be reconsidered.

8.5. Panels' differential loading

139. The 68 Units of Assessment were assigned for assessment to 60 panels, some being responsible for more than one UoA. The numbers of submissions and of researchers returned in the submissions varied enormously between panels. UoA 43 Business & Management Studies had the biggest load on both measures, with 97 submissions and 2555 Category A and A* staff. UoA 46 Middle Eastern Studies had fewest submissions at 11, and UoA 54 Russian, Slavonic & East European Languages fewest researchers to assess at 77. These differences in workload were not compensated for by larger panels, as numbers of members were within the range only of 6 to 20. Researchers submitted per member ranged from 8 in Russian, Slavonic & East European Languages to 196 in Business & Management Studies. (Sub-panels served to spread the load for those panels which constituted them.) The consequence was not that the former panel completed its business with a small fraction of the latter panel's effort. Rather the degree of scrutiny accorded to submissions expanded to utilise the capacity available. One panel with a light 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%.

140. A related issue (identified by Rogers 6.3) is the comparative framework within which submissions were judged. In e.g. Business & Management Studies, submissions were assessed against a broad and diverse background of research presented to the panel. Submissions to a 'small' language panel were measured against a very small body of research work. The broad assessment in English or History contrasted with the narrower one in American Studies - where much of the work was concerned with literature in English or with historical questions. The difference in size of panels' loads questions the consistency of assessment at least across cognate areas.

141. The resources available to each panel were broadly the same: all panel members and sector secretaries' institutions were paid the same, irrespective of load. Rogers table 1 gives figures for the direct cost per researcher (i.e. those costs attributable to the panel), ranging from £606 for UoA 45 American Studies and £591 for UoA 49 Celtic Studies to £32 for UoA 43 Business & Management Studies and £27 for UoA 3 Hospital-based Clinical Subjects. Indeed we have been told (though we have not verified) that lighter loaded panels demanded even higher levels of support: that as they proceeded to second readings, they requested outputs again, and that they generated cross-referrals of individual outputs (not researchers) for a further opinion.

142. This raises the question whether the grades awarded by the heavily loaded panels are any less reliable than the grades awarded by lightly loaded panels. We have not noticed that asserted, other than in the suggestion that where reading was selective it was the better works which were read, resulting in higher grades compared with UoAs where a high percentage of outputs was read. If the grades awarded by the heavily loaded panels command acceptance in their respective communities, should less thorough scrutiny be adopted in other UoAs? Should the allocation of resources in support of panels be more unequal, so as to be more proportionate to the numbers of submissions and researchers under assessment?

8.6. Confidentiality and declarations of interest

143. Panel members were required to sign a confidentiality agreement (Panel Members' Handbook, section 4). It is evident from some of the more serious challenges to gradings - and the many anecdotes in circulation - that the agreement was not universally observed. HEFCE's solicitors raise the following issues for any future research review which includes a peer review element or otherwise requires confidentiality to be preserved by people other than staff:

- ensuring that the wording of confidentiality undertakings is as strong and clear as possible;
- ensuring that the concept of collective responsibility for decisions by panels is clearly explained and also embedded in statements of the assessment criteria and working methods;
- in particular, emphasising that it is the responsibility of panel members to raise concerns about a decision, whether about its substance or the process by which it was made, *during panel discussions*;
- ensuring that the reasons why confidentiality undertakings are so important is explained/emphasised to panel members regularly before and during the assessment process;
- considering in advance of panel proceedings what the general policy should be where a breach of confidentiality by a particular Panel member is identified or strongly suspected;
- emphasising the duration of the need for confidentiality. Discussions about results before they are announced or about the basis of them before the feedback has been received and digested are very vulnerable to the inference that they indicate that the panel themselves are uneasy about their decision, that the feedback does not give the real reasons for the decision or that the panel had a lax approach to process generally.

9. The assessment process

9.1. Communications

144. Institutions prepared their submissions using the RDA and sent them to the RAE Team in electronic format - which was vital to the Team's subsequent processing of the submissions' content. But the submissions were distributed to Panel Members only on paper. Only Panel Secretaries were provided with copies on CD. Neither group had on-line access. The Panel Members' Handbook forbade faxing of paper copies of RAE data and information relating to the panel working methods or the assessment of submissions, unless the recipient was involved in the assessment concerned and would receive the information directly from the fax machine. Email was permitted to facilitate administrative processes, but not to distribute or discuss RAE data and any information relating to panel working methods or the assessment of submissions. RAE data might not be loaded into any computer (or other storage medium) which was not under the direct control of one of the funding bodies without written authority from the RAE Manager.

145. These prohibitions on using the customary means of communications in the academic world must have been frustrating and with little doubt caused considerable inefficiencies (e.g. in the cross-referral process). We appreciate that there must be concerns about (a) data being accidentally or deliberately changed and (b) the confidentiality of Panels' deliberations, though we have not looked at the advice which led to the restrictions imposed. But the planning of any future assessment exercise must pay more attention to enabling 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.

146. We have not established how often and in what circumstances the RAE Manager permitted RAE data to be loaded onto a computer outwith the control of a funding body, but it is evident that some panels initiated extensive transcription from the paper copies, with all the inherent potential for error.

9.2. Timetable

147. Panels received submissions in May and reported their grades by the end of October 2001. These dates were the same as in 1996, but the actual time available to panels was shorter because two new steps were introduced for 2001, consultation with non-UK-based experts and discussion by Umbrella Groups, which meant that near-final grades were required around the middle of September. Furthermore, the volume of assessment activity was greater than in 1996, not in terms of submissions received or researchers named, but of outputs read (and therefore requested), of referrals to specialist advisers, of cross-referrals between panels, of preparing workbooks, etc. Generally Chairs' criticisms of assessment processes, reported below, arose from the massive peaking of activity during the summer, 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. As we have said elsewhere, the risk to the Exercise from staff resignation or absence - and also from IT system failure - was high. Aligning the assessment period with the financial/academic year, rather than the calendar, and making the submission date early in January while keeping the publication of grades to early December might make

an essential difference (see Rogers 10.2-4). But the peak would still fall in the summer vacation, which is particularly harsh for those with young families.

9.3. Adequacy of the data collected

148. Panel Chairs seem to have been broadly satisfied with the range of data collected in RA1 to 4. Points made about RA2, research outputs, were:

- allow A* staff to submit four publications
- require at least one publication per year and the total count of publications of all returned staff
- having only four publications was insufficient: require all the publications of each researcher, to explore the depth as well as quality of the output [from a Panel with an extremely heavy load]
- identify multi-authored papers cited by more than one author in the same submission
- ensure greater clarity about internally published items
- collect more information on the researcher's contribution to a multi-authored paper
- use (at least for medicine) the standard internationally accepted format for citations (UoA 1 Clinical Laboratory Sciences minutes).

149. Five Overview Reports commented on the number of outputs to be submitted per researcher. UoA 53 felt that the expectation of four outputs for A and C category staff encouraged premature publication. UoA 59 History feared that it encouraged submission of the same research twice in different formats (and languages) and that institutions might be automatically disqualifying category A staff with fewer than four publications (while more willingly entering category C staff). However, UoA 29 Electrical & Electronic Engineering considered that four was too few and led to RA2 data being of less use than in 1996. The inclusion of only two outputs by A* researchers created some difficulties for UoA 38 Economics when making an assessment of the depth of research for those individuals.

150. Developments in electronic publication have implications for future RAEs, e.g. interactive research resources can only be properly assessed in their live form, but the RAE has operated a fixed cut-off date some months before outputs are assessed. The terms of reference of JISC's Scholarly Communications Group include considering how the RAE should take into account developments in scholarly communication; see its submission to the RAR.

151. As to comments on RA1 (research-active staff), RA3 (research students and studentships), RA4 (research income), and RA5 and 6, this list includes a few points gleaned from other sources:

- give grants awarded as well as spend
- give income rather than expenditure
- give all financial data for the calendar year [because the assessment period was for calendar, rather than academic/financial, years]
- put all research income on RA4, whether or not through the institution's books [this exclusion may be the source of the reported discrepancies between RA4 and RA5/6]
- restrict returns of research income and research students to those associated only with the returned staff
- return only those research students whose supervisor was research-active.

- give standard analyses with totals by year, per capita, named researcher for sums over a given threshold, total postgraduate numbers for period, average length of postgraduate research enrolment, success rate
- collect more metrics on each assessed person, e.g. total number and type of performance/recognition indicators (publications/plenary lectures/medals/grants/etc), and collate onto RA1/2
- give the date of the researcher's PhD if recent, to identify early-career staff
- relate the permitted length of RA5/6 (which some Contacts want expressed in words or characters) more closely to the number of researchers [i.e. the banding is wrong]

152. The question must be asked whether the use which the panels made of the data in RA3 and 4 (and the FTE of research assistants and research students in RA1) is worth the costs entailed. With the exception of the Engineering Panels and UoA 11 Professions Allied to Medicine, all panels gave most weight in assessment to the research outputs in RA2, often with the rest of the information contextualising the outputs and/or influencing borderline decisions. The majority of the data in RA3 and 4 relate to inputs and, particularly outside the sciences, often hard to interpret (see Rogers 7.7 and 10). It should be possible, in consultation with some Contacts, to identify the data elements which were most onerous (e.g. the FTE of research assistants and research students supervised by each researcher returned; splitting research expenditure across financial years) and to review their utility in more detail than has been possible here.

9.4. Presentation of data

153. For circulation to Panel Members, the information on each researcher on RA1 was conflated into RA2. Some Chairs regretted that RA1 was not issued, but our understanding is that RA1, or something very close to it, was one of the standardised 'assessment aids' which Panels were invited to request in advance. Each submission as circulated was prefaced by a table of figures, summarising the numbers of staff and research outputs submitted, and the values in RA3 and 4, also given per FTE and headcount A, A* and C staff. The same data were given for the UoA in total. A couple of Chairs said the presentation could be improved, without being specific, and the Operational Review Meetings called for the Data Systems Working Group to include representatives from panel members, to ensure that the summary statistics met the panels' requirements. UoA 50 English said that the aids arrived too late to be incorporated into the assessment and in any case were too bulky to carry to meetings.

9.5. Ad hoc statistical analyses

154. Panels were invited at their November 2000 meetings to specify any additional statistical analyses which they wanted, so that ASG could prepare for these in advance of the submissions arriving. The Engineering Panels decided on generating initial grades from the quantifiable data and negotiated appropriate analyses with ASG. So far as we know, these ad hoc statistical analyses specified in advance were provided in a timely fashion. But when panels met after receiving the submissions, they made further requests of statistics. ASG was by then heavily committed to the planned processing of the data and did not have the capacity to respond to these requests within the timeframe that panels needed.

9.6. Audit queries

155. Panel members raised over 300 queries about information in submissions, for investigation by the RAE Team. According to the Data Validation Officer's report, most of

these queries arose from misinterpretation of the RAE Guidance rather than deliberate attempts to deceive. Chairs collectively questioned the efficiency of the system for responding to the queries, and asked that it should be better supported in any future exercise. A couple of Chairs complained at the slow response to their queries and at the queries not being pursued hard enough. The RAE Team's reply to the latter was that some queries relied on anecdotal information ('X has left Y University and is working in Italy') and that Y produced evidence that X was on the payroll.

9.7. Selection of staff for submission

156. The major issue raised by six panels in their Overview Reports was the effect of departments selecting fewer of their staff for submission, compared with 1996, as one explanation of the increase in gradings seen in most UoAs.

It was clear that many Institutions operated a carefully selective policy in deciding which staff members should be returned. To what extent selective return of staff should be retained in future RAEs is a matter for consideration following the completion of the present exercise. (UoA 62 Philosophy)

Another panel considered that 'the situation whereby not all staff must be submitted devalued the process and recommended that other systems be sought' (UoA 14 Biological Sciences). A third panel felt that

unless these anomalies are rectified, the relative assessments of those departments returning high proportions of their research staff and those returning low proportions of their research staff will continue to be biased and unfair. Most panel members were very concerned about the above point. (UoA 29 Electrical & Electronic Engineering)

Other possible explanations of the upwards drift of grades were some HEIs not submitting to UoAs where they had scored low in 1996, and the cumulative effect of research funding from previous RAEs.

157. Universities UK [RAR] considered that selecting research-active staff did not recognise the team effort required to undertake research. Janet Lewis, in *THES*, 10 Jan. 2003, saw the RAE as entirely geared towards the individual, not the team, and so based on an outmoded model of knowledge production. This touches on a point by UoA 1 Clinical Laboratory Sciences, that further thought is needed on how panels should integrate judgements on individual items of output, to derive judgements on individual researchers and hence on research groups and whole submissions.

9.8. Use of RA5 and 6

158. Panel Chairs were asked how valuable were RA5 and RA6 in reaching the Panel's judgements. The large majority, 79%, said they were excellent or good, and only 6% that they were poor.

159. The comments reveal a greater split, 15 saying that RA5 and RA6 were important or essential etc, 15 said they were of little or marginal use. The common ground would be that they could be very helpful in resolving borderline grading cases (6), and that they could provide as useful context and overview (5). One problem was that, as HEIs have 'learned to play the game', they have become better at writing 'hype', so the statements failed to discriminate. RA5 was more useful than RA6, according to two panels, but both had a minor influence, compared with RA2. They were 'difficult to evaluate and give weight to',

according to one panel, but Engineering Panels did define weights. The permitted length could be reduced (2). One panel mentioned that the software was insufficiently flexible to accommodate RA5 and RA6 information.

160. A more standardised presentation was also suggested. An example of the possible need for this can be found over the element of self-assessment introduced for 2001. RAE 2/99 stated that this would be facilitated: 'Accordingly, panels will specify a structure to the textual parts of the submissions (RA5 and RA6) which facilitates self-assessment under a number of headings. Panels will make these specifications within a basic common framework.' However, panel's published criteria (RAE 5/99) varied considerably in the extent of facilitation, some merely inviting HEIs to include something brief (e.g. UoA 20/21 Earth Sciences: 'Where a department wishes to do so, a brief statement of self-assessment may be included'), others requiring an objective and self-critical evaluation (e.g. UoA 18 Chemistry and UoA 15/16/17 Agriculture etc.) and comparison with 1996 research plans (e.g. UoA 9 Pharmacy).

161. Five of the published Panel Overview Reports contain comments on RA5 and RA6.

162. Panels reported that some RA5 and RA6 were too long, yet not very informative (UoA 52 German etc.), too subjective (UoA 53 Italian) or exaggerated (UoA 55 Iberian etc.), and that in some cases the information contradicted that in RA3a, RA3b and RA4 (UoA 52 German etc. and UoA 55 Iberian etc.) and in RA2 (UoA 64 Art & Design). One panel suggested that more guidance needs to be given on what constitutes an 'esteem factor' (UoA 52 German etc.). Two reports (UoA 53 Italian and UoA 64 Art & Design) considered that sections on research strategy were of limited value and not always coherent. Further, in the only reference to self-assessment:

there was also a substantial variation in the quality of response to the self-assessment and SWOT analysis invited by the panel. Strengths identified in these were sometimes inconsistent with the quality and volume of research cited elsewhere in the submission and weaknesses, potential threats or opportunities were not always identified. The panel was concerned to note that policies and mechanisms which might foster the strengths and address the weaknesses or imbalances were not always identified in strategies for post 2001 (UoA 64 Art & Design).

The Engineering Panels normalised the quantitative data (dividing by FTE researchers) and gave scores for the vitality and standing of departments as evidenced by RA5 and 6, and weighted them so as to produce an initial grading as a basis for discussion. UoA 29 Electrical & Electronic Engineering found that 'The RA5 and RA6 data were the best correlated with the final rankings, and were given equal ratings of 20%.'

163. Panel Chairs were also asked: 'Were you able to take appropriate account of "staff circumstances" as set out in RA6?' 47 replied, the large majority agreeing or strongly agreeing (38) and only three disagreeing.

164. Several Chairs commented on some institutions using this heading very constructively and others failing to put it to advantage - or being a bit opportunistic. Ten Chairs felt that the information was only relevant in a few instances, though the allowable cases probably were identified. One panel had allocated special responsibility for this issue to two members and this had worked well. The major use of this data was for new and young researchers, and six Panel Chairs valued this information, and another considered that maternity leave information

had been useful. An Overview Report, UoA 63 Theology, expressed concern 'lest the RAE process be seen as giving inadequate credit to younger and less well established staff who have not yet published four quality publications' - which suggested that RA6 was not judged effective. The RAE Manager considered that for the future panels should state explicitly how they would take 'staff circumstances' into account and what were their normal expectations of early-stage researchers (Rogers 7.27-28).

9.9. Panel proceedings

165. The samples of Panel meeting files which we examined indicated that the meetings were organised and managed effectively, and that they were conscious of the RAE schedule within which they were operating. Panel Secretaries took slightly differing approaches to the presentation of information in the minutes of meetings, but not to such an extent as to suggest the need for a stronger centrally-provided template.

166. The final preparation of the Panel Workbooks, setting out in detail how they had reached their grading conclusions, 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. But one Chair considered them a waste of time, another unduly onerous to complete. The one Overview Report comment was that they 'facilitated the ready and consistent application of the Panel's criteria to all submissions and helped formally record the Panel's decisions' (UoA 61 Library & Information Management). Another panel, UoA 40 Social Policy & Administration, felt that the workbooks overemphasised factors other than research output.

9.10. Arrangements for panel meetings

167. At the Operational Review Meetings in May 2002, 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. A dozen Chairs repeated these complaints in their questionnaire responses. The explanation we received was that many panels, understandably because of travel, wanted to meet in London, but that the maximum cost per head allowable there under Civil Service rules was (unrealistically) low. In 1996, arranging accommodation was devolved to Panel Secretaries - though, as they were drawn only from the funding bodies, the arrangements were being made from four locations, with some degree of local pooling of administrative support. The engagement for 2001 of staff from institutions as Secretaries probably pointed to centralised booking; and Funding Council staff who served in both 1996 and 2001 were pleased to be relieved of that task, allowing them to concentrate on the tasks requiring the skills and experience for which they had been selected.

168. We offer the following comments:

- other analogous public bodies, such as the Research Councils, have similar needs for (residential) meetings: their experience should be tapped
- particularly if a future exercise were conducted with fewer panels of wider scope, each panel might be given its own budget - but the arrangements for (sector) Secretaries should include also some local clerical support
- if a future assessment exercise again falls in the summer, the better quality university

conference accommodation in a couple of locations might be block-booked

9.11. Research Output Sourcing System

169. The most time-consuming element of a panel member's load was probably reading all or a sample of the research outputs cited in the submissions under consideration. Panel members were asked to consult the items by the means readily available to them, normally through their institutional libraries or on-line. HEIs were required to make available, on request, any item cited in their submissions, and for items not readily available to a member, the RAE Team established a central Research Outputs Sourcing System (ROSS). The system had a security protected web-based interface, linked directly to a copy of the RAE Database, through which the member could request research outputs. The system also allowed panel members and institutional Contacts to track the progress of requests for items. To protect the identity of the panel member requesting an item, and to ensure that the correct items had been supplied by HEIs, all outputs requested were sent to the HEFCE office, unpacked and checked, and repackaged for dispatch to panel members. In total, over 45,000 items of research output (equivalent to nearly 25% of the total cited in the submissions) were requested through the system.

170. ROSS was suspended for the receipt of new requests, from a Thursday to the following Monday, at the height of the demand for items, in early August. The reason was the backlog of items received from institutions awaiting redistribution to panel members. The backlog was estimated at some 9000 items, when the capacity to clear was about 1000 a day. Staff worked 8 a.m. to 8 p.m. over the weekend. The constraint on capacity was more physical space than staff numbers.

171. The suspension of ROSS is firmly embedded in the collective memory and was repeatedly mentioned by Panel Chairs. But some admitted that ROSS worked better than expected, was less time-consuming than other means of accessing items, and gave a better service than 1996's arrangements. Indeed, ROSS may have been the victim of its own success. Given that the typical proportion of outputs which panels committed themselves to reading was 25%, and allowing for second readings in some panels and for cross-referrals, ROSS may have provided a majority of items read. We suspect that some members used it in preference to consulting copies in their own institution's library, when summer vacation opening hours were likely daytime Monday to Friday, because they could read material when and where suited them. Certainly a good few Contacts thought this was happening. One Contact with nine submissions quantified the extreme variance in the number of requests made by different panels: from zero in one UoA to almost a half of the entire submission for another.

172. Contacts were divided roughly evenly between those who found ROSS worked satisfactorily, those uncertain and those who did not. But the comments, amongst the lengthiest, are broadly similar: surprise at the high volume of requests, the high cost in staff time and postage, coping just about, occasional losses or mislaying of items in handling at Bristol, and items returned damaged or not at all. (One problem noted was that a new output number was generated when same output was passed to another assessor.)

173. Because ROSS was a computerised system, and if the data have been kept, a detailed analysis would be possible of the patterns of requests.

174. 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. The latter certainly would hit problems of copyright; the former might also if institutions deposited photocopies (indeed if institutions formed collections of cited outputs (see below) using photocopies, that might be open to challenge).

175. The RAR Team had heard that Panel Members had placed a substantial load on their home institution's libraries, particularly by requiring inter-library loans. We pursued this point by asking Toby Bainton, the Secretary of SCONUL (Society of College, National and University Libraries), to send an email enquiry to his members; we appreciate his assistance. Fourteen replies described the library's involvement. That appears to have been largest by way of: verifying bibliographical descriptions of outputs to be cited in submissions; forming a special collection of the cited works (including by ILL); and servicing the requests coming via ROSS. However, six libraries described special arrangements to assist Panel Members in accessing outputs, including the provision of private reading space in the library. Two specialised libraries registered a noticeable demand from panel members in other institutions. Several universities gave the library some additional funding.

9.12. Specialist advisers

176. The published list names 464 specialist advisers; the Manager's report says 350 were consulted. In a question to Panel Chairs, we listed specialist advisers with other means of panels gaining external advice. A dozen in their comments referred to them: four were generally positive and six negative. Most comments were on administrative matters, for example: 'I had to contact specialist advisers myself because of [RAE Team's] inertia', and 'administrative support from [RAE Team] was weak or non-existent. They were obviously overwhelmed'; the advisers had too little time and were poorly paid [though they were unpaid in previous RAEs]. Four commented on the quality of advice: it was variable (4), some not being as useful as expected, or not received at all, and one panel 'needed good commentary to be able to have the confidence to take external views on board.'

177. Only one Overview Report made a substantive comment, and that was procedural: 'All scores from other panels or Specialist Advisors were reported back to the Computer Science panel though a nominated link member of the panel; in this way, the panel avoided misunderstandings between different scoring systems' (UoA 25 Computer Science).

178. The Operational Review meetings considered that in the main 'the Specialist Advisors process had worked well, and had resulted in useful advice to panels. Some panels would have liked the flexibility of running the process for requesting advice itself.'

9.13. Cross-referrals

179. Comments from Panel Chairs specifically on cross-referrals (30) outnumbered those on specialist advisers by a factor of over 2: 1. Further, they were almost entirely negative (21 vs 3). Panels felt that the volume of cross-referrals had increased since 1996, outstripping the system and resources to handle them.

180. The mandatory cross-referrals, those requested by institutions, were relatively straightforward, in that they were flagged in the submissions and the whole submission could be forwarded at an early stage to the second panel. UoA 1 Clinical Laboratory Sciences wanted a

threshold of size for these: referrals of one researcher or even a few outputs were difficult to assess in context. The major difficulty lay with cross-referrals generated by panels. Because of the concerns about the security of electronic communication, the arrangements were entirely paper-based. They started with the Chair of the sending panel completing a form, typically identifying the researcher whose work should be cross-referred and giving guidance on the advice sought. The form went to the RAE Team which forwarded it, with the RA1/2, to the Chair and Secretary of the receiving panel; the Chair chose which member(s) of the panel should assess the outputs and notified the Team; the Team then informed the assessor(s). The Panel Members Handbook (2.11) stated the expectation that the assessor(s) would prepare a written report which would go to a face-to-face meeting of members from the two panels or, for the occasional cross-referral, would be discussed by telephone.

181. The processes were criticised for being too centralised and complex, and therefore slow. 'Central organization of the cross-referral system was far too cumbersome, leading to long delays, compounded by a veto on the use of emails between panels.' 'In aggregate the support from [the RAE Team] was inadequate. We arranged cross-referrals ourselves.' 'We found ourselves unexpectedly "dumped on" by some other panels, which generated considerable quantities of unanticipated work.' Some panels seemed to cross-refer too much, others not as much as receiving panels had expected. Reports on cross-referrals were of variable quality and timeliness, depending on the member of the other panel (2). 'We were at the mercy of the very varying efficiency of different Panels in conducting their business. We could have had a crisis here and only narrowly averted it.' Three Panel Chairs mentioned problems across panels having different criteria and approaches: 'problems of incorporating heterogeneous ways of assessment by different panels', 'and there was a definite difficulty where a small panel (accustomed to reading all cited outputs) was reporting to a big panel (accustomed to reading only a small proportion of cited outputs) or vice versa.'

182. Only five panels made more evaluative comments in their Overview Reports (UoA 5-8 Pre-Clinical etc., 29 Electrical & Electronic Engineering, 38 Economics, 59 History and 65 Communication, Cultural & Media Studies). There is no inkling here of the weight of criticism in Chairs' questionnaire responses; indeed the chair of one of the two panels represented in both lists described the cross-referral arrangements as a 'nightmare' in the questionnaire. One Panel Overview considered that cross-referral process worked well (UoA 29 Electrical & Electronic Engineering), another that the standard of research activity in the cross-referrals received to be broadly comparable in range to that in submissions made directly to it (UoA 5-8 Pre-Clinical etc.), and a third that the recommended outcomes received back from UoA 43 Business & Management Studies were similar to those submitted to the Panel but with slightly less at international level (UoA 38 Economics). Two Panel Overviews contain more significant points:

The panel also received and assessed many items cross-referred by other panels. In some cases the panel felt that it would have been more appropriate had these items been returned under History by the institutions concerned. In a few cases, the Panel found that it did not have the expertise that other panels assumed it possessed. This was so, for example, with some works in mathematics, and with work written in Welsh on recent and contemporary history. (UoA 59 History)

The chairs of Panels 64, 65 and 66 co-ordinated their activities and agreed that there had been some dispersal of work on film, television and video as a result of institutional choices. They concluded that some consideration needed to be given to how this dispersal across the panels might be handled in future. (UoA 65 Communication, Cultural & Media Studies)

183. We cannot tell, without examining the Team's logs, whether the arrangements for cross-referral were bound to be inefficient, within the accepted constraints, or whether more resources in the RAE Team would have made enough difference to avoid the complaints. If this element is likely to be a feature of a future RAE with a similar configuration of UoAs, there should be an analysis of the flows of cross-referrals, as it may pinpoint deficiencies in the configuration of UoAs and in the composition of individual panels.

9.14. Non-UK-based advisers

184. Our question to Panel Chairs whether the mechanisms (beyond the grade definitions) to promote comparability and consistency across UoAs (Umbrella Groups, non-UK-based advisers) were effective. Significantly, slightly more, 37%, thought not, compared with 29% thinking they were effective, but with as many as 35% uncertain. The feedback specifically on non-UK-based advisers is found in the comments. They were mostly negative (15 vs 3). 'Waste of time' occurred several times. There were suggestions that non-UK experts were not properly briefed, had too little understanding of the criteria and process (4), were often out-of-date on UK research, needed closer integration (4) into the process whereas previously they had been panel members, had too little time to do the job (3), were variable in the quality of their advice (2), were under-paid (2), and relied too much on their existing opinions about individual or departments reputations (3).

185. Chairs' opinions minuted from the Operational Review Meetings were more measured:

Views on the usefulness of non-UK-based advisors were mixed. Many panels reported that the advice had a positive effect on the assessment process, in that it confirmed the panel's decisions. In some cases non-UK advice had changed a panel's final grade. Non-UK advice seemed to be most useful where a panel had asked specific questions of the advisor. Many panels felt that non-UK advisors had not been adequately briefed on their role, and as such the quality of advice was mixed. Much greater thought to the aims and role of non-UK advisors needed to be given in future exercises.

186. Virtually all in the Overview reports are confined to saying that the non-UK views confirmed the Panel's judgements about grades 5 and 5*, and the 4/5 borderline. Some qualified this with phrases such as 'supportive', 'broad agreement', 'large measure' and 'no major inconsistency.' Only five panels made more evaluative comments (UoA 9 Pharmacy, 40 Social Policy, 53 Italian, 59 History and 63 Theology). Again, there is no inkling here of the weight of criticism in the questionnaire responses; of the two panels appearing in both lists, one chair made similar comments in the questionnaire, but the other used the phrase 'waste of time' in the questionnaire. Three Overview Reports found the non-UK-based advisers 'useful', 'helpful' and to have 'worked well'. Two Panel Overviews contain more significant points:

The non-UK advisers, who were given an awkward task, as they were not expected to read research outputs (in fact were not allowed the time to do so), could only give generic judgements, but were largely in agreement with our provisional grades. (UoA 53 Italian)

The use of overseas experts to review the Panel's initial grades was an interesting innovation.... The scholars consulted were neither asked, nor were they in a position, to duplicate the work of the panel; in particular they had no time to inspect the research outputs. Some of them wished to make use of criteria that the panel explicitly did not use, notably the proportion of research-active staff, the overall range of research topics covered by a department, and the number of monographs as compared to other types of publication. (UoA 59 History)

9.15. Umbrella Groups

187. Comments concerning Umbrella Groups (32) considerably outnumbered those on non-UK experts. They were fairly equally split between those which were positive (10), negative (14), and equivocal. However, the positive comments were about being ‘helpful’ or ‘useful’, and only one said they ‘worked effectively’. Analysing the equivocal comments revealed that many combined opposite views, for example: ‘Useful but of limited impact ‘ and ‘the umbrella group meetings were valuable occasions for clarification of many issues. Whether they promoted comparability and consistency is less clear.’ One Panel Chair suggested that systematic use of sub-panels shared amongst cognate UoAs would probably help.

188. Where there were real discrepancies amongst panel criteria, practices or grading profiles, the Umbrella Group process had no force (9). For example:

This proved to be the most serious problem. Panels operated and applied their criteria differently, resulting in widely divergent profiles of ratings.

A major discrepancy still remains, in which some panels grade according to groups of people, and others, like ours, grade in terms of individuals. I believe that the latter system is going to produce lower grades overall.

The meeting did not focus on arriving at comparability and consistency across UoAs but simply noted the different distributions. It was clear there were three extreme cases but there was no attempt to discuss changes to bring these into line. After the meeting some panels revised their distributions up, I think because they had seen the higher ratings of other panels but this was not consistent.

Umbrella meetings were moderately useful except that there was a failure to ensure that there actually was a similar distribution of grades across similar panels.

At the Umbrella Group meeting some Panel Chairs lost their nerve when they saw others’ inflated grades and inflated in their turn (while the inflaters went away and deflated theirs.) It was astonishing to see Chairs changing grades which had been arrived at over several months at the very last minute under peer pressure and this was certainly the worst aspect of the exercise for me.

RAE Team tried to promote comparability and consistency, but I am not sure they succeeded. Even when it became clear at our Umbrella Group meeting that panels had been operating quite different criteria (some more or less ignored RA5 and RA6, interpretations of international and national excellence varied significantly) no lead was given by RAE Team as to what was to be done to iron out these inconsistencies.

189. The timing of the Umbrella meetings was criticised by several Panel Chairs (6), for example :

The Umbrella Group meeting came too late (early October 2001) for effective comparability. By that stage panels had done their grading and sent out to the International Experts a sample of submissions. To expect wholesale changes to that grading as a result of the Umbrella meeting was to expect too much. The meeting should have been held after the initial reading of the outputs and submissions and before gradings had become too set in stone (i.e. early September).

190. The four panels whose Overview Reports made evaluative comments (UoA 9, 29, 32 and 53) were positive about Umbrella Groups, finding it useful in calibration between different UoAs (2), and enabling them better to delineate and review grade definitions and boundaries (2). There is no indication of the negative comments which outnumbered the

positive in the Chairs' questionnaire responses. The one panel common to both lists made similar comments in both places. One Overview Report made more significant points:

Umbrella panels were valuable for setting our assessments with a frame of reference in cognate disciplines....In particular, the ranges of the weightings for RA2, RA3, RA4, RA5 and RA6 were constrained by the Umbrella group and the Manager.... The Umbrella panel meeting clearly indicated that this UoA was marking on the hard side so the panel took this into account when arriving at its final grades. (UoA 29 Electrical & Electronic Engineering)

191. It is possible that in writing Overview Reports and replies to us, Chairs were unduly mindful of the last phase of the Umbrella Groups' work and forgetful of a greater utility when they were drafting the criteria.

192. The British Academy [RAR] believes that the Umbrella Groups should have a larger role in future in standardising practice and ensuring consistency.

9.16. Interdisciplinary research

193. There has been continuous controversy around the RAE's capability to do justice to interdisciplinary research. We asked Panel Chairs: 'Were the arrangements and criteria appropriate to enable the panel to give interdisciplinary research proper consideration?', and their response was emphatically in the affirmative: 34% saying 'Strongly yes' and 50% 'Yes'. The one Chair saying 'No' did not elaborate. Of the 21 who did comment, a clear majority claimed 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. Other comments were:

the problem is when work falls into cracks between disciplines rather than being genuinely cross-disciplinary.

(from an Area Studies Chair) we would have appreciated more guidance on handling interdisciplinary research. We did our best, and genuinely tried to be fair to it, but it took more time to evaluate, and time was at a premium.

(from a lightly-loaded Panel) there were some difficulties to be overcome where the two disciplines were very different in the minuteness of their examination.

194. This confidence in the equitable treatment of interdisciplinary work is not as widespread in the institutions. Admittedly, only 82% of Contacts felt able to answer the question, 'Did the RAE arrangements enable interdisciplinary research to be submitted appropriately?', and 38% of those were uncertain; but 36% said 'No' or 'Strongly no' as against only 25% 'Yes' or 'Strongly yes'. Several of the longer comments capture points made by others and warrant quotation:

We felt that the treatment of some of at least one of our units which mainly produces interdisciplinary work demonstrated the problems of the system very effectively. The problems lie in panel coverage both in terms of the expertise available or called on, and the perceived UoA boundaries. In addition, from a purely technical point of view, it is not possible to represent the activity of a member of staff who is formally contracted to work across more than one UoA effectively as s/he can only count as category A in one of the UoAs (as opposed to category A in both and 50% fte in one and 50% fte in another). We had a number of appointments like this (intended to develop interdisciplinary working)

and this resulted in endless debates over who would 'own' the member of staff concerned.

The informal feedback that we have received suggests that work cross-referred to other panels was often assessed at a lower grade than the rest of the submission. Interdisciplinary work was seen as achieving excellence in neither discipline. A detailed study is needed to see if this pattern was reflected across the RAE. Informal feedback also indicated to us two cases where the positive outcomes of a cross-referral were ignored by the panel that had referred the work.

The process is now so competitive that research that did not fit squarely within one UoA or another was perceived as being a risk and in danger of contributing to the tail. In most cases where there was any chance that the research would be assessed by those who did not fully understand its nature, it was not submitted.

There is much work to be done to tackle this conundrum. Perhaps the issue is not so much about arrangements as about attitudes towards i-d research by Panel members assessing from within their definition of mainstream research – i-d research can be very strong whilst still at the margins of the consciousness of researchers who sit firmly within one discipline.

No, we were not convinced that inter-disciplinary research was recognised as relevant to particular UoAs, nor that the cross-referral system allowed its assessment on equal terms with mono-disciplinary submissions. In several instances, a large number of words in RA5 and 6 were dedicated to explaining the inter-disciplinary nature of activity; yet less weight appeared to be given to these textual parts of the submission than we had anticipated.

No: although it was possible to draw attention to interdisciplinary research the arrangements dictated that research was submitted to a single panel in the first instance, even though it might be cross-referred to others during the assessment process. This helped to create a strong impression that interdisciplinary research did not essentially fit the basic structure of RAE (and there is evidence that this structural bias sometimes influenced the evaluation and eventual rating of interdisciplinary research).

... Concern also for subjects that are interdisciplinary in nature, e.g. Development Studies, where a sub-panel was created but the criteria was driven by that of its host panel [Geography] rather than a separate set of criteria.

The CBI [RAR] commented in similar terms:

Much cross-disciplinary work ends up being assessed in an inappropriate way—this is a serious failing considering that innovation is most likely to take place at the interfaces between subjects and where expertise from many disparate subjects can be brought to bear on problems. Everything possible must be done to help break out of the traditional research silos in universities.

195. Elsewhere (Rogers 8.3) the point has been made that panels' understanding of journals' refereeing and editorial standards must have influenced their assessment of outputs not read but reviewed - with the relative standing of journals being one way by which the standard of international excellence was operationalised. But for interdisciplinary work (and for emerging fields) the hierarchy of journals may be less well defined than for established fields, and conversely the high-prestige journals in established fields may not accept interdisciplinary work.

196. No one offered the opinion that RAE 2001 was better at handling interdisciplinary research than RAE 1996. After the latter, the funding bodies commissioned a large study from *Ea* Evaluation Associates Ltd on 'Interdisciplinary research and the Research Assessment Exercise' (March 1999, issued as RAE 1/99). Two of its main messages were that:

1. Departments and researchers widely believe the RAE inhibits interdisciplinary research. Many institutions and individuals are changing the way they do research as a result of the RAE.
2. Despite these beliefs, there is no evidence that the RAE systematically discriminated against interdisciplinary research in 1996.

Our (limited) evidence suggests that many in institutions still believe the first and that panels believe the second. Whether perceptions have changed since the 1996 RAE we cannot tell.

197. The *Ea* report made five specific recommendations which were accepted by the Funding Bodies:

1. The Funding Bodies and Panel Chairs should take steps to ensure the representativeness of panels across all forms and areas of research, including interdisciplinary research.
2. Panel methodologies should specifically address assessment of interdisciplinarity.
3. New mechanisms for boundary critical submissions should be introduced based on a new proforma (5C).
4. The Funding Bodies should introduce monitoring mechanisms to ensure the effectiveness of cross-referral.
5. Feedback and reporting should embrace interdisciplinary research.

198. We have not attempted to establish how each of these was implemented and then to assess its effectiveness. 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.

9.17. Collaborative research

199. We asked Panel Chairs, 'Did the assessment mechanisms disadvantage research, of which evidence was submitted, which had been undertaken collaboratively with partners within and beyond higher education?'¹ 41% replied 'Strongly yes' and 36% 'Yes', i.e. over three-quarters of the Panel Chairs responding were of the view that collaborative research was not disadvantaged. The comments came only from Chairs of social science and humanities panels, pointing to the requirements of collaborating funders or a 'popular' audience possibly meaning that the outputs were less polished or 'academic' than work appearing through conventional scholarly channels; and to the need for detailed guidance as more research in the humanities became collaborative.

200. The Contacts were asked: 'Did the RAE arrangements enable collaborative research to be submitted appropriately?' Nearly 20% did not respond, and of the rest 42% were uncertain, but 'Yes' and 'Strongly yes' at 34% outweighed 'No' and 'Strongly no' at 24%. It is possible that the question was read to refer only to collaborative research which was the subject of a joint submission by two or more institutions (of which there were few), whereas the question to Chairs more clearly referred to all collaborative research.

201. The comments picked up issues which should be addressed for any future RAE, as collaboration between institutions is being encouraged:

- The arrangements are unsatisfactory when the collaboration is between groups in different

¹ The negative wording clearly confused some Chairs whose 'Yes' answers were, from the comments, meant to be 'No' and have been amended. The other four 'Yes's' were without comments.

institutions in different UoAs - yet this is likely to be (increasingly) the pattern of interdisciplinary collaboration.

- The competitive character of the RAE mitigates against collaborative research and though the arrangements for the submission of collaborative research were satisfactory they were grievously underused. 'On the one hand, funders are increasingly urging us to collaborate; on the other hand, the RAE rules are generally perceived to discourage inter-institutional research projects.'
- Joint submissions between 'unequal' partners is seen as high risk, i.e. if one institution is small or does not have a strong research track record.
- Where all the staff were employed on joint contracts with the two partner institutions, each person had to be split, artificially, between two separate submissions.
- It is very unsatisfactory that non-administering institutions cannot claim the research income that was associated with their involvement in joint projects, other than to make mention of it in RA5/6.

202. Two comments in RAR submissions related to applicable research. The Joint University Council for Social Administration, Social Policy and Social Work Education asserted that the outcome judgements of the RAE indicated that collaborative research was not as highly valued as other forms, referring particularly to small-scale local projects involving service providers and to work disseminated to a practitioner audience. The Department of Health offered a wider ranging conclusion: 'The RAE has taken a narrow view of the research contribution of individuals, which are aggregated to the level of a single institution but no further. The result has been a disincentive to the development of those multidisciplinary and multi-site networks that are best fitted to tackle really substantial problems in health care.'

203. Some Contacts raised the issue of duplicate citations arising from collaboration within a UoA, e.g.: 'Thus, if two colleagues in the same department jointly write four pieces, the panel is likely to comment unfavourably on repetition. If one of those colleagues were to be submitted to another UoA or to work in a different institution but be submitted to the same UoA, repetition would not be an issue.'

204. In summary, the arrangements are seen to be satisfactory where the partners decide to make joint submissions; the true extent of collaboration may not be revealed; and the RAE may operate against the policy push for greater collaboration. Therefore further attention to assessment in this area is needed.

9.18. Research into teaching and learning

205. In its RAR submission, the Institute for Learning and Teaching stated:

Those committed to understanding and improving the professional practice of teaching in higher education, including ILTHE's members, report that they are discouraged from undertaking research to improve their teaching practice because of the perception, accurate or not, that such applicable research into academic practice is not valued alongside other forms of research.

The Staff and Educational Development Association in its submission pressed for the creation of an additional Unit of Assessment and for nominees to each panel or sub-panel with expertise in research into teaching and learning in that discipline.

206. This was said, notwithstanding the study of the pattern of 'pedagogic research'

submitted to the 1996 RAE as part of the preparations for 2001. (The term should perhaps be avoided for the future, given the age profile of h.e. students.) In consequence, all panels stated that they would accept such research in their field and would assess it on a par with other work. We did not solicit any evidence on this point and have noted only three comments in Overview Reports:

UoA 48 European Studies: There were some differences in the patterns of work submitted in 2001. The most striking difference was the sharp fall in the proportion of pedagogical research that the Panel had to handle compared to 1996 and 1992. It is possible that more of this work found its way to other Panels that were seen as more appropriate. Equally, this decline may reflect a greater caution on the part of universities in returning this type of work, perhaps prompted by the criticisms levelled by Panels after the RAE in 1996. It would be a matter of regret if the feedback after the last RAE had led to an exclusion of pedagogical research in RAE 2001 when what we sought was an improvement of research in this field.

UoA 51 French: We were impressed by the overall quality and range of what we read, but would wish to express some concern over certain recurrent problems. For instance, some items submitted were broadly pedagogic in character. Where such items clearly embodied research outcomes, they were rated accordingly, but assessment was more difficult when they seemed rather to represent a resume of existing scholarship, and – despite our explicit request – not all submissions clearly specified the research element in this type of publication.

UoA 64 Art and Design: As there was little evidence of pedagogic research, it was assumed that much of this work had been submitted to [UoA 68 Education]. In general, that work that was submitted to Art and Design reflected the pedagogy of teaching and course development, rather than the pedagogy of learning.

UoA 68 Education's Overview Report contained no reference to research in teaching and learning in higher education.

10. After the assessment

10.1. Overview Reports

207. In response to the question, 'Do you consider that the subject overview report has proved valuable to your subject community and beyond?', nearly half the Chairs responding were uncertain, mainly for lack of feedback from their subject communities. (It was intended that the reports should be printed in one volume, but given the large audience, each reader with an interest in a small part, this might not have added much to web publication.) Though only 11% went as far as to say 'No', some of the higher scores were accompanied by fairly negative comments, e.g., 'More a chore than a useful service'; 'Limited value, but keep'. Others noted that panels were not supported in preparing the reports, e.g. by funding for another meeting; if the Team's guidance had been followed to the letter, the report would have been largely platitudes; and if all panels had done a good job, then the Research Councils might be using the reports as a resource. Two chairs had concerns that mistaken messages might be taken from reports.

208. For the future, repetition of the same is not an option: either Overview Reports are dropped or more effort, with more licence, is put into them.

10.2. Feedback to institutions

209. In response to the question, 'Given the constraints on what might be said, do you think the provision of feedback to institutions was worthwhile?', a similar proportion of Chairs replied in the affirmative as on the Overview Reports, 41%, but fewer 'uncertains' increased the 'no's' to 27%. The most positive comments were on the line: 'some value, needs to be seen to be done, not much effort.' Other complained of the feedback's potential utility being emasculated by fear of judicial review, of the lack of time and secretarial support and of insufficient clarity of purpose.

210. We asked Contacts: 'Was the feedback from panels on your submissions helpful?' No other question elicited so low a percentage of positive responses from them, at 23%. One Contact encompassed the points made by many: 'Highly variable. Some was useful (e.g. from UoA 41/42 [?40/41 Social Policy/Social Work]); much was anodyne in the extreme. As a Panel Secretary I fully understand why. However, if feedback is to be provided in the future, more thought MUST be given to how it can be useful without being potentially litigious.' UoA's 33 Built Environment, 34 Town & Country Planning and 60 History of Art were also commended, by an art college.

211. Again and more strongly, for the future, repetition of the same is not an option. The implications of the Freedom of Information Act 2000 must be considered: it may well not be possible to find any exemption in the Act which would protect the workbooks from disclosure. The funding bodies can note that no challenge to either the 1996 or the 2001 RAE has gone to judicial review.

212. Two panels in their Overview Reports, UoAs 60 History of Art and 67 Music, commented on the facility in the generic criteria for formal 'flagging' work within a submission of much higher standard than the rest. Neither felt it would be appropriate, and

preferred instead to mention it in feedback to the HEI. We have examined panels' use of flagging.

10.3. Complaints about the outcome

213. 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 application for which was withdrawn on the eve of the hearing). The formal position seems to be that an institution should complain to the body from which it receives grant; but inevitably any detailed work has to be done from the RAE records in Bristol. The ragged disbanding of the RAE Team meant that there were inadequate arrangements at Northavon House for dealing with complaints, giving rise to the solicitors' advice as follows:

There is a need for a clear protocol for the handling of responses/queries/challenges to the ratings or results of any assessment exercise. Such a protocol should cover what to say to initial enquirers about the available avenues of appeal or challenge, the timetable for responding to queries/challenges, what information should be sought from complainants who indicate they wish to complain/challenge, what documentation and information should be obtained from Panel Chairs and others in [the funding bodies] to facilitate the drafting of a response, and a policy on when and how to obtain legal advice. The protocol should be shared with Panel Chairs and Panel members before the Panel disbands, as they need to be aware that their time and input might be required after the publication of results and feedback to enable [the funding bodies] to respond to queries and challenges, and to understand why such responses are required.

10.4. Matching the key principles: institutions' assessment

214. We asked Contacts for an assessment of the RAE's achievement against each of the key principles laid down at the outset. The results can at best be regarded as a straw poll: the high proportion of scores of 3 (Adequate: neither good nor bad) may represent 'don't know', respondents not having consulted RAE 2/99 for the explanation of the principles. Even so, the comparative proportions of 'Yes' (3.5 to 5) and 'No' (1 to 2.5) are instructive:

RAE 2/99, para. 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 on which 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. [Several elements in the RAE process bear on this important issue and they are listed under the

‘General conclusions’ in the Executive Summary.]

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.

Nine responses from Contacts made comments explicitly on consistency, of which one was ‘adequate’, and the others negative. ‘There were inconsistencies in approach and standard between panels that were not justified in subject terms.’ Assessment was not always consistent across cognate areas, related panels used different methods and criteria, and interpreted grade definitions and guidelines differently.

Annex 1 Email and questionnaire of 13 January 2003 to Panel Chairs

Dear RAE Panel Chair,

Tom Sastry's message of 8 January explained that the HEFCE has asked Universitas Higher Education Management Consultants to undertake a review of the operation of the 2001 Research Assessment Exercise. Our task is to assess the performance of the 2001 RAE against its overall objective of providing ratings of quality of research conducted in HEIs. The behavioural, funding and other impacts of the RAE, and alternative approaches to the assessment of research, fall outside our remit. Our focus is on the managerial and operational, rather than on the philosophical or on policy - the latter being the concern of the Research Assessment Review (RAR) in progress under Sir Gareth Roberts's leadership.

Our report nevertheless will be fed into the RAR's deliberations, and our timetable has been determined by the RAR's. We have to produce our report in late February. So what we can do in collecting evidence is limited. We are reviewing the published documents, some of the working files, and those RAR submissions which comment on the 2001 RAE's operation; and we are interviewing a few of the officials involved. We are conducting two questionnaire surveys: this one, addressed to Panel Chairs, and a second addressed to the 'RAE Contacts' in the universities and colleges which made submissions. We will appreciate replies by **FRIDAY 31 JANUARY**.

We will be very grateful if you can complete the attached questionnaire, 'RAE_chairs.doc'. We are asking for your personal responses, from your experience as Panel Chair, though you may of course, if you wish, consult fellow members or advisers (perhaps giving priority to the sub-Panel Chairs). Your reply will be confidential to the Universitas team preparing the report. We will need to know which panel you chaired, to put some replies in context. But we will hold your replies in confidence and will not pass them to HEFCE or the other funding bodies, nor will we identify panels in drawing on them in our report.

Working with the Research Policy Team at HEFCE, we have drawn up the questions in the attached document. We have kept their number down by concentrating on issues of likely concern to the RAR and by grouping aspects of the RAE process together. You are asked, as a minimum, to assign a score in answer to each question. Those scores taken from all the replies will at the very least identify areas for further investigation. Under each question, please add any comments you wish.

Please return your answers by saving the document (without changing the version of Word, if you are using Word) and attaching it to an email addressed to <farrant@universitas.co.uk>, e.g. as a reply to this message.

Yours sincerely,

John Farrant

UNIVERSITAS

Higher Education Management Consultants

PO Box 5060 Milton Keynes MK6 3ZX United Kingdom

RESEARCH ASSESSMENT EXERCISE 2001: OPERATIONAL REVIEW

QUESTIONNAIRE TO PANEL CHAIRS January 2003

Please return to John Farrant, attached to an email addressed to [<farrant@universitas.co.uk>](mailto:farrant@universitas.co.uk), by
FRIDAY 31 JANUARY 2003

Please answer each question using the one of these two five-point scales appropriate to it:

5	Strongly yes	Excellent
4	Yes	Good
3	Uncertain	Adequate; neither good nor bad
2	No	Poor
1	Strongly no	Bad

If you do not have an opinion, enter DK or Don't know.

Enter the appropriate digit in the box to the right of the question. Enter any comments in the box beneath the question. Please use only blank, unshaded boxes and do not add or delete boxes.

Do not feel that you should find something to say in each comment box. Reserve comments for issues on which you feel particularly strongly or on which you have an important insight to add. If you consider comments relevant to more than one question please indicate where this is the case. If you are really pressed for time, we will still appreciate having only your numerical scorings: aggregated across panels, they will flag issues calling for further examination.

Serial number(s) of UoAs covered:	
Title of panel:	

Question number	Question	Score
	The Panel	
1	Were the procedures for the nomination and appointment of Panel Chair and members appropriate?	
comments>		
	Statement of criteria and working methods	
2	Did the processes involving the generic criteria and working methods (RAE 2/99 and the Guidance to Panel Chairs and Members), which shaped your Panel's statement, allow the Panel to make judgments appropriate to the field and consistent with related fields?	
comments>		
	Information in the submissions	
3	Were the quantitative data, and their presentation, appropriate and sufficient?	
comments>		
4	How valuable were RA5 and RA6 in reaching the Panel's judgements?	
comments>		
5	Were you able to take appropriate account of 'staff circumstances' as set out in RA6?	
comments>		
	Provision to panels of information additional to that in the submissions	
6	How appropriate, in aggregate, were the arrangements for sub-panels, specialist advice, cross-referrals and sourcing research outputs?	
comments>		
	Administrative support	
7	Panel secretariat: did the secretary and assistant secretary have enough time, training and support to give the level of service required?	
comments>		
8	The central RAE Team: assess the efficiency of the services it provided. Please comment on individual services (meeting arrangements, audit requests, sourcing outputs, ad hoc requests for statistics, etc.)	
comments>		

9	Did the funding bodies and the RAE Team do as much as reasonably possible to provide information to the sector about the workings of the RAE?	
comments>		
Assessment phase		
10	Was the workload on panel members manageable?	
comments>		
11	Did the generic grade definitions and the advice on their application make the grading system readily workable?	
comments>		
12	Were the mechanisms (beyond the grade definitions) to promote comparability and consistency across UoAs (Umbrella Groups, non-UK-based advisers) effective?	
comments>		
13	Was the panel's decision-taking enhanced by access to expertise and advice of research users?	
comments>		
14	Were the arrangements and criteria appropriate to enable the panel to give interdisciplinary proper consideration?	
comments>		
15	Did the assessment mechanisms disadvantage research, of which evidence was submitted, which had been undertaken collaboratively with partners within and beyond higher education?	
comments>		
After the outcomes were published		
16	Do you consider that the subject overview report has proved valuable to your subject community and beyond?	
comments>		
17	Given the constraints on what might be said, do you think the provision of feedback to institutions was worthwhile?	
comments>		

Are you willing for us to follow up your response, by telephone or email?	YES/NO
If yes, please give the telephone number at which we should contact you.	

THANK YOU VERY MUCH FOR COMPLETING THIS QUESTIONNAIRE. Please return it to John Farrant, attached to an email addressed to farrant@universitas.co.uk , by FRIDAY 31 JANUARY 2003

Annex 2 Email and questionnaire of 14 January 2003 to RAE Contacts

Dear RAE Contact,

The joint funding bodies' Review of Research Assessment in Higher Education (the RAR) is in progress, steered by the group chaired by Sir Gareth Roberts. Several studies are being conducted in conjunction with it, and the Higher Education Funding Council for England has commissioned Universitas to undertake one of them. This a review of the operation of the 2001 RAE. Its aim is twofold: to provide, for the sake of accountability, an independent appraisal of RAE 2001, and to identify lessons for the conduct of a successor exercise. The behavioural, funding and other impacts of the RAE, and alternative approaches to the assessment of research, are outside its remit, being the province of the RAR itself.

Our report will be fed into the RAR's deliberations, and our timetable has been determined by the RAR's. We have to produce our report in late February. So what we can do in collecting evidence is limited. We are reviewing the published documents, some of the working files, and those RAR submissions which comment on the 2001 RAE's operation; and interviewing a few of the officials involved. We are conducting two questionnaire surveys: this one, addressed to the RAE Contacts in the universities and colleges which made submissions, and a second addressed to the Panel Chairs. We will appreciate replies by FRIDAY 31 JANUARY.

We will be very grateful if you can complete the attached questionnaire, 'RAE_contacts.doc' (in Word 7). We are asking for your personal observations, from your vantage point as RAE Contact, rather than an institutional response formulated after extensive consultation. But we are asking you to speak also for the Data Contact (if that was a different person). Please also include an assessment of how colleagues in academic departments coped with the operational aspects of the RAE, for example, whether the Guidance on Submissions was clear and whether it enabled all your submissions to be effective. If your institution named two contacts, this message is going to both, but a joint response will be appreciated.

Your reply will be confidential to the Universitas team preparing the report. We will need to know your institution, in order to put your replies in context (e.g. whether you made few or many submissions). But we will maintain the confidentiality of the views you express and will not pass replies to HEFCE or the other funding bodies, nor will we identify institutions in drawing on the replies in our report.

We have drawn up the questions in conjunction with the Research Policy Team at HEFCE. We have kept their number down by concentrating on issues of likely concern to the RAR and by grouping aspects of the RAE process together. You are asked, as a minimum, to assign a score in answer to each question. Those scores taken from all the replies will at the very least identify areas for further investigation. Under each question, please add any comments you wish. If you consider comments relevant to more than one question, please indicate where this is the case. Please return your answers by saving the document (without changing the version of Word, if you are using Word) and attaching it to an email addressed to <farrant@universitas.co.uk>, e.g. as a reply to sender (not all recipients) to this message.

Yours sincerely,

John Farrant

UNIVERSITAS

Higher Education Management Consultants

PO Box 5060 Milton Keynes MK6 3ZX United Kingdom

RESEARCH ASSESSMENT EXERCISE 2001: OPERATIONAL REVIEW

QUESTIONNAIRE TO RAE CONTACTS January 2003

Please return to John Farrant, attached to an email addressed to [<farrant@universitas.co.uk>](mailto:farrant@universitas.co.uk), by
FRIDAY 31 JANUARY 2003

Please answer each question using the one of these two five-point scales appropriate to it:

5	Strongly yes	Excellent
4	Yes	Good
3	Uncertain	Adequate; neither good nor bad
2	No	Poor
1	Strongly no	Bad

If you do not have an opinion, enter DK or Don't know.

Enter the appropriate digit in the box to the right of the question. Enter any comments in the box beneath the question. Please use only blank, unshaded boxes and do not add or delete boxes.

Do not feel that you should find something to say in each comment box - which will enlarge as you type. Reserve comments for issues on which you feel particularly strongly or on which you have an important insight to add. If you consider comments relevant to more than one question please indicate where this is the case. If you are really pressed for time, we will still appreciate having only your numerical scorings: aggregated across panels, they will flag issues calling for further examination.

Name of institution:	
----------------------	--

Question number	Question	Score
	The Panel	
1	Were the arrangements for nominating Panel members appropriate?	
Comments>		
	Guidance on Submissions (RAE 2/99)	
2	Was the Guidance clear and timely, to you and colleagues?	
Comments>		
	Statements of Panels' criteria and working methods	
3	Were the arrangements for consultation on the Panels' draft criteria and working methods appropriate?	
Comments>		
4	Did the statements of criteria and working methods present any problems, while submissions were being prepared?	
Comments>		
	Preparing the submissions	
5	Was the supplementary guidance, in the Briefing Notes and website, clear and timely?	
Comments>		
6	Were the briefing seminars helpful?	
Comments>		
7	How do you rate the assistance the RAE Team gave in response to ad hoc enquiries?	
Comments>		
8	Did the structure and processes of the RAE enable your institution to make effective submissions?	
Comments >		
9	Did the RAE arrangements enable interdisciplinary research to be submitted appropriately?	
Comments >		
10	Did the RAE arrangements enable collaborative research to be submitted appropriately?	
Comments >		
	The RAE Data Application	
11	In which category do you place your institution's use of the	

	RDA? Enter A, B or C in the box to the right. A RA tables completed centrally, after data collected using own tools B RDA mounted centrally and networked to departments for completion C RDA distributed for 'stand-alone' use by departments	
Comments>		
12	Did you automatically populate any parts of the table from your own databases (as distinct from re-keying or manually copying and pasting)? Enter Yes or No.	
Comments>		
13	Did you encounter technical problems in using the RDA?	
Comments>		
14	Overall, how do you rate the performance of the RDA, including the documentation and other forms of support?	
Comments>		
	After the submissions were in	
15	Was the process of data checking and verification carried out without imposing unreasonable demands?	
Comments>		
16	How well did the Research Outputs Sourcing System function from your perspective?	
Comments>		
17	Was the feedback from panels on your submissions helpful?	
Comments>		
	The key principles	
18	How do you rate achievement of each, as described in RAE 2/99?	
a	Peer review	
b	Clarity	
c	Consistency	
d	Continuity	
e	Credibility	
f	Efficiency	
g	Neutrality	
h	Parity	
j	Transparency	
Comments>		

Name of the person completing the questionnaire	
Are you willing for us to follow up your response, by telephone or email?	YES/NO
If yes, please give the telephone number at which we should contact you.	

THANK YOU VERY MUCH FOR COMPLETING THIS QUESTIONNAIRE. Please return it to John Farrant, attached to an email addressed to farrant@universitas.co.uk , by FRIDAY 31 JANUARY 2003

Annex 3 Summary and recommendations from *Data collection for the Research Assessment Exercise 2001: A review by Ewan S. Page (2001)*

Summary

Judged by the results of the operation, the data collection for the Research Assessment Exercise 2001 (RAE 2001) was even more successful than that in 1996. The institutions were able to return their submissions promptly, indeed most a little earlier than required, and the Higher Education Funding Council for England (HEFCE) processed the data and distributed material to panellists a week or more before the target date.

The Data Collection Steering Group (DCSG) established by HEFCE contributed usefully to the specification, construction and testing of the software, and so to the improved experiences with the software of users in the institutions. Most of the changes in software and in procedures introduced since RAE 1996 were welcomed but some brought new faults or new difficulties for the institutions. Although the primary aim for a future RAE must be to satisfy the needs of the assessment panels more attention should be given to easing the task of making the submissions for institutions. The software must be made as error-free and as robust as possible, even at the expense of some of the desirable but not essential features; a DCSG strengthened by more computing professionals could play an important part in achieving this.

Every effort should be made to avoid changes in the types of information sought, the definitions or the significance that may be attributed to particular items (as indicated by Notes of Guidance to Units of Assessment) to be placed upon them. Any changes, apart from abandonment of some element, are likely to cause much work for little, if any, advantage and may lead to ambiguity or even conflict.

The RAE has developed into a mature and efficient process for evaluating the research achievements of institutions and has provided the basis for distributing public funds for research which has demonstrably attained the objectives of improvement in quality and quantity.

Panels should be further assisted by clear comparative summaries of the statistical information with the first distribution of material. Software for this purpose should be constructed ready for use on receipt of the submissions.

Recommendations

The following recommendations are presented in the order that action is needed if they are to be adopted. Some require action immediately, before staff are assigned to other tasks; some quickly before panels are dispersed; some immediately after that; others as HEFCE prepares well in advance for the next RAE.

1. In order to retain the option of modifying the RDA software used in 2001 for the next RAE, HEFCE should ensure that the internal documentation of the programs is sufficiently detailed.
2. HEFCE should prepare specimen summaries of comparative statistics so that the comments of the present assessment panels can be sought at their last meeting after they have completed their assessments but before they disband.
3. At this meeting panels should be asked to suggest what information they have received that their successors could do without, what different presentation of existing data would make their task easier and what other information would be very important for their successors to have. The panels should be assured that their views, especially any which lead to a reduction of the burden of recording on the institutions, are very valuable.
4. At this meeting panels should be asked to review the notes of guidance they produced for their Units of Assessment and to draw attention to any elements which need amendment in a future exercise, bearing in mind that the fewer changes there are the better.
5. HEFCE should review the suggestions of the panels immediately, adopting no changes requiring more data, more detailed data or data in a different form unless there are the most compelling reasons for doing so.
6. HEFCE should announce their decisions about the data as soon as possible so that institutions can adjust their recording practices if necessary. HEFCE should aim to make this announcement by Spring 2002.
7. HEFCE should review their experience of handling the publications requested by panellists, their estimates of total and peak traffic and record their findings for the guidance of those planning the next RAE.
8. HEFCE should plan for the next RAE on the assumption that special software will be needed to collect the data and distribute it quickly to panels and that it will be constructed in-house, whether or not the RDA is to be adapted or a fresh start made.
9. No less time should be allowed for the construction, full documentation and testing of software.
10. A Data Collection Steering Group should be formed at the beginning of the planning stage and given the terms of reference of DCSG supplemented by 'To approve the detailed scheme for the programming and testing of the software'.
11. The next DCSG should contain two more computing specialists from institutions. The officer responsible for the RAE should be a member of the DCSG but it would be advisable to have as Chairman some other senior person.
12. DCSG should resist proposals for changes in the software whether arising from the panels or the institutions unless there are the most compelling reasons for them. DCSG should weigh carefully the options of basing the software on the RDA with only minimal

modifications or producing new minimalist software supporting templates for data entry or some combination which would cater for the range of institutions submitting.

13. DCSG should maintain the commitment to undertake preliminary tests of the software in their parent institutions as well as to participate in the sector-wide testing. It should satisfy itself that sufficient other institutions will take part to guarantee adequate coverage of the modes of operation and types of system that will be represented in the RAE itself.

14. The software tested sector-wide should not be altered except to fix errors.

15. HEFCE should review the capacity of panellists to have ready access to publications listed in the submissions in the light of the current technology in the run up to the next RAE and take this into account in the composition of the panels.