August 2006/32

Issues paper
Purpose

This report is for information

This study investigates how disability, age, sex and ethnicity are related to selection of staff for inclusion in the 2001 Research Assessment Exercise (RAE2001), and the possible reasons for the differences in selection rates found. It examines the question of whether the process of selecting staff was fair, or if some staff were disadvantaged.

Selection of staff for inclusion in RAE2001

HIGHER EDUCATION hefce FUNDING COUNCIL FOR ENGLAND

Acknowledgement

The bibliometric analysis reported here was based on data prepared by Evidence Ltd. Jonathan Adams and Louise Jackson also provided advice on interpreting these data.

Contents

Selection of staff for inclusion in RAE2001

To	Heads of HEFCE-funded higher education institutions
Of interest to those responsible for	Equality and diversity management, Human resources management, Institutional strategic planning, Research management
Reference	2006/**32**
Publication date	August 2006

Enquiries about HEFCE research and RAE policy	Paul Hubbard tel 0117 931 7334 e-mail p.hubbard@hefce.ac.uk
Enquiries about HEFCE equality policies	Yasir Mirza tel 0117 931 7316 e-mail y.mirza@hefce.ac.uk
Technical enquiries relating to data, modelling, etc	Mark Gittoes tel 0117 931 7052 e-mail m.gittoes@hefce.ac.uk

Summary

Purpose

1. This study investigates how disability, age, sex and ethnicity are related to the selection of staff for inclusion in the 2001 Research Assessment Exercise (RAE2001), and the possible reasons for the differences in selection rates found.

Key points

Background and scope

2. The RAE assesses the quality of research in higher education institutions (HEIs) in the UK through a process of peer review. This study is based on an analysis of data from the last completed exercise in 2001. Institutions were able to select the staff whose research outputs were to be included in this assessment from their 'eligible staff', that is those academic staff employed by the institution, who were not employed to carry out another individual's research programme.

3. There have been concerns about the possible impact of the RAE on equal opportunity policies. Is the process of selecting staff fair, or are some staff disadvantaged? The scope of our analysis for this study is limited to answering this question. We have not, for example, attempted to ascertain whether RAE panels assessed the work of different groups of academics fairly, or whether the process of accepting or rejecting an article is fair, or whether the research process as a whole is biased or not.

Methodology

4. To assess whether there were disadvantages for certain groups of academics in the process of being selected for inclusion in RAE2001, we first considered to what extent these groups of academics were associated with a 'department' that their HEI had decided to submit for assessment. We then took just those staff who were associated with these 'submitting departments' and, by using statistical models, tried to compare staff on a 'like-for-like' basis.

5. This statistical modelling is not straightforward, for two reasons:

   a. To give a definitive answer as to whether the selection of staff was fair, we would need a measure of research output quality, and we do not have this. Our only alternative was to rely on proxies for such a measure.

   b. Ambiguities are introduced into the analysis by including grade and other measures of employment status in the modelling. This is because the proportions of staff in different grades are also an equal opportunities issue. To help with the interpretation of the results we therefore present two sets of analysis: 'full' models, in which all the available variables are included and 'restricted' models which exclude variables relating to employment status.

We also used bibliometric data to assess the relative research strength of the men and women, and of staff from ethnic minorities, whose work was submitted to RAE2001.

Results and conclusions

6. The selection rate for staff with recorded disabilities was 58 per cent, slightly lower than for staff not so identified (59 per cent). When other attributes were taken into account, the analysis showed disability was not a significant factor in the propensity to be selected.

7. Both simple and 'like-for-like' comparisons showed that the proportion of staff selected varied significantly by age. The general pattern was that staff aged over 30 were more likely to be selected. This result suggested either that institutions were unwilling to select staff who had not had long enough to build up a substantial research record, or that, in general, researchers in the early parts of their careers might not produce outputs of the highest possible research quality.

8. There was a large difference in the selection rates for men and women, which were 64 per cent and 46 per cent respectively. These overall figures hid differences by age. When other factors are taken into account, including grade and other employment related variables, the 'like-for-like' comparisons still showed that men had significantly higher selection rates than women over the middle age range between 30 and 47. However, bibliometric measures of research strength of the selected staff showed no great differences between men and women if the most cited staff are excluded. Though not

conclusive, these results were consistent with an explanation of the lower selection rate of women being due to a lower proportion of women having a research record that leads them to be selected, rather than bias in the selection process.

9. The simple unadjusted comparisons showed selection rates of around 58 per cent to 60 per cent for staff from different ethnic groups, apart from staff from Black ethnic groups, who had a lower rate of 37 per cent. This lower rate was partly the result of a higher proportion of these staff being employed in departments which did not make an RAE2001 submission. But even when non-submitting departments were excluded, the selection rate for staff from Black ethnic groups was much lower than for others. The results from the two models used in the analysis gave different results when comparing staff by ethnic group. The 'full' model showed no significant difference in the selection rates of ethnic groupings when compared on a 'like-for-like' basis, while the 'restricted' model, which excluded employment variables showed that selection rates for 'Black', 'Asian' and 'Other' ethnic groupings were all lower than expected. We could not be certain whether this result from the restricted model reflected an unmeasured difference in quality of research, or whether it was evidence of an unjustifiable bias. However, evidence from the bibliometric analysis suggested that it was the former, the weakness of the proxies for research quality, rather than bias in selection. The bibliometric measures of research strength of the selected staff showed no great differences between researchers from different ethnic groups.

10. We have put in place comprehensive guidelines to ensure that, as far as possible, the next RAE in 2008 is fair, and seen to be fair, to all academic staff. However, these measures will not directly address any underlying factors which result in different groups of staff being less likely to produce research outputs of the standing expected for selection to the RAE.

11. To address these wider concerns we should look to other equal opportunities policies as the instruments for change. We, working with the Equality Challenge Unit and the Leadership Foundation for Higher Education, have an ongoing programme to support institutions in ensuring there are equal opportunities through a range of measures, for example, the introduction of more flexible working arrangements. A consequence of these initiatives should be to reduce the impact of career breaks on individual academics' research programmes. These policies are being introduced in the context of a trend for increasing proportions of women and academics from ethnic minority groups to be in the more senior academic positions, including professorial posts.

12. The differences in rates of selection for submission to the RAE are highly visible. It would be unfortunate if this transparency led to the conclusion that the RAE was the cause of inequalities in research careers. Without the relatively stable public funding through block grant that the RAE underpins, it would, in our view, be more difficult for HEIs to improve the career opportunities for academic staff.

Action required

13. No response to HEFCE is required in relation to this document.

Introduction

14. The Research Assessment Exercise (RAE) assesses the quality of research in higher education institutions (HEIs) in the UK through a process of peer review. The first exercise took place in 1986 and was repeated with progressive modifications in 1989, 1992, 1996 and 2001. A further assessment is planned for 2008. This study is based on an analysis of data from the RAE in 2001, referred to as 'RAE2001'.

15. For RAE2001 the full range of academic disciplines was divided into 68 subject areas or units of assessment (UOAs). HEIs could make one or more submissions within each of these UOAs. The RAE2001 peer review panels, on the basis of the evidence provided, assigned a quality rating to each submission on a seven point scale: 1, 2, 3b, 3a, 4, 5 and 5*, with 5* being the highest rating. For each of their submissions, institutions selected 'research active' staff for inclusion from among their 'eligible staff'. Eligible staff are those academic staff employed by the institution who were not employed to carry out another individual's research programme. Typically eligible staff are on a lecturer grade or are more senior.[1]

16. This selection of certain staff over others by institutions for inclusion in RAE2001 is analysed in this report. While the quality ratings that are subsequently assigned through the RAE process relate to the work of 'departments'[2] and not of individuals, the process of selecting staff within institutions necessarily identifies individual staff as having been judged by the HEI to have a research record suitable for inclusion in the submission, or not. This is inevitably seen by the staff concerned, and by their peers, as having a bearing on their status and career prospects. Therefore it is not surprising that there have been concerns about the possible impact of the RAE on individuals and whether there is equality of opportunity. Is the process of selecting staff fair, or are some staff disadvantaged?

17. A series of measures were adopted to address these concerns. For the 1996 RAE institutions were invited to highlight reasons why certain members of staff included in submissions had reduced research activity. The patterns of selection for the 1996 RAE were then analysed in terms of grade, sex and age as part of the HEFCE review of research (HEFCE, 2000). The measures were strengthened for the 2001 RAE: a separate section on the submission return ('Form RA6') enabled HEIs to give background information about staff circumstances, and panels making assessments were encouraged to take these circumstances into account.

---

[1] Details of the guidance, procedures, submissions and results of RAE2001 are at www.hero.ac.uk/rae.

[2] 'Department' is used throughout this report to mean all the eligible staff associated with a UOA at a particular HEI, whether there were one or more submissions from the HEI within that UOA or not. These staff may not be part of a single administrative unit at the HEI.

18. It was also agreed that RAE2001 should be subject to more careful monitoring with respect to equal opportunities. In support of this objective the Higher Education Statistics Agency (HESA) staff record was enhanced so that for all eligible staff covered by the record, including those not selected, were associated with a unit of assessment. It is these HESA data that underpin most of the analysis reported here.

19. These HESA data were used to calculate rates of 'research activity' for a report published by the Association of University Teachers (AUT, 2004). That analysis did not distinguish between those academic staff who potentially could have been selected for inclusion in a submission, and those (such as research assistants) who were ineligible for selection. While we recognise that there may well be equal opportunities issues surrounding who is recruited or promoted to a position where they are employed as an independent researcher in their own right, it seems to us unhelpful to confound such issues with the concerns about who is selected or not selected for inclusion in the RAE. Because of this, and other differences in method of analysis,[3] the results reported here are not compared with those published by the AUT.

20. We looked at the selection rates for different groups in RAE2001 and, by using statistical models, tried to compare staff on a 'like-for-like' basis. In addition, we used bibliometric data to assess the relative research strength of the men and women, and of staff of different ethnicities, whose work was submitted to RAE2001.

21. The scope of this analysis is therefore limited. We have not, for example, attempted to ascertain whether RAE2001 panels assessed the work of different groups of academics fairly. In our use of bibliometric data, we did not attempt to directly assess whether the process of accepting or rejecting an article is fair, or whether decisions on which articles to cite fairly reflected the research strengths of individual staff. The use of bibliometrics is justified solely in terms of the association between bibliometric measures and RAE2001 quality ratings. We do not need to make assumptions, one way or the other, about these wider issues so long as it is appreciated that this investigation does not address the questions as to whether the research process as a whole is biased or not, but rather whether there are specific disadvantages for certain groups of academics in the process of being selected for inclusion in the RAE.

Determining rates of selection

22. In this section we investigate the factors associated with an individual's probability of being selected for inclusion in RAE2001, with particular reference to differences

[3] As well as including research assistants and other staff who could not be selected, the AUT study used data for the academic year two years after RAE2001. The AUT figures do not refer to individual staff as they are based on counts of HESA records which include duplicates where one individual has several contracts of employment. The figures reported in our report are based on a revised and edited version of the 2000-01 HESA staff record following cross-checks with RAE2001 submission data, whereas the AUT figures are based on the original HESA data.

between individuals with recorded disabilities, between different ethnic groups, and between men and women and different age groups.

The population

23. We are concerned with those staff employed by UK HEIs for academic duties who are not employed to carry out another individual's research programme. We refer to these persons as 'eligible staff' or simply 'staff'. Some RAE2001 submissions included the output of others, referred to as 'Category C' individuals, who carried out research within departments of the HEI but were not employed by the HEI for academic duties. These individuals, who accounted for 6.2 per cent of all persons whose research outputs were assessed, were not included in this investigation. There is no information available concerning individuals who potentially could be counted within Category C but were not included in any submissions.

Data sources

24. The main data source used in this analysis was the HESA individualised staff record for 2000-01. This record excluded staff working with a full-time equivalent (FTE) of less than 25 per cent. We estimate these excluded staff to total about 18.8 per cent of all eligible staff and 1.6 per cent of staff selected as research active.[4]

25. The HESA individualised staff record includes details of the individual as well as the post held. In addition it shows whether the person was eligible for submission to RAE2001, and, if eligible, whether they were selected. All eligible staff, whether selected or not, are attributed to an RAE2001 submission which is identified by a UOA and a multiple submission code.

26. The RAE2001 database was also used in this analysis. This provided the aggregate full-time equivalents of eligible staff, both selected and not selected, for each RAE2001 submission.

Cross-checking and data exclusions

27. Aggregations of the 2000-01 HESA data were cross-checked with the RAE2001 submissions database. For some institutions significant discrepancies were revealed, and subsequent enquiries confirmed that there were inaccuracies in the HESA data. Some institutions were able to correct and resubmit their HESA data; however, for some UOAs within institutions, there remained a significant difference between how many staff the HESA record indicates were entered into RAE2001 and how many the RAE2001 submission database held. Further details of these cross-checks are at Annex B.

[4] This estimate of excluded eligible staff was made by taking using the 2003-04 staff record and taking the proportion of staff active on 31 March 2004 at grade lecturer and above with FTE of less than 25 per cent. The figure for excluded selected staff was taken directly from the RAE2001 database. Both statistics are based on headcounts.

28. To reduce the risk of these data quality problems determining the conclusions of our analysis, we examined three sets of data:

a. All HESA records of eligible staff after making corrections from the cross-checking exercise.

b. Data as in 'a' above with data from highly suspect UOAs within HEIs or complete HEIs removed.

c. Data as in 'a' above with data from moderately suspect UOAs within HEIs or complete HEIs removed.

29. In this report, only the results from data as in 'b' are shown.

30. Staff associated with UOAs in which their HEI did not submit to RAE2001 were excluded from the statistical models and some of the tabulations. We refer to these groups of staff as 'non-submitting departments'. Staff associated with UOAs in which their HEI did submit to the RAE are referred to as 'submitting departments'.

31. Table 1 shows how the initial data extraction and the exclusions described above determine the overall numbers of staff, departments and institutions presented in this report.

Table 1: Numbers of eligible staff, 'Departments' (UOAs within HEIs), and HEIs

	Eligible staff	**Departments (UOAs within HEIs)**	**HEIs**
Eligible staff recorded in the 2001-02 corrected HESA record	82,900	3,697	160
HEIs with large proven errors and 'highly suspect' UOAs within HEIs removed	**74,358**	**3,349**	**143**
Excluding staff associated with non-submitting departments	**62,829**	**2,270**	**141**

Table 1 notes: Each row of Table 1 is defined as a subset of the previous row. Counts are based on numbers of staff at institutions. Duplicate records of staff within an HEI have been excluded. Note that HESA reference volume (HESA, 2002) statistics are based on counts of contract records including duplicate records for the same individual.

Factors that determine whether staff are selected

32. In the results presented below we show the percentage of staff selected for inclusion in RAE2001 for different groups. For example, we find that for men and women the rates were 64 per cent and 46 per cent respectively. While this, in itself, is of interest,

clearly such simple comparisons of selection rates may reflect differences in the patterns of employment between different groups of staff, which do not have a direct connection with the RAE. For example, the selection pattern varies significantly by unit of assessment. In Nursing 71 per cent of the eligible staff were female and the average percentage of staff selected was 13 per cent, whereas in Electrical and Electronic Engineering, where only 7 per cent of the eligible staff were female, the average proportion of staff selected was 63 per cent.

33. To be selected, a member of staff must be associated with a body of research activity that their HEI has decided to submit for assessment. Usually this will involve being part of a department which makes a submission. We can make a distinction between an individual not being selected as a result of being associated with a non-submitting department, and not being selected even though they are associated with a submitting department. In the latter case, it is more likely that the decision will be perceived as being about that person's individual research output. We considered both selection processes by tabulating selection rates for all eligible staff and selection rates for just those staff who are associated with submitting departments.

34. For staff who are associated with a submitting department, being selected will, in part, depend on the quality of their output as determined by the institution. In addition, each institution will have made its own decision as to the threshold level of research quality that individual staff had to achieve to be selected. We cannot assume that this threshold level was the same for different institutions, or even different submissions within the same institution. Clearly, an individual located in a department with a very high threshold level of research quality will be less likely to be selected, all other things being equal. Finally, staff may fail to be selected because of some prejudice or bias against them.

35. We attempted to take account of the differing achievements of different groups of staff, and the different quality thresholds through the construction of statistical models.

36. We simultaneously allowed for the following attributes:

- age; sex; ethnicity; disability
- PhD holder; clinical status; highest qualification in a relevant subject
- location of the individual in the previous year
- grade; contract status; mode of employment.

37. The first four of these variables (age, sex, ethnicity and disability) define the groups we were interested in. The other variables are our best proxies for research quality. Some of these factors, in particular grade and other aspects of employment status, are themselves issues where equal opportunities may be in question; this makes the inclusion of such variables problematic, particularly as being selected for the RAE may improve someone's chances of being promoted to a higher grade. We therefore provide the results of both a 'full statistical model' which includes all these variables, and a 'restricted model' which does not allow for grade, contract status and mode of employment.

38. To allow for varying quality thresholds for different institutions, we constructed the statistical model in such as way that it allows for variation at the HEI level, for the UOAs across all HEIs and UOAs within an HEI, or department, level as well as by individual staff. Where an HEI made two or more submissions within one UOA these have been combined to simplify the model structure. (Given the infrequency of such cases, this should not greatly affect the results.) Details of the modelling are at Annex C.

39. The results of the modelling are presented in terms of 'odds ratios'. A fuller explanation of the 'odds ratio' statistic is given below, along with the presentation of the results for staff with and without disabilities.

Selection rates for staff with and without recorded disabilities

40. As shown in Table 1, in our population there are 74,358 eligible staff, and 62,829 staff who are eligible and associated with submitting departments. Tables 2 and 3 show the corresponding numbers of staff with and without recorded disabilities, and the numbers and percentages that were selected for inclusion in RAE2001.

Table 2: Selection rates for staff with recorded disabilities (including non-submitting departments)

Disability?	Selected	All eligible staff	% Selected	Index
Without disabilities	43,159	73,591	59%	(ref) 1.00
With disabilities	445	767	58%	0.97
Total	**43,604**	**74,358**	**59%**	**n/a**

Table 3: Selection rates for staff with recorded disabilities (excluding non-submitting departments)

Disability?	Selected	All eligible staff	% Selected	Index
Without disabilities	43,159	62,185	69%	(ref) 1.00
With disabilities	445	644	69%	0.99
Total	**43,604**	**62,829**	**69%**	**n/a**

41. We can see that the percentage of staff with recorded disabilities selected was very similar to that for staff without recorded disabilities, whether non-submitting departments are included or not.

42. Using statistical models we can explore the extent to which the selection rates can be compared on a 'like-for-like' basis after allowing for other factors. The results of this modelling can be most conveniently presented as a 'selection index'. Table 4 shows how this index is calculated for the actual 'raw' figures shown in Table 2.

Table 4: Derivation of the selection index

	Without disabilities (ref)	With disabilities
Percentage selected	59%	58%
Percentage not selected	41%	42%
Selected/Not selected (odds ratio)	1.42	1.38
Odds ratio relative to odds ratio of reference group (selection index)	1.42/1.42=1.00	1.38/1.42= 0.97

43. If the selection rate for staff without and with disabilities had been the same, the selection index would have been exactly equal to 1.00. The value 0.97 indicates that staff with disabilities had a slightly lower selection rate than the reference group, staff without disabilities. The equivalent selection index when excluding staff at UOAs within HEIs that were not submitted (Table 3) is 0.99. These 'actual' selection indices are 'unadjusted' in that they do not allow for other factors. Table 5 shows this actual index from table 3 along with the indices from the statistical models which take other factors into account.

Table 5: Selection indices comparing staff with and without recorded disabilities (excluding non-submitting departments)

Disability?	Actual	Restricted model	Full model
Without disabilities	(ref) 1.00	(ref) 1.00	(ref) 1.00
With disabilities	0.99	0.99	0.97

Table 5 notes:
No significant differences from1.00 at 5% level

44. The restricted model does not take account of factors relating to employment status, while the full model does (see paragraph 36). We see that the selection indices for both restricted and full models were not significantly different from 1.00. There is therefore no evidence to suggest an advantage or disadvantage for those staff with a disability.

45. There could be differences in the selection rates for staff with and without disabilities for specific groups, defined by age or other characteristics. None were found in our study. However, given the relatively small number of staff with a recorded disability, such effects would have to be large to be detectable.

Selection rates for men and women

46. Tables 6 and 7 show the selection rates for men and women. We can see marked differences. Men had a much higher selection rate. Twenty per cent of women in the population of eligible staff were associated with non-submitting departments, compared to 14 per cent of men, but a large difference in selection rates is still found when staff associated with non-submitting departments are excluded (see Table 7).

Table 6: Selection rates for men and women (including non-submitting departments)

Sex	Selected	All	% Selected	Index
Women	10,759	23,243	46%	(ref) 1.00
Men	32,845	51,115	64%	2.09
Total	**43,604**	**74,358**	**59%**	**n/a**

Table 7: Selection rates for men and women (excluding non-submitting departments)

Sex	Selected	All	% Selected	Index
Women	10,759	18,644	58%	(ref) 1.00
Men	32,845	44,185	74%	2.12
Total	**43,604**	**62,829**	**69%**	**n/a**

47. These selection indices are 'unadjusted' and do not allow for other factors. Before presenting the output from the statistical models, we first consider the joint effect of sex and age on selection rates.

Selection rates for men and women by age

48. Figures 1 and 2 show that the rates of selection varied by age for both men and women. The broad pattern was the same for both sexes. Selection rates increased sharply up to about 30, declined gradually from mid-30s to the mid-50s, followed by a small rise in later years.

Figure 1: Selection rates for men and women by age (including non-submitting departments)

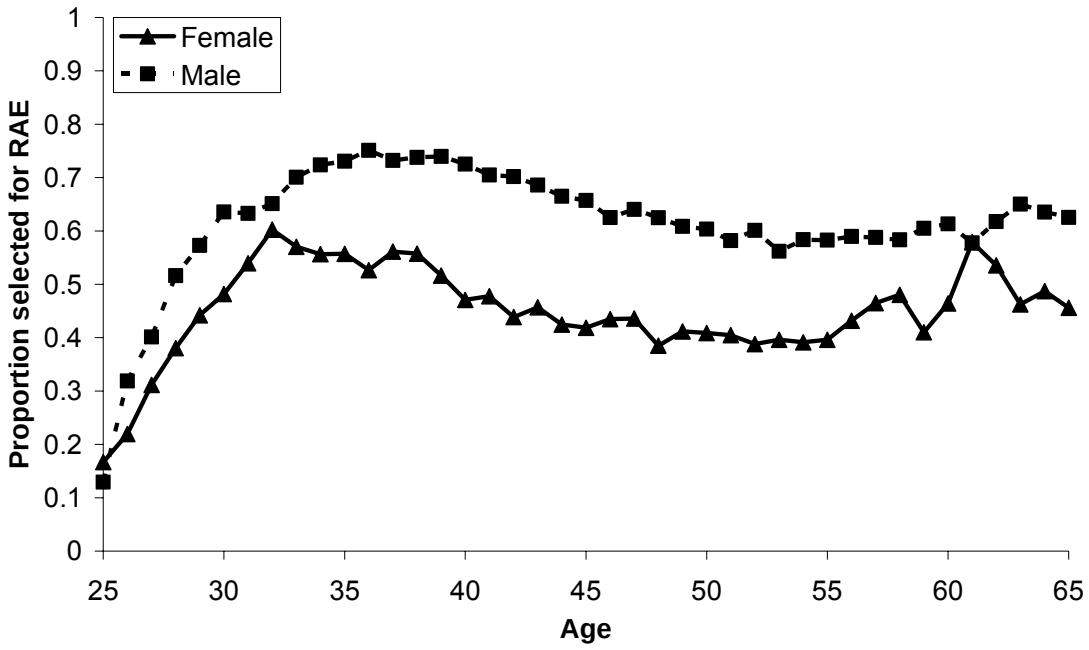
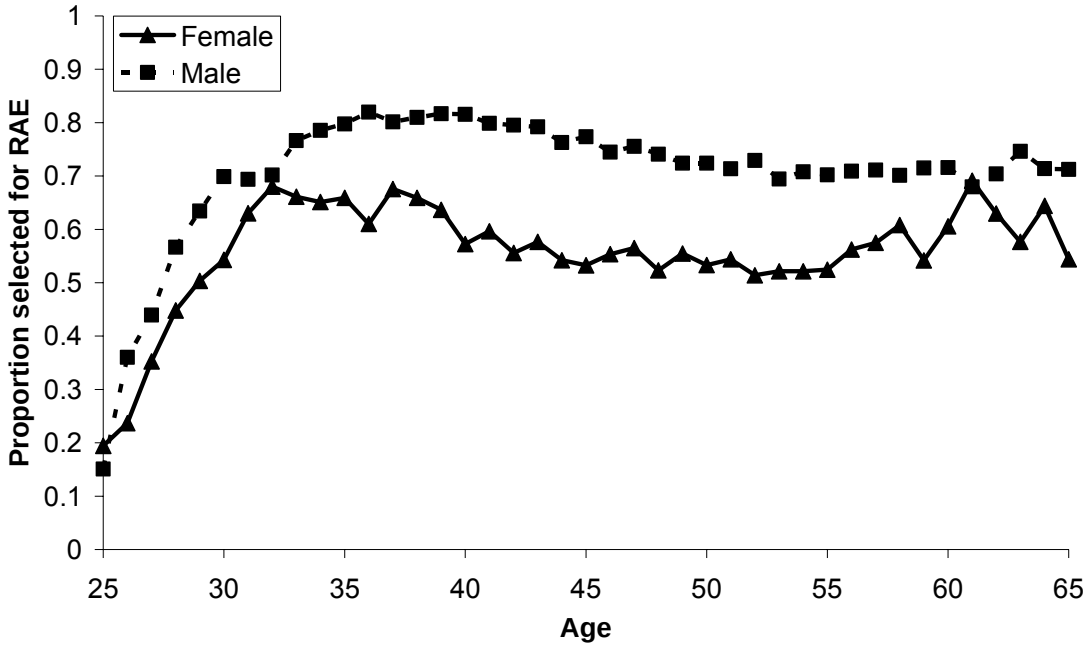
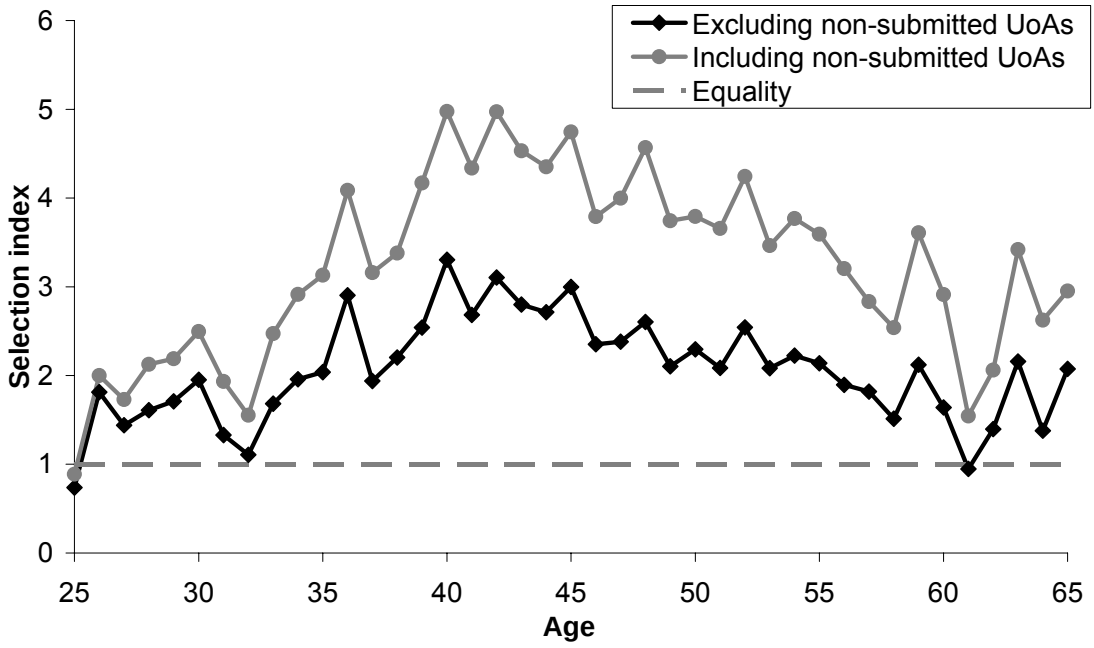


Figure 2: Selection rates for men and women by age (excluding non-submitting departments)



49. However, as Figures 1 and 2 show, the relationship between age and selection rate was not exactly the same for men and women. Therefore, the relative rates of men compared to women must have also varied by age, with the biggest differences in the middle years, between about 35 and 55. This is illustrated in Figure 3, which shows how the selection index varied by age.

Figure 3: Actual selection indices by age



50. When we used statistical models to take other factors into account, we found that the adjusted selection index also depended on age. It is therefore not meaningful to ascribe a single value to compare men with women. Figures 4 and 5 show the selection indices by age after adjusting for other factors with the restricted and full models respectively.

Figure 4: Restricted model selection index by age (excludes employment status variables)

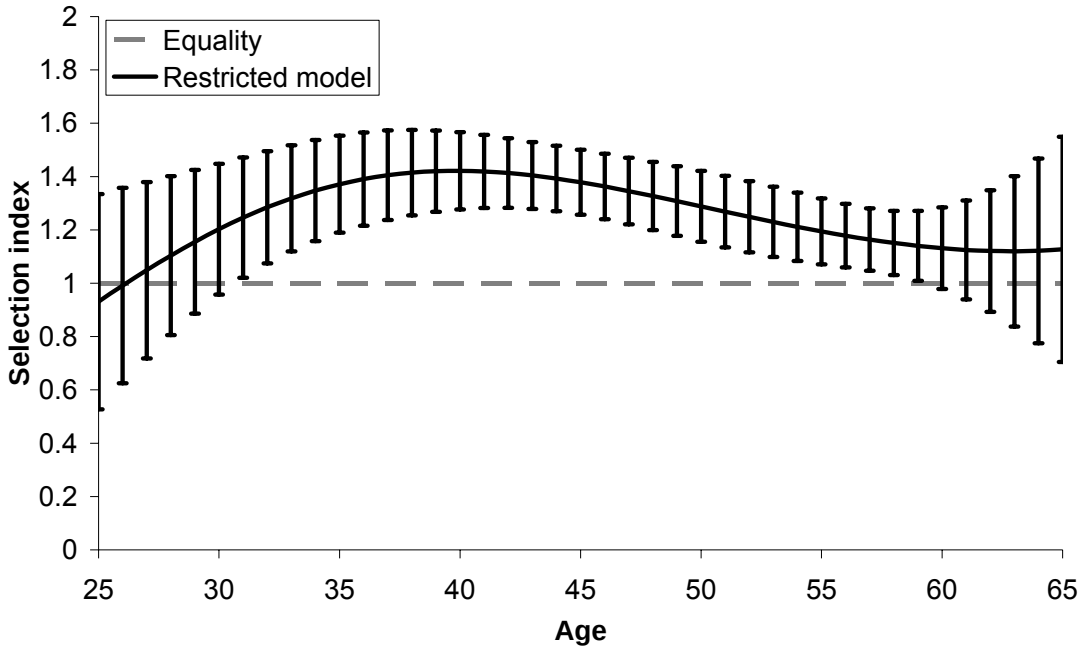
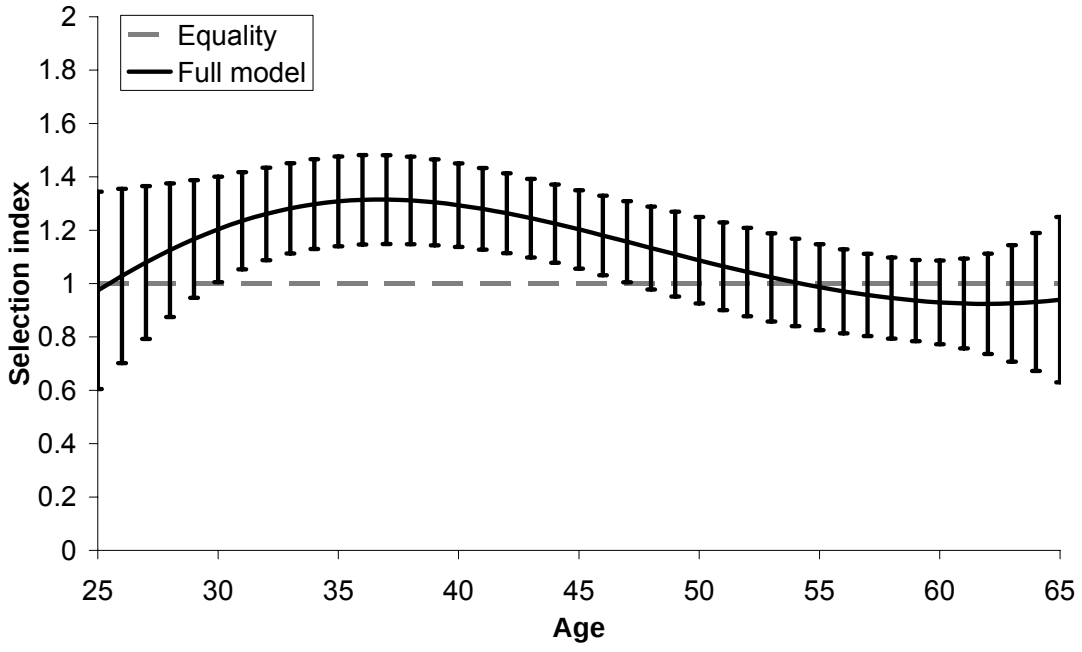


Figure 5: Full model selection index by age (includes employment status variables)



51. Figures 4 and 5 show non-linear relationships between age, and the selection index. The error bars show the 95 per cent confidence interval of the index for each age. Where the interval is completely above 1, it shows that there is a statistically significant unexplained advantage towards men.

52. We can see that there was a statistically significant advantage for men in the middle of the age range for both restricted and full models. Using the restricted model, women are shown to have been disadvantaged between 31 and 59. With the full model the age range was somewhat narrower, from 30 to 47.

53. There is no conclusive way of deciding whether the restricted or the full model is the most appropriate. If a bias existed which disadvantaged women with respect to grade, and also in the decision as to whether they should be selected for RAE2001, the restricted model would be more appropriate. Conversely, if grade is a good proxy for research strength, the full model may be more appropriate. However, in this case, even the full model shows that, at least for staff in mid-career, men had higher selection rates than we would expect after taking into account all the factors that are available to us.

Variability of selection index by HEI and UOA

54. The statistical modelling also showed that there was significant variation in the selection index between men and women that was due not only to the age of individual being considered for selection but also to the higher education institution and the UOA within that institution. There was no strong evidence that the UOA as such, across the sector, was associated with the relative selection rates of men and women.

Interpreting the 'like-for-like' differences

55. We have presented the comparisons of the selection rates for men and women in terms of 'selection indices' which have a simple relationship to the coefficients from the statistical modelling. Table 8 shows what these results mean in terms of expected selection rates for staff with some example sets of attributes.

Table 8: 'Full model' selection rates for 'typical' staff

	Male	Female
32 year-old lecturer in Law without a PhD	53%	47%
40 year-old senior lecturer in Physics with a PhD	83%	79%
55 year-old professor in History with a PhD	88%	88%

Table 8 notes
The other attributes of these staff are: Full-time; Permanent contract; Teaching and research function; Subject area of highest qualification related to the subject area of the UOA; Non-senior post holder; Non-clinical status; Employed at the same institution in the previous year; and within a 4 rated UOA.

Selection rates by ethnicity

56. Fewer than 3,000 staff in the population analysed were from ethnic minorities so, in order to ensure sufficient numbers were available, a simplified classification into four ethnic groupings was used. These were: White, Black, Asian and Other, as well as 'information refused'. (See Annex B for definitions of these groupings.) Tables 9 and 10 show the rates of selection for each ethnic grouping.

Table 9: Selection rates by ethnicity (including non-submitting departments)

	Selected	All	% Selected	Index
White	36,328	62,310	58%	(ref) 1.00
Black	267	723	37%	0.42
Asian	1,547	2,581	60%	1.07
Other	801	1,362	59%	1.02
Refused	4,661	7,382	63%	1.23
Total	43,604	74,358	59%	n/a

Table 10: Selection rates by ethnicity (excluding non-submitting departments)

	Selected	All	% Selected	Index
White	36,328	52,319	69%	(ref) 1.00
Black	267	537	50%	0.44
Asian	1,547	2,241	69%	0.98
Other	801	1,170	68%	0.96
Refused	4,661	6,562	71%	1.08
Total	43,604	62,829	69%	n/a

57. The selection rates for each of these groupings were similar, except for staff from ethnic groups under the heading 'Black' who had an overall selection rate of 37 per cent. Part of the reason for this relatively low figure for Black staff is that a higher proportion, 26 per cent, were in non-submitting departments. Table 11 shows the percentage of eligible staff in non-submitting departments for each of the ethnic groupings.

Table 11: Percentage of staff in non-submitting departments by ethnicity

	Non-submitting departments	All departments	% Non-submitting departments
White	9,991	62,310	16%
Black	186	723	26%
Asian	340	2,581	13%
Other	192	1,362	14%
Refused	820	7,382	11%
Total	11,529	74,358	16%

58. However, even when non-submitting departments are excluded, as in Table 7, we find that staff included in the Black grouping had a lower selection rate.

59. Table 12 shows the selection indices for each ethnic grouping after taking into account other factors from both the restricted and full models. The results from the restricted model shows that the selection rate for the three ethnic groupings 'Black', 'Asian' and 'Other' were all lower than expected. For the Black grouping this is not surprising, given the low actual selection rates, but the actual selection rates for Asian and Other groups were very similar to that for staff from the White group. It must be that the combination of individual characteristics of staff in the Asian and Other groups – like having a PhD, and the departments that they are working in – means that we would expect them to have higher than average selection rates, which were not found. When the full model is used, we find no significant differences in the selection rates between the ethnic groupings. This is largely due to the addition of 'grade' as a variable in the model. Staff from ethnic minorities tend to be on lower grades, and when grade is taken into account their expected selection rates are lower.

60. There are a number of possible interpretations of these findings. It could be that ethnic minority staff were unfairly disadvantaged with respect to both grade and selection to RAE2001. On the other hand, if we assume that grade is a reliable proxy for research strength, the findings could be interpreted as showing that, on average, ethnic minority staff had lower achievements in research, compared to what was expected from the restricted model which did not include grade. We do not have the evidence to decide which of these, or other interpretations, hold true. However, the bibliometric analysis, described below, is consistent with the second interpretation: that the research strength cannot be sufficiently allowed for with the restricted model, and that the full model is needed.

Table 12: Selection indices comparing staff from different ethnic groups (excluding non-submitting departments)

	Actual	Restricted model	Full model
White	(ref) 1.00	(ref) 1.00	(ref) 1.00
Black	0.44	** 0.78	0.87
Asian	0.98	** 0.90	1.01
Other	0.96	* 0.88	0.93
Refused	1.08	** 0.90	0.96

Table 12 notes

* indicates significantly different from 1.00 at the 5% level

** indicates significantly different from 1.00 at the 1% level

61. It is possible that, in addition to these differences in selection rates between staff from different ethnic groupings, there could have been age-specific differences, or differences specific to some other attribute. No evidence was found for such differences, but, given the relatively small numbers of staff from ethnic minorities, such age-specific ethnically relative selectivity would need to be quite marked to be detected.

Bibliometric measures of research outputs of groups of selected staff

62. When using the restricted model we find that ethnic minority staff had lower selection rates and, even after taking into account the factors in the full model, there was a significant difference in the selection rates for men compared to women for staff in the age range 30 to 47. The question arises, do these differences simply reflect a difference in the quality of research, or are they evidence of unjustifiable biases? Unfortunately as already noted, the variables available to us, with the possible exception of 'holding a PhD', are somewhat distant from what we really need – an accurate and objective measure of the 'research strength' of each eligible member of staff, both selected and not selected.

63. Such measures are not available to us, but, at least for some subjects, we can approximate to such a measure for selected staff by using bibliometrics. Essentially, such measures assess research by counting how frequently articles are subsequently cited.

64. In using bibliometrics to further explore the reasons behind differences in selection rates we recognised the following limitations:

a. Bibliometrics give only an imprecise measure of research strength. They certainly could not be used to assess an individual researcher or even a particular submission. The best that can be achieved is to get an assessment of the average research strengths of large numbers of researchers.

b. Bibliometrics can only be used for those UOAs where the usual form of research output is an article in a journal which is included in citation databases.

c. Bibliometrics summarise judgements as to whether an article should be accepted for publication, and by authors of articles published subsequently as to whether to refer to it. We do not know whether these judgements are biased.[5]

d. We only have details of the publications of staff that were selected. We have no bibliometric data for those who were not selected.

65. Despite these limitations, bibliometrics provided the only practical way of further exploring the reasons for differences in selection rates with the available data. If there was large-scale discrimination against women or ethnic minorities, so that they had to achieve a higher quality of research output to be selected, we would expect to find that

[5] A number of studies provide evidence that men are more likely to cite men, women more likely to cite women. See, for example Ferber (1988), Davenport (1995) and Hakanson (2005). This would be expected if men and women tended to focus on different areas of research, but it could also arise through systematic differences in deciding what was worth citing.

the publications of women, or researchers from ethnic minorities, had a higher average impact than the publications of white men.

Bibliometrics – data and methodology

Data sources

66. The principal source of data was the details of published outputs submitted to RAE2001, which is available via the HERO web-site (www.hero.ac.uk). These include 99.3 per cent of the outputs that were assessed, the remainder remaining unpublished due to the possible risks to the personal safety of the researchers. Thirteen units of assessment were initially selected for our analysis, though for two (UOAs 42, Sociology and 59, History) it was subsequently found that the relationships between bibliometric measures and RAE2001 quality were too weak for the bibliometrics to be used.

67. Additional information was obtained from the Thomson Scientific National Citation Report database and the HESA 2000-01 Individualised Staff Record.

Linking to citation databases

68. The selected RAE2001 data were linked to the Thomson Scientific National Citation Report database by Evidence Ltd. A description of this matching and the bibliometrics that were appended to the RAE2001 data are at Annex D.

Linking to the HESA

69. The RAE2001 output data only include limited information about the researchers: their institution and the submission they were associated with, their age and sex. To append additional information the data were linked to the HESA 2000-01 staff record. Full details are at Annex E. This linking was only partial, and is estimated to include 17 per cent mismatches, but the additional information did provide an opportunity to investigate the relationship between bibliometrics and research quality for different groups of individuals within RAE2001 submissions.

Data analysed

70. Table 13 shows the numbers of individuals and outputs that were extracted for our analysis. Researchers who were not employed by the HEIs for academic duties – 'Category C' individuals – were excluded from part of the analysis.

71. Thirteen UOAs were initially selected. Of these, nine were subjects where bibliometrics were known from previous work to provide a good measure of research quality and together covered a broad range. UOAs from the social sciences and the humanities were added to see if it was possible to extend the coverage, and Food Science and Technology was included because, though the number of researchers is not large, there was a relatively high proportion of women. Of these 13 UOAs, 11 were

included in the bibliometric analysis report here. (Annex D gives a full explanation of the UOA selection.)

Table 13: Staff and outputs extracted and matched for bibliometric analysis

Data set		Person count	Output count
	Complete RAE2001 database	53,845	205,285
	Published RAE2001 database	53,445	203,748
	11 selected UOAs (excluding Sociology and History)	12,261	47,347
	Matched to NCR 2003 database	10,130	32,841
1	**Matched and UOA/institution present in original data**	**9,375**	**30,439**
2	**Match to HESA staff records**	**7,118**	**22,850**

Table 13 notes
Each row was taken as a subset of previous row
Dataset 1 includes only those data described at paragraph 28b

Bibliometrics and other measures of research quality

72. There is a consensus, confirmed through a review of RAE2001 (HEFCE, 2003) and the subsequent consultation exercise (Funding bodies, 2004), that the most authoritative measures of research quality are through peer review; and therefore the RAE quality ratings are widely accepted as the most authoritative measure of research quality. Even if this were not accepted, given that a rational criteria for deciding which staff should be selected for the RAE would be to aim to achieve a high RAE quality score, we can only justify using bibliometrics if they are reasonably consistent with research quality as indicated by the RAE.

Relationship between bibliometric measures and RAE quality scores

73. We assessed a large number of possible bibliometric measures for different UOAs to see which, if any, corresponded to RAE2001 quality ratings. The data were grouped by UOA and quality rating, which resulted in 73 combinations. We excluded combinations with less than 50 individuals, which left 40 combinations. We then examined the rankings by average bibliometric measure against RAE2001 quality rating.

74. We failed to find any satisfactory measures for History and Sociology. For the remaining subjects for which bibliometric data had been prepared we found that the following measures had the most consistent relationship with RAE2001 quality ratings:

a. Average citation count relative to a field baseline.

b. Best citation count relative to a field baseline.

c. Proportion of papers over a field baseline.

75. Of these three measures, the average citation count compared to the citation count for the field is the most straightforward and seems to be the most robust. The tables and charts presented in the main body of this report use this measure, though values for all three measures are tabulated at Annex G. Figures 6a, 6b, 6c and 6d show the relationship between these measures and RAE2001 quality ratings for nine subjects where there are sufficient numbers of assessed staff across a range of RAE ratings.

Figure 6a: Citation rates relative to field and RAE2001 quality ratings (Clinical Laboratory Sciences, Community-based Clinical Subjects)

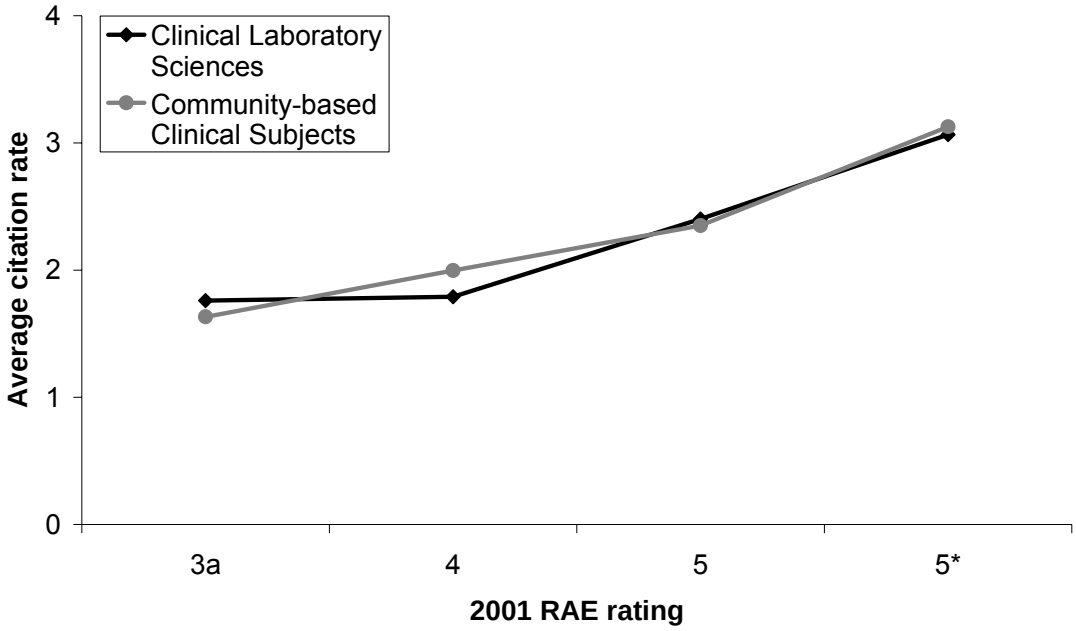


Figure 6b: Citation rates relative to field and RAE2001 quality ratings (Biological Sciences, Chemistry, Environmental Sciences)

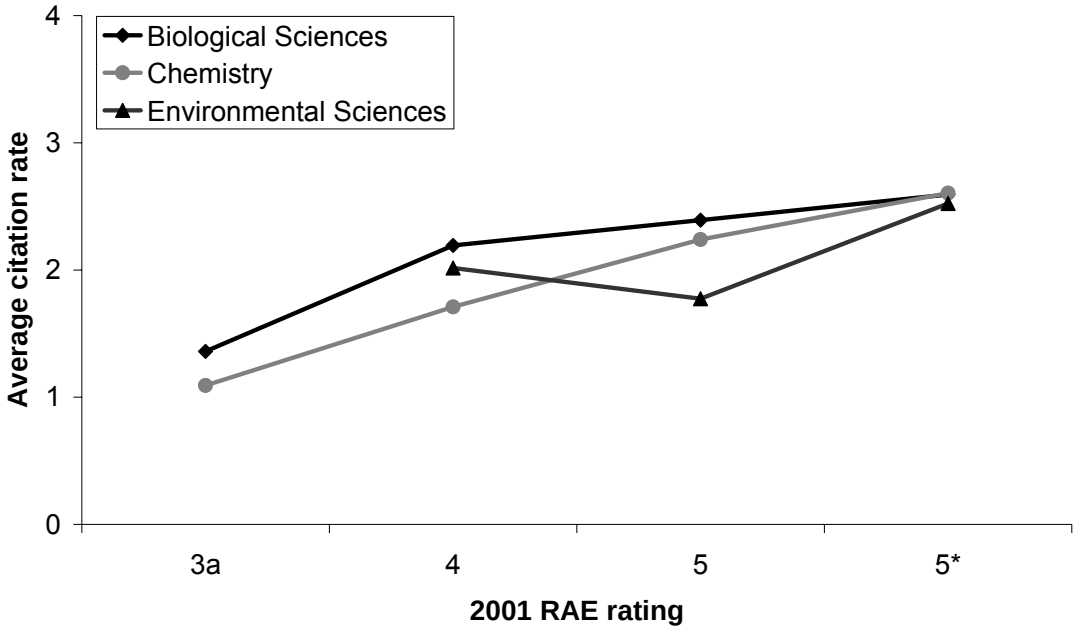


Figure 6c: Citation rates relative to field and RAE2001 quality ratings (Civil Engineering; Mechanical, Aeronautical and Manufacturing Engineering)

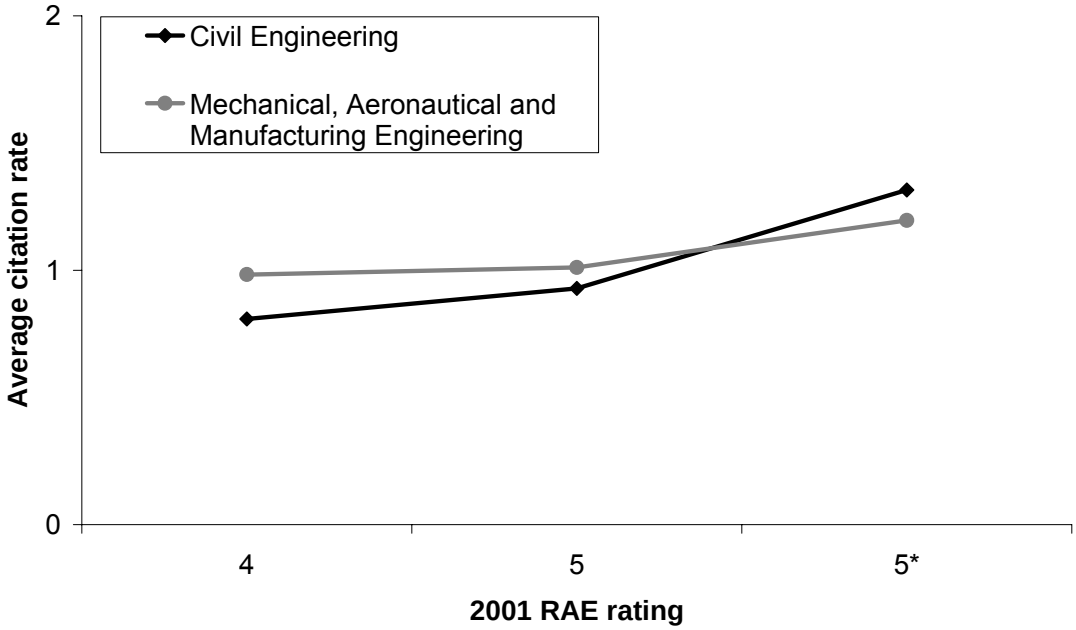
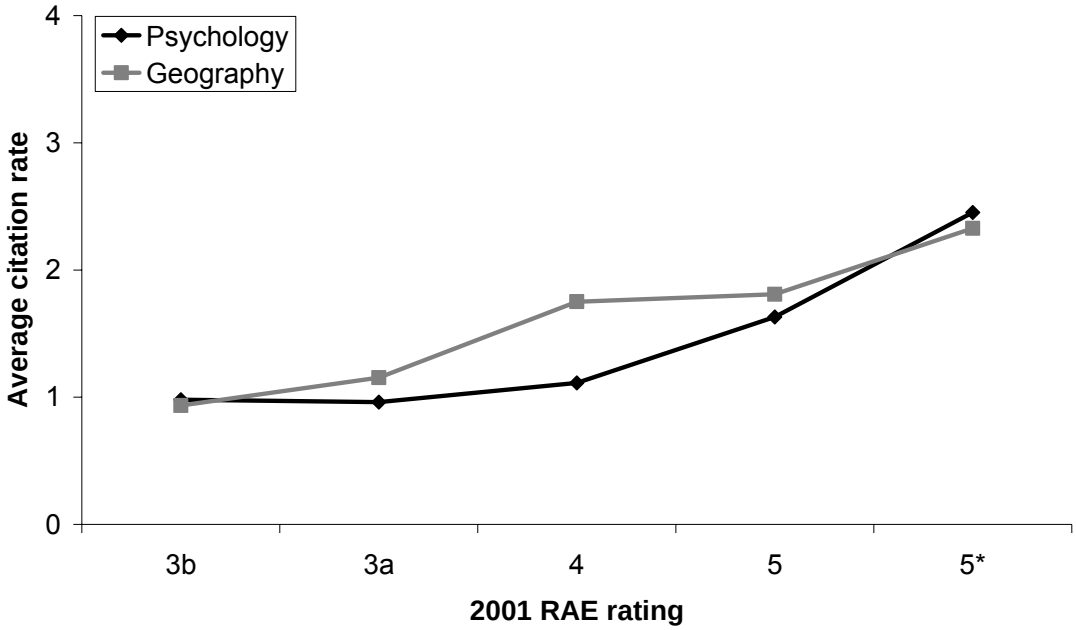


Figure 6d: Citation rates relative to field and RAE2001 quality ratings (Psychology, Geography)



Notes for Figures 6a-6d

Based on dataset 1 as listed in Table 13

Excludes UOA of given RAE rating with less than 50 individuals

76. The association between bibliometric and RAE rating at this aggregate level is clear. However, this disguises quite a wide variation when the relationship is examined at the level of individual submissions. This is illustrated in Figure 7.

Figure 7: Citation rates relative to field and RAE2001 quality ratings for individual Chemistry submissions

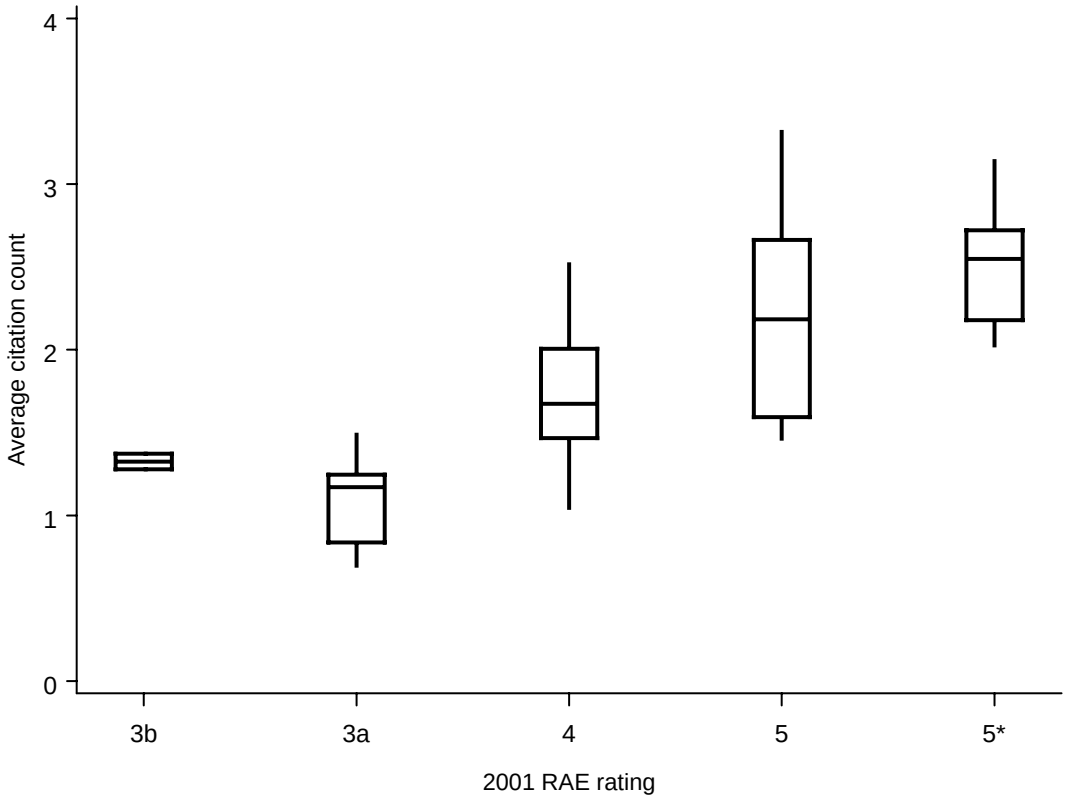


Figure 7 notes

Based on dataset 1 as listed in Table 13

Each box shows the lower quartile (25 per cent), median (50 per cent) and upper quartile (75 per cent). The end of the lower tail shows the minimum value, and the maximum value is shown at the end of the upper tail.

77. We can see from Figure 7 that there is a wide overlap in the bibliometrics measures across different RAE2001 ratings, which means using that data at the level of individual submissions cannot provide reliable results.[6]

Selection of staff and outputs for submission to RAE2001 and bibliometrics

78. Unfortunately, we only have information on the outputs of those staff who were selected for assessment. There is, however, some comparative information about the articles that were submitted to RAE2001 and articles that were not submitted from the

[6] The weakness of the association between bibliometric measures and RAE ratings at the level of individual submissions is in part explained by the distribution of citations. The distributions tend to be highly skewed, with a long tail of papers with very high citation values (Adams, 2006). This means that, within a single submission, the average citation value may be largely determined by a relatively small number of individual papers. However, taking the average of log(1 + citation measure) to reduce the impact of this tail produced only a very small improvement in the alignment between bibliometric measures a RAE ratings.

same HEI, in the same field, over the same period. Jonathan Adams of Evidence Ltd reported that the citation rates for submitted articles compared with the citation rates for all articles were 1.9 and 1.6, for Chemistry and Psychology respectively (Adams, 2004).

79. These non-submitted articles include articles by selected staff not included in what they consider to be their best four outputs, and articles by staff who were not selected. Thus the non-selected articles were judged to be of lower quality by either the individual researchers or the department. Given even a limited shared understanding of what constitutes high quality research, this means that the submitted outputs should have been better than those not submitted, in terms of the criteria and judgements of RAE 2001 panels. The fact that a difference was found in the citation rates provides further evidence of an association between bibliometrics and research quality as determined through the RAE2001 process.

Comparing groups of individuals within submissions

80. In comparing different groups of staff, we need to make comparisons within submissions because, as noted previously, we cannot assume that the standard required to be selected would be the same for each HEI, or for each UOA within an HEI. On the other hand, as we have seen, bibliometrics are not strongly associated with RAE2001 ratings at the level of individual submissions. To work within these constraints we have, in effect, summed the within-submission differences, weighting these individual submission differences by the strength of the evidence. This was achieved through simple multi-level models, details of which are provided at Annex F. Simple overall average differences are also shown for comparison.

Comparing professors with staff on lower grades

81. In analysing selection rates, the ‘full’ model included grade as one of the factors but the ‘restrictive’ model did not. Including grade is problematic, because grade, like selection to the RAE, is an equal opportunities issue. On the other hand, it seems likely that grade may act as a proxy for research strength. Table 14 shows the overall differences between professors and non-professors alongside the within-submission differences for each UOA.

Table 14: Citation rates relative to field for professors and other staff

Subject	Professor	Non-professor	Overall difference	Within-submission difference
Clinical Laboratory Sciences	3.32	2.36	0.96	** 1.06
Community Based Clinical Subjects	2.71	2.13	0.58	** 0.71
Psychology	1.95	1.49	0.46	** 0.43
Biological Sciences	2.60	2.21	0.39	** 0.47
Food Science and Technology	2.36	1.06	1.30	* 1.30
Chemistry	2.48	1.98	0.50	** 0.50
Environmental Sciences	1.74	1.42	0.32	0.26
Pure Mathematics	1.43	1.01	0.42	** 0.34
Civil Engineering	1.05	0.97	0.08	0.1
Mech, Aero & Man Engineering	1.02	1.03	-0.01	-0.02
Geography	2.19	1.57	0.62	** 0.61
All 11	**2.23**	**1.77**	**n/a**	**** 0.50**

Table 14 notes:
Based on dataset 2 as listed in Table 13
* indicates significantly different from 1.00 at 5% level
** indicates significantly different from 1.00 at the 1% level

82. The simple overall differences are very similar to the sum of within-submission differences derived from the model. We see that, apart from one UOA, the citation rate for professors is higher than for staff on lower grades, For seven UOAs and overall, the difference is significant at the 1 per cent level. The only subject where the citation rate is higher for non-professors is Mechanical, Aeronautical and Manufacturing Engineering, and the difference is small and not significant. This provides some independent evidence that, in general, grade is a proxy for research strength, and provides some justification for taking the results from the full rather than restricted selection model.

Comparing men and women

83. Table 15 shows the differences in citation rates relative to the field between men and women both overall and within submissions.

84. Again we find that the overall and within-submission differences are similar. We see that overall men had significantly higher citation rates. Higher citation rates are also found for men than women in nine of the 11 subjects, though these differences are only significant for three subjects.

Table 15; Citation rates relative to field between men and women

Subject	Male	Female	Overall difference	Within-submission difference
Clinical Laboratory Sciences	2.71	2.01	0.71	** 0.71
Community Based Clinical Subjects	2.42	2.01	0.41	* 0.38
Psychology	1.67	1.50	0.17	0.12
Biological Sciences	2.28	2.19	0.09	0.18
Food Science and Technology	1.52	1.10	0.42	0.44
Chemistry	2.02	2.29	-0.27	-0.28
Environmental Sciences	1.64	1.15	0.49	* 0.42
Pure Mathematics	1.16	0.86	0.3	0.32
Civil Engineering	0.99	0.93	0.06	0.01
Mech, Aero & Man Engineering	1.06	0.88	0.18	0.15
Geography	1.70	1.72	-0.02	-0.02
All 11 subjects	**1.88**	**1.84**	**n/a**	**** 0.23**

Table 15 notes:

Based on dataset 1 as listed in Table 13

* indicates significantly different from 1.00 at 5% level

** indicates significantly different from 1.00 at the 1% level

85. However, the overall average is not the most appropriate statistic. We want to know whether women had a higher research quality threshold compared to men. If this were the case we would expect those women who were just, and only just, selected to have a higher quality of research output than the men who were 'only just' selected. The staff with the highest research output quality, who were bound to be selected, are not relevant. Estimating the 'threshold' research quality is problematic. Either we have to make an assumption about the form of the distribution of research quality, or we have to rely on a small number of observations to make an estimate. We have no basis for assuming the form of this distribution and, as we have seen, bibliometrics only work when estimating the average of a relatively large number of articles.

86. Table 16 shows the results of taking a simple approach to this problem. The selected staff, both men and women, within each submission were ranked by their research strength as measured by the citation rates to their articles. Within each submission, the staff in the lower half of the range (in terms of citation rates) were used to calculate the differences between men and women.

Table 16: Citation rates relative to field between men and women (Staff in the lower 50% of citation rates within submissions)

Subject	Male	Female	Overall difference	Within-submission difference
Clinical Laboratory Sciences	1.09	1.06	0.02	0.02
Community Based Clinical Subjects	0.97	0.94	0.03	0.00
Psychology	0.73	0.67	0.06	0.03
Biological Sciences	0.95	1.01	-0.05	-0.02
Food Science and Technology	0.68	0.65	0.02	0.03
Chemistry	0.89	0.95	-0.06	-0.08
Environmental Sciences	0.72	0.57	0.15	0.07
Pure Mathematics	0.52	0.56	-0.05	0.01
Civil Engineering	0.39	0.31	0.08	-0.02
Mech, Aero & Man Engineering	0.44	0.42	0.02	0.02
Geography	0.71	0.65	0.06	0.02
All 11 subjects	**0.79**	**0.84**	**n/a**	**0.00**

Table 16 notes:
Based on dataset 1 as listed in Table 13
No within-submissions differences are significantly different from zero

87. When the comparison is restricted in this way, the relative research strength of men and women seems to be similar. This suggests that the overall higher values for men shown in Table 15 were indeed the result of a higher proportion of men among staff with very high citation rates.

88. The 50 per cent selection is rather arbitrary. Unfortunately, if we take smaller proportions, in an attempt to get closer to the minimum citation rates for selection, the numbers become too small to make meaningful estimates, at least when treating each UOA separately. Figure 8 shows how this overall difference in citation rates changes within different percentages of staff included in the estimation.

Figure 8: Overall estimates of within-submission differences in citation rates relative to field between men and women by percentage of staff excluded

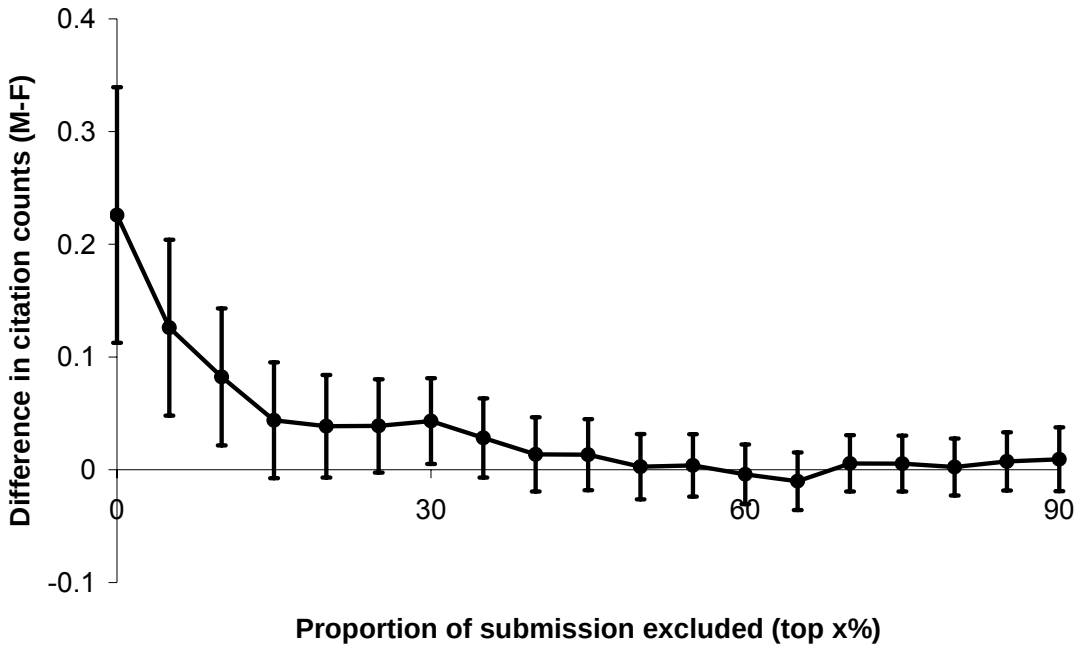


Figure 8 notes
Based on dataset 1 as listed in Table 13

89. Figure 8 suggests that, at least overall, the citation rates of men and women are roughly equivalent when we select staff at the lower end of the citation rate distributions within submissions.

Comparing staff from different ethnic groups

90. Because ethnicity is not captured on the RAE2001 submission, we can only use those records that we were able to link to the HESA staff data. As a result, meaningful statistics on individual UOAs could not be derived, and the analysis was limited to the overall differences in citation rates. Also, the data were too sparse to derive results for different ethnic groupings.

91. Table 17 shows those overall citation rates for whites and for all staff associated with the Black, Asian or Other groupings. It at first appears that the non-whites had lower citation rates but, as the within-submission difference calculations show, this is due to the distribution of ethnic minority staff across the submissions. The within-submission difference is not significant, and this result holds when we only included varying proportion of staff, ranked by bibliometric research strength. Table 17 shows the figures for the lower 50 per cent of citation rates for illustration.

Table 17: Citation rates relative to field between different ethnic groups (Summary of all 11 UOAs)

Population	White	Non-white	Overall difference	Within-submission difference
All staff	1.84	1.64	n/a	-0.01
Staff in the lower 50% of citation rates within submissions	0.80	0.66	n/a	0.01

Table 17 notes:

Based on dataset 2 as listed in Table 13

No significant differences

Are there biases in selection? Some provisional conclusions

92. The limitations of this study need to be appreciated. A difference in selection rates between one group of staff and another does not necessarily mean that one group has been treated unfairly. While we can try to compare different groups on a 'like-for-like' basis, we only have rough proxies for the relative quality of the research output of different groups of staff. Conversely, even where our analysis shows no overall difference in the selection rates for staff with different attributes, this does not mean that there have been no individual cases of bias. Any allegations of unfair treatment would have to be investigated on the evidence of the particular case. All that we can assess through this analysis is the evidence for a general bias against groups of individuals. Bearing these points in mind, we found the following with respect of disability, sex and age, and ethnicity.

Disability

93. The selection rate for staff with recorded disabilities was 58 per cent, slightly lower than for staff not so identified (59 per cent). This small difference is in part accounted for by the fact that staff with disabilities are more likely to work in departments which have not made an RAE2001 submission. When other attributes are taken into account, disability is not a significant factor in propensity to be selected.

Age and sex

94. Both simple and 'like-for-like' comparisons show that the proportion of staff selected varied significantly by age. The general pattern is that staff aged over 30 were more likely to be selected. This result cannot be explained by the age profile of 'research assistants', since most of these staff were not eligible and were not included. This result suggests that either institutions were unwilling to select staff who had not had long enough to build up a substantial research record, despite the provisions and guidance designed to ensure that they would be selected, or, in general, researchers in the early parts of their careers may not produce outputs of the highest possible research quality.

95. Overall, there was a large difference between the selection of men and women, with selection rates of 64 per cent and 46 per cent respectively. Within these overall figures there were differences by age, with the biggest differences between men and women between the ages of 40 and 55. When other factors are taken into account, the

like-for-like comparisons still show that men had significantly higher selection rates than women over the age range 31 to 59, or 30 to 47 when grade and other employment variables are included.

96. Bibliometric measures of research strength show no great difference between men and women among those selected for submission if the small group of very highly-cited staff are excluded. The bibliometric method we used does not take into account the number of articles associated with each member of staff, and thus gives due credit to submitted staff whose cited outputs are fewer but of high quality in keeping with the approach that the panels should have adopted. Though not entirely conclusive, the results produced by this method are consistent with an explanation that the lower selection rate of women, particularly in the mid-career age range, was due to a lower proportion of women having a research record that leads them to be selected, rather than a bias in the selection process.

97. This could be due to more women taking career breaks. If this is the case, and if such breaks do lead staff whose research is generally of a high quality to have less than four suitable outputs, then they should still be selected, with appropriate explanations as provided for in the guidance. We cannot be entirely sure whether the differences in selection rates were due to a reluctance on the part of those making the submissions to make use of these provisions, or whether such career breaks, or other factors, impacted on the quality of research output leading to a lower proportion of women than men having outputs of suitable quality and number to be submitted within the terms of the guidance. The evidence from the bibliometric analysis is consistent with the latter explanation.

Ethnicity

98. The simple unadjusted comparisons show similar selection rates of around 58 per cent to 60 per cent for staff with known ethnicity, apart from staff included in the Black grouping which had a much lower rate of 37 per cent. This lower rate was partly the result of a higher proportion of these staff being employed in departments which did not make an RAE2001 submission, but even when non-submitting departments are excluded, the selection rate for staff from Black ethnic groups is much lower than for others.

99. The results of modelling selection are not clear cut. When a model using all the available factors, including grade, (the ‘full model’) is used, we find no significant differences in the selection rates for different ethnic groupings. However, if ethnic minorities were unfairly treated with respect to selection for the RAE, they may also be unfairly treated in relation to promotion. Indeed, given selection for the RAE may be a factor in gaining promotion, it is possible that unfair treatment with respect to the RAE could cause an unfair lack of promotion. For this reason, including grade and other employment factors is problematic. We found that if a restricted model was used, which did not include grade, not only Black but also Asian and Other ethnic groupings had lower than expected selection rates. On the other hand, the evidence from the bibliometrics shows that grade was associated with research strength and that, therefore, a model without these employment factors may not be sufficient to take account of the

quality of research output of individual researchers. Finally, we find no difference in the research strength, as measured using bibliometrics, of staff selected for RAE2001 from different ethnic groups, which is consistent with these staff being treated fairly in the selection process.

The RAE and equal opportunities

100. The differences in rates of selection for submission to the RAE between male and female staff are highly visible, and, rightly, have been subject of much comment. The evidence, though not conclusive, suggests that these differences are not specific to the RAE, but rather reflect more deeply rooted inequalities in the research careers of men and women. This study has also shown that selection rates also vary by ethnicity, which is a cause for concern, though it seems likely that the observed differences are due to differences where staff are employed, and in research strength, rather than a bias against particular groups. While there is clearly a need for further measures to address the specific issue of selection to the RAE, it would be unfortunate if the transparency of the RAE led to an unmerited conclusion that it was the cause of inequalities in research careers. Indeed, without the relatively stable public funding through block grant that the RAE underpins, HEIs would be much more dependent upon shorter term project funding and it would be much more difficult to improve the career opportunities for academic staff.

The 2008 RAE and beyond

101. Three out of four staff selected for RAE2001 were male. This in part reflects the fact that 68 per cent of staff who could have been selected were male, and in part, as we have seen, that a smaller proportion of female staff were selected. Even if there were no changes, either to the assessment protocols or the practice of institutions, we would expect to see the proportion of women among selected staff to increase from 25 per cent to about 30 per cent.[7] In part this is due to the increasing proportion of eligible staff who will be female, projected to be 37 per cent by 2008. We would also expect the relative selection rate for women compared to men to increase given the increasing proportion of women in the higher grades, with, for example, the proportion of professors who are women rising from 15 per cent to 20 per cent.[8]

102. Changes to the RAE for 2008 should further reduce the disadvantages experienced by women. Firstly, rather than score research quality as a single value, panels will produce a research quality profile. This means that the risk of losing recognition and funding for high quality research through the selection of staff with lower quality research output is removed. This is not to say that institutions will not continue to be selective in their choice of staff when preparing RAE submissions, but the pressure to be very highly selective should be greatly reduced.

---

[7] Based on 'current recruitment' protocol (HEFCE, 2002). Figures from 2008-09 used, 25,652 female and 44,517 male staff. For the latest figures on past trends see HEFCE, 2006.

[8] Reference as in previous footnote. Figures used in 2008-09, 2,819 female and 11,242 male professors.

103. More directly, the 2008 RAE gives increased prominence to equal opportunities. Taking advice from the Equality Challenge Unit, the 2008 RAE team has produced guidance for panels (Funding Bodies 2005a) and institutions (Funding Bodies 2005b), designed to ensure that the maximum number of staff who are conducting excellent research have their work included in submissions, and that staff whose research output has been limited by, for example, absence due to maternity leave, are included and their work assessed taking these circumstances into account. Institutions are required to develop their own code of practice which should ensure that there is transparency in the process of deciding which staff should be selected. Furthermore, institutions are required to see if there are any prima facie imbalances in the proportions of staff selected and to account for any such imbalances.

104. As with RAE2001, monitoring at a sector-wide level will be carried out by the funding bodies using the HESA staff record. Enhancements to this record mean that all academic staff employed by the HEIs will be included within the record. For the first time, the data on selected research active staff submitted to the RAE will include the individual staff identifier used within the HESA staff record, and HESA will ensure that the two returns are consistent during the HESA data collection. In addition, institutions will be asked to ensure that their HESA records are consistent with their internal monitoring as described above. These measures should ensure that data of the required accuracy are available to carry out a full investigation of selection rates following the 2008 RAE.

105. These measures should help to ensure that the next RAE in 2008 is fair to all academic staff. But if, as seems likely, the observed differences in rates of selection were largely due to the relative quality of the research output of different groups, these changes to the RAE process will not of themselves greatly reduce the differences in rates of selection. There is therefore a need to tackle the underlying causes for the differences in research output. For this reason, our equal opportunities policies, which go beyond those specifically related to the RAE, are relevant to this discussion.

106. HEFCE has an ongoing programme giving general support to HEIs' equal opportunities policies. Funding has been provided specifically for rewarding and developing staff, and this funding was conditional on institutions producing human resources strategies which include equal opportunities policies. HEFCE also supports the Equality Challenge Unit (ECU), the Leadership Foundation for Higher Education and other related organisations which have ongoing initiatives to support institutions in embedding and mainstreaming equality and diversity. Examples include the ECU's range of good practice guidance related to employment equality and cultural change within HEIs, the Leadership Foundation's funding of senior diversity champions in the higher education sector, and our good management practice project on flexible employment options.[9] A direct impact of this last initiative should be to reduce the impact of career breaks on individual academics' research programmes.

107. Looking beyond the 2008 RAE, discussions are already under way as to how research should be assessed after this exercise (DfES, 2006). Before any new operations are confirmed and implemented they will be subject to equality impact

[9] See the 'Flexible Employment Options' project at www.staffs.ac.uk/feo.

assessments (HEFCE, 2004), and such assessments will be informed by the analysis of the 2001 RAE reported here and our planned analysis of the 2008 RAE.

References

Adams, J (2004), 'Relating bibliometrics to UK research assessment grades', Oxford Workshop on the use of Metrics in Research Assessment, Oxford, September 2004

Adams, J (2006), 'How good is the UK research base?', Higher Education Policy Institute (HEPI) research report, 22 June 2006. Available at www.hepi.ac.uk.
Also to be published as: Adams, J, Gurney, K and Marshall, S (2007, in press), 'Profiling citation impact: a new methodology', Scientometrics

AUT (2004), 'UK academic staff 2002-03: Gender and research activity in the 2001 Research Assessment Exercise'
Available at: www.aut.org.uk/media/pdf/genderandrae_jun04.pdf

Davenport, E and Snyder, H (1995), 'Who cites women? Whom do women cite?: an exploration of gender and scholarly citation in sociology'. Journal of Documentation, 51 (1995) 404-410.

DfES (2006), 'Reform of higher education research assessment and funding'. Available at: www.dfes.gov.uk/consultations

Ferber, M (1988), 'Citations and Networking', Gender and Society, 2 (1), (1988), 82-89

Funding Bodies (2004), 'RAE 2008: Initial decisions by the UK funding bodies', Reference RAE 01/2004. Available at: www.rae.ac.uk/pubs/2004/01/

Funding Bodies (2005a), 'Equality briefing for panel chairs, members and secretaries', Reference RAE 02/2005. Available at: www.rae.ac.uk/pubs/2005/02/

Funding Bodies (2005b), 'Guidance on submissions', Annex G, Reference RAE 03/2005. Available at: www.rae.ac.uk/pubs/2005/03/

Hakanson, M (2005), 'The impact of gender on citations: an analysis of College and Research Libraries, Journal of Academic Librarianship, and Library Quarterly, 66 (4), 312-322.

HEFCE (2000), 'Review of research', HEFCE 00/37, Annex J. Available at: www.hefce.ac.uk/pubs/hefce/2000/00_37.htm.

HEFCE (2002), 'Academic staff: trends and projections', HEFCE 02/43. Available at: www.hefce.ac.uk/pubs/hefce/2002/02_43.htm

HEFCE (2003), 'Joint consultation on the review of research assessment' HEFCE 2003/22. Available at: www.hefce.ac.uk/pubs/hefce/2003/03_22.htm

HEFCE (2004), 'Conducting impact assessments for equal opportunities in higher education', HEFCE 2004/37. Available at: www.hefce.ac.uk/pubs/hefce/2003/04_37.htm

HEFCE (2006), ‘The higher education workforce in England: A framework for the future ‘ HEFCE 2006/21. Available at: www.hefce.ac.uk/pubs/hefce/2006/06_21.htm

HESA (2002), ‘Reference Volume: Resources of Higher Education Institutions 2000/2001’

Annex A
Terminology and abbreviations

Terms

Academic staff	Staff employed under a contract of salaried employment with the HEI whose primary employment function is teaching, research or both.
Department	The staff associated with a **submission** or submissions**,** if multiple submissions are made, from an HEI to the same UOA. Even if there is no **submission**, the **eligible staff** will still be associated with a particular **UOA.** These staff will often be part of a single administrative unit, but this is not a necessary condition and will not always be the case.
Eligible staff	This term refers to staff who are eligible for inclusion in the **submission** to the RAE, that is staff whose research outputs may be included in the **submission**. These are those **academic staff** who are not employed by the HEI to carry out another individual's research programme, so **research assistants** are not usually eligible. Note that the glossary of terms in RAE 03/2005 (www.rae.ac.uk/pubs/2005/03/) provides a more restrictive definition.
Non-submitting department	A department where there are no **submissions**.
Quality rating	This is a measure of the quality of research described in a submission. There is a seven point scale from the lowest assigned quality 1, to the highest 5*: 1, 2, 3b, 3a, 4, 5, 5*. In the 2008 RAE the quality rating will be replaced by a quality profile showing the proportion of activity in four levels.
Submitting department	A department where there is at least one **submission**.
Research assistants	Staff who are employed by an HEI to carry out another individual's research programme. They are often on short-term projects funded by research grants. Exceptionally they may take the role as a principal investigator, and in such circumstances they may be classed as **eligible staff**, though generally they will not be.
Research outputs	Publicly available assessable outcomes of the research of **selected staff** or, if confidential, available to be assessed. Each **selected staff** may submit up to a maximum of four **research outputs**. Bibliometrics can only be used for those research outputs which take the form of articles in the academic journals included in citation databases.

Selected staff	**Eligible staff** whose **research outputs** are included in an RAE **submission**.
Selection index	When using simple summary statistics this a ratio of odds ratios based on the selection rate of one particular group of staff and the selection rate of a reference group of staff. S_j x (100 – S_r) / S_r x (100 – S_j) Where S_j = selection rate of j^{th} staff group S_r = selection rate of reference staff group When based on a model the selection index is the exponential of the coefficient identifying the staff group.
Selection rate	100 x (Number of **selected staff**) / (Number of **eligible staff**)
Submission	A set of information provided to the RAE by an HEI pertaining to a **UOA**. The submission received a **quality rating.** In a few cases HEIs made more than one submission for one UOA; these are referred to as multiple submissions.
Unit of assessment (UOA)	One of 68 discipline areas to which 2001 RAE **submissions** may have been made by an institution.
UOA within HEI	The **submissions** associated with a **UOA** for a particular HEI. Usually identical to a **submission**. Used as an approximation to a **submission** for most of the analysis in this report.

Abbreviations

FTE	Full-time equivalent
HEFCE	Higher Education Funding Council for England
HEI	Higher education institution
HESA	Higher Education Statistics Agency
n/a	Not applicable
RAE	Research Assessment Exercise in general
RAE2001	Research Assessment Exercise that took place in 2001
Ref	The reference group used to calculate the selection index
UOA	RAE unit of assessment

Annex B
HESA data – definitions and quality checks

1. This annex gives details of the derivation of the base data used in constructing the data set used in the modelling. Throughout the annex, fields taken from the HESA record are given in capitals using the field names from the HESA coding manual. Data used in the modelling were derived from a modified version of the 2000-01 HESA Individualised Staff Record.

Data cleaning

2. The modified version of the HESA record was first restricted to those staff recorded as category A, A* or C (RESACT = 1, 2) and then further restricted to only include those staff who were in post on the census date and had been assigned to a UOA (UOA ≠ "")

3. Data cleaning occurred in two stages. In the first stage English HEIs were queried where there were significant differences between RAE returns and the data submitted to the HESA staff record. Where the HESA staff record was shown to be in error, corrections to the HESA staff record provided by the HEI were incorporated into the copy of the record held by HEFCE. In addition, staff attributed to multiple submissions on the HESA database where only a single submission to the RAE was made were re-coded to be in the single submission made. Where a member of staff has two records within the HESA staff record for 2000-01, only one record was kept for the purposes of this analysis. The record to keep was decided via the following hierarchy:

- submitted to the 2001 RAE
- most senior
- full-time over part-time
- permanent over non-permanent.

4. Despite this data cleansing some suspect data remained in the copy of the HESA staff record held by HEFCE. To reduce the risk of these data quality problems determining the conclusions of our analysis three sets of data were used:

a. All HESA records of eligible staff after making corrections from the cross-checking exercise.

b. Data as in 'a' above with data from highly suspect UOAs within HEIs or complete HEIs removed.

c. Data as in 'a' above with data from moderately suspect UOAs within HEIs or complete HEIs removed.

5. The identification of highly or moderately suspect UOAs was based on the headcount of category A and A* staff selected who were in post on the RAE census date as derived from the RAE and HESA returns.

6. Staff recorded as selected in units of assessment where the HEI made no submission were treated as not selected, and the records treated as unsafe when determining whether an institution's data as a whole could be treated as safe.

7. UOAs were treated as highly suspect if the difference between the number of staff submitted on the HESA and RAE returns was more than 2 and this represented at least 15 per cent of the staff submitted to the RAE.

8. UOAs were treated as moderately suspect if the difference between the number of staff submitted on the HESA and RAE returns was more than 2 or this represented at least 45 per cent of the staff submitted to the RAE.

9. Data for a whole HEI were removed if more than a third of the staff recorded as submitted were either in UOAs that were suspect or in UOAs that were not submitted to the exercise.

Ethnicity groupings

10. In this analysis seven ethnicity groupings were used. The groupings were derived from the more detailed classification used on the HESA staff record using the mapping given in Table B1.

Table B1: Mapping to ethnicity groups

Ethnicity group	Detailed descriptors
White	ETHNIC = 10
Black	ETHNIC = 21, 22, 29
Indian	ETHNIC = 31
Chinese	ETHNIC = 24
Other Asian	ETHNIC = 32, 33, 39
Other	ETHNIC = 80
Information refused	All other staff

Grade

11. Four broad grade groups were defined for use in this analysis. The grade groups were primarily based on the GRADE field on the HESA staff record. Where no grade was given a proxy was derived based on banded salary data. The mapping used is given in Table B2.

Table B2: Mapping to grade

Grade group	Detailed descriptors
Professor	GRADE = 04, 34, 39, 42, 63, 71 or (GRADE = 29, 59, 65, 99 and salary > £40,000)
Senior lecturer	GRADE = 03, 33, 38, 41, 62, 72 or (GRADE = 29, 59, 65, 99 and SALARY > £35,000 and SALARY ≤ £40,000)
Lecturer	GRADE = 01, 02, 31, 32, 40, 61, 73 or (GRADE = 29, 59, 65, 99 and (SALARY > £30,000 and SALARY ≤ £35,000 or (SALARY < £35,000 and TERMS = 1)))
Research assistant and other	GRADE = 05, 06, 35, 36, 37, 64, 74 or (GRADE = 29, 59, 65, 99 and SALARY < £35,000 and Terms ≠ 1)

Other groupings

12. Only two modes of employment were used in the model; all part-time and casual or hourly paid staff (MOE = 2, 3, 9) were grouped together as part-time.

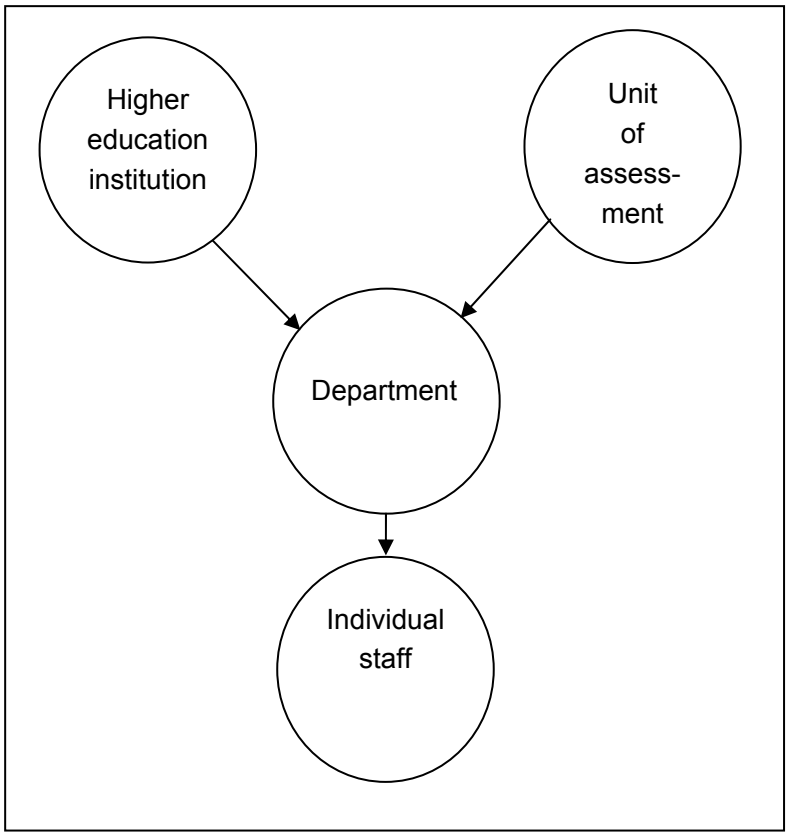
13. All staff who were not on either permanent or fixed term contracts (TERMS ≠ 1, 2) were grouped together as casual.

14. All staff who were not recorded as having a disability (DISABLE = 2, 3) were treated as not disabled.

Annex C
Model of staff selection

1. The statistical models from which these results are derived are cross-classified multi-level models. The schematic for the structure for the models is given in Figure C1.

Figure C1 Schematic of the structure for the model



2. Figure C1 shows that individual staff are assumed to be within a department within a higher education institution. Individual departments are also assumed to be within a unit of assessment, giving a cross-classification at the higher level.

3. The statistical form of the full model is given in Figure C2.

Figure C2 Full model form

$$\text{selected}_{ijkl} \sim \text{Binomial}(\text{cons}_{ijkl}, \pi_{ijkl})$$

$$\begin{aligned}\text{probit}(\pi_{ijkl}) = {} & \beta_{0jkl}\text{cons} + \beta_1\text{Age}_{ijkl} + \beta_2\text{age.age}_{ijkl} + \beta_3\text{age.age.age}_{ijkl} + \\ & \beta_4\text{age.age.age.age}_{ijkl} + \beta_5\text{T1}_{ijkl} + \beta_6\text{T2}_{ijkl} + \beta_7\text{Rating1}_{ijkl} + \beta_8\text{Rating2}_{ijkl} + \\ & \beta_9\text{Rating3a}_{ijkl} + \beta_{10}\text{Rating3b}_{ijkl} + \beta_{11}\text{Rating5}_{ijkl} + \beta_{12}\text{Rating5*}_{ijkl} + \\ & \beta_{13}\text{GradeGp1}_{ijkl} + \beta_{14}\text{GradeGp2}_{ijkl} + \beta_{15}\text{GradeGp3}_{ijkl} + \beta_{16}\text{PartTime}_{ijkl} + \\ & \beta_{17}\text{PF1}_{ijkl} + \beta_{18}\text{PF2}_{ijkl} + \beta_{19}\text{Male}_{ijkl} + \beta_{20}\text{Univ}_{l} + \beta_{21}\text{Age.Rating5}_{ijkl} + \\ & \beta_{22}\text{Rating3a.GradeGp3}_{ijkl} + \beta_{23}\text{Rating5.GradeGp1}_{ijkl} + \\ & \beta_{24}\text{Rating5.GradeGp2}_{ijkl} + \beta_{25}\text{Rating5.GradeGp3}_{ijkl} + \\ & \beta_{26}\text{Rating5*.Gradegp3}_{ijkl} + \beta_{27}\text{Rating2.PF2}_{ijkl} + \beta_{28}\text{Rating3a.PF1}_{ijkl} + \\ & \beta_{29}\text{Rating3b.PF1}_{ijkl} + \beta_{30}\text{Rating5*.PF1}_{ijkl} + \beta_{31}\text{T1.Rating3b}_{ijkl} + \\ & \beta_{32}\text{T1.Rating5}_{ijkl} + \beta_{33}\text{T2.Rating3b}_{ijkl} + \beta_{34}\text{T2.Rating5}_{ijkl} + \\ & \beta_{35}\text{Age.GradeGp2}_{ijkl} + \beta_{36}\text{Age.GradeGp3}_{ijkl} + \beta_{37}\text{Age.T2}_{ijkl} + \\ & \beta_{38}\text{Age.Age.T2}_{ijkl} + \beta_{39}\text{Age.PartTime}_{ijkl} + \beta_{40}\text{Age.PF1}_{ijkl} + \\ & \beta_{41}\text{Age.Age.PF1}_{ijkl} + \beta_{42}\text{Age.PF2}_{ijkl} + \beta_{43}\text{GradeGp2.PF1}_{ijkl} + \\ & \beta_{44}\text{GradeGp3.PF1}_{ijkl} + \beta_{45}\text{GradeGp2.PF2}_{ijkl} + \beta_{46}\text{PartTime.PF2}_{ijkl} + \\ & \beta_{47}\text{Age.Male}_{ijkl} + \beta_{48}\text{Age.Age.Male}_{ijkl} + \beta_{49}\text{Age.Age.Age.Male}_{ijkl} + \\ & \beta_{50}\text{WithPhd}_{ijkl} + \beta_{51}\text{UG1.WithPhD}_{ijkl} + \beta_{52}\text{UG2.WithPhD}_{ijkl} + \\ & \beta_{53}\text{UG3.WithPhD}_{ijkl} + \beta_{54}\text{WithPhD.Age}_{ijkl} + \beta_{55}\text{WithPhD.Age.Age}_{ijkl} + \\ & \beta_{56}\text{Univ.WithPhd}_{ijkl} + \beta_{57}\text{PartTime.WithPhD}_{ijkl} + \beta_{58}\text{CommonSub}_{ijkl} + \\ & \beta_{59}\text{CommonSub.WithPhD}_{ijkl} + \beta_{60}\text{Senior}_{ijkl} + \beta_{61}\text{Clinical}_{ijkl} + \\ & \beta_{62}\text{Prev\_ORes}_{ijkl} + \beta_{63}\text{Prev\_Other}_{ijkl} + \beta_{64}\text{GradeGp1.Previous\_ORes}_{ijkl} + \\ & \beta_{65}\text{GradeGp2.Previous\_ORes}_{ijkl} + \beta_{66}\text{GradeGp3.Previous\_ORes}_{ijkl} + \\ & \beta_{67}\text{GradeGp1.Previous\_Other}_{ijkl} + \beta_{68}\text{GradeGp2.Previous\_Other}_{ijkl} + \\ & \beta_{69}\text{GradeGp3.Previous\_Other}_{ijkl} + \beta_{70}\text{Black}_{ijkl} + \beta_{71}\text{Asian}_{ijkl} + \beta_{72}\text{Other}_{ijkl} + \\ & \beta_{73}\text{Refused}_{ijkl} + \beta_{74}\text{Disable}_{ijkl}\end{aligned}$$

$$\beta_{0jkl} = \beta_0 + f_{0l} + v_{0kl} + u_{0jkl}$$

$$\begin{bmatrix} f_{0l} \end{bmatrix} \sim \text{N}(0,\ \Omega_f)\ :\ \Omega_f = \begin{bmatrix} \sigma^2_{f0} \end{bmatrix}$$

$$\begin{bmatrix} v_{0kl} \end{bmatrix} \sim \text{N}(0,\ \Omega_v)\ :\ \Omega_v = \begin{bmatrix} \sigma^2_{v0} \end{bmatrix}$$

$$\begin{bmatrix} u_{0jkl} \end{bmatrix} \sim \text{N}(0,\ \Omega_u)\ :\ \Omega_u = \begin{bmatrix} \sigma^2_{u0} \end{bmatrix}$$

$$\text{var}(\text{selected}_{ijkl}|\pi_{ijkl}) = \pi_{ijkl}(1 - \pi_{ijkl})/\text{cons}_{ijkl}$$

$$\textit{Deviance(MCMC)} = 42557.39 (62829 \text{ of } 62829 \text{ cases in use})$$

where *i* represent the individual, *j* represents the sector-wide unit of assessment, *k* represents a particular unit of assessment within a particular institution (*l*). The variables in the model are defined in Table C1.

4. The statistical form of the restricted model is given in Figure C3.

Figure C3 Restricted model form

$$\text{Selected}_{ijkl} \sim \text{Binomial}(\text{Cons}_{ijkl}, \pi_{ijkl})$$

$$\begin{aligned}\text{probit}(\pi_{ijkl}) = {} & \beta_{0jkl}\text{Cons} + \beta_1\text{Age}_{ijkl} + \beta_2\text{Age.Age}_{ijkl} + \beta_3\text{Age.Age.Age}_{ijkl} + \\ & \beta_4\text{Age.Age.Age.Age}_{ijkl} + \beta_5\text{Rating1}_{ijkl} + \beta_6\text{Rating2}_{ijkl} + \beta_7\text{Rating3a}_{ijkl} + \\ & \beta_8\text{Rating3b}_{ijkl} + \beta_9\text{Rating5}_{ijkl} + \beta_{10}\text{Rating5*}_{ijkl} + \beta_{11}\text{PF1}_{ijkl} + \beta_{12}\text{PF2}_{ijkl} + \\ & \beta_{13}\text{Male}_{ijkl} + \beta_{14}\text{Univ}_{l} + \beta_{15}\text{Age.Rating5}_{ijkl} + \beta_{16}\text{Rating2.PF2}_{ijkl} + \\ & \beta_{17}\text{Rating3a.PF1}_{ijkl} + \beta_{18}\text{Rating3b.PF1}_{ijkl} + \beta_{19}\text{Rating5*.PF1}_{ijkl} + \\ & \beta_{20}\text{Age.PF1}_{ijkl} + \beta_{21}\text{Age.Age.PF1}_{ijkl} + \beta_{22}\text{Age.PF2}_{ijkl} + \beta_{23}\text{Age.Male}_{ijkl} + \\ & \beta_{24}\text{Age.Age.Male}_{ijkl} + \beta_{25}\text{Age.Age.Age.Male}_{ijkl} + \beta_{26}\text{WithPhd}_{ijkl} + \\ & \beta_{27}\text{UG1.WithPhd}_{ijkl} + \beta_{28}\text{UG2.WithPhd}_{ijkl} + \beta_{29}\text{UG3.WithPhd}_{ijkl} + \\ & \beta_{30}\text{WithPhd.Age}_{ijkl} + \beta_{31}\text{WithPhd.Age.Age}_{ijkl} + \beta_{32}\text{Univ.WithPhd}_{ijkl} + \\ & \beta_{33}\text{CommonSub}_{ijkl} + \beta_{34}\text{Commonsub.WithPhd}_{ijkl} + \beta_{35}\text{Clinical}_{ijkl} + \\ & \beta_{36}\text{Previous\_ORes}_{ijkl} + \beta_{37}\text{Previous\_Other}_{ijkl} + \beta_{38}\text{Black}_{ijkl} + \beta_{39}\text{Asian}_{ijkl} + \\ & \beta_{40}\text{Other}_{ijkl} + \beta_{41}\text{Refused}_{ijkl} + \beta_{42}\text{Disable}_{ijkl}\end{aligned}$$

$$\beta_{0jkl} = \beta_0 + f_{0l} + v_{0kl} + u_{0jkl}$$

$$\left[f_{0l}\right] \sim \text{N}(0,\ \Omega_f)\ :\ \Omega_f = \left[\sigma^2_{f0}\right]$$

$$\left[v_{0kl}\right] \sim \text{N}(0,\ \Omega_v)\ :\ \Omega_v = \left[\sigma^2_{v0}\right]$$

$$\left[u_{0jkl}\right] \sim \text{N}(0,\ \Omega_u)\ :\ \Omega_u = \left[\sigma^2_{u0}\right]$$

$$\text{var}(\text{Selected}_{ijkl}|\pi_{ijkl}) = \pi_{ijkl}(1 - \pi_{ijkl})/\text{Cons}_{ijkl}$$

Deviance(MCMC) = 47839.(62829 of 62829 cases in use)

The subscript and variables definitions are as in the full model

Table C1 Variables in the full and restricted models

Type	Model variable name	Description
Continuous	Age	Individual's age (in years)
Dummy/Categorical	T	Terms of employment: Permanent(1); Fixed term(2); Casual/hourly(REF)
	Rating	RAE rating of department: 1, 2, 3b, 3a, 4(REF), 5, 5*
	GradeGp	Individual's grade: Professor (1); Senior lecturer(2); Lecturer(3); Researcher(REF)
	PF	Primary employment function: Teaching only (1); Research only (2); T and R(REF)
	Previous	Employment in previous year: Current HEI (REF); Other research institution (Previous_ORes); Other (Previous_Other)
	UG	Group of UOAs: Arts & vocational (1); Clinical (2); Humanities, social sciences & languages (3); Engineering & sciences(REF)
	Refused, Black, Asian, Other	Ethnicity of individual: Black; Refused; Asian; White(REF); and Other
Single dummy	PartTime	On a part-time contract [Full time-contract(REF)]
	Male	Male [Female(REF)]
	Univ	Pre-1992 HEI [Post-1992 HEI(REF)]
	WithPhD	Individual holds a PhD as their highest qualification [Does not hold a PhD(REF)]
	CommonSub	An individual's subject of highest qualification is common to individuals in the associated UOA [Not associated(REF)]
	Clinical	Staff on clinical rates [Not on clinical rates(REF)]
	Senior	Senior management post holder [Not a senior management post holder(REF)]
	Disable	Disability recorded [No disability recorded(REF)]
Structural	Cons	One for all individuals
	F	Random effect relating to a particular institution
	V	Random effect relating to a particular unit of assessment within an institution
	U	Random effect relating to the sector wide unit of assessment

Table C1 note: Those categories marked with REF are the reference categories for each categorical or dummy variable and are not formally included in the model structure.

5. The parameter estimates and the associated standard errors for each model are given in Table C2.

Table C2 Parameter estimates for the full and restricted models

Parameter	Full model		Restricted model	
	Estimate	SE	Estimate	SE
Cons	-12.40	2.00	-8.29	1.68
Age	1.03	0.17	0.69	0.15
Age.Age	-0.03	0.01	-0.02	0.00
Age.Age.Age	0.00	0.00	0.00	0.00
Age.Age.Age.Age	0.00	0.00	-1.18	0.00
T1	0.62	0.14		
T2	-0.93	0.38		
Rating1	-0.07	0.18	-0.11	0.18
Rating2	-0.37	0.08	-0.45	0.07
Rating3a	-0.34	0.06	-0.31	0.05
Rating3b	0.87	0.30	-0.21	0.04
Rating5	1.43	0.34	0.33	0.08
Rating5*	0.10	0.06	0.29	0.05
GradeGp1	1.04	0.06		
GradeGp2	1.96	0.15		
GradeGp3	1.03	0.12		
PartTime	-0.47	0.12		
PF1	1.83	0.70	1.74	0.67
PF2	-0.28	0.14	-1.44	0.10
Male	-4.15	1.32	-4.25	1.22
Univ	0.78	0.15	0.75	0.14
Age.Rating5	-0.01	0.00	0.00	0.00
Rating3a.GradeGp3	0.15	0.05		
Rating5.GradeGp1	0.33	0.08		
Rating5.GradeGp2	0.20	0.07		
Rating5.GradeGp3	0.16	0.07		
Rating5*.GradeGp3	0.20	0.05		
Rating2.PF2	0.41	0.18	0.45	0.17
Rating3a.PF1	0.58	0.14	0.55	0.14
Rating3b.PF1	0.37	0.13	0.47	0.12
Rating5*.PF1	0.62	0.13	0.70	0.13
T1.Rating3b	-1.05	0.29		
T1.Rating5	-0.82	0.33		
T2.Rating3b	-0.98	0.29		
T2.Rating5	-0.96	0.33		
Age.GradeGp2	-0.04	0.00		
Age.GradeGp3	-0.03	0.00		
Age.T2	0.04	0.02		
Age.Age.T2	0.00	0.00		
Age.PartTime	0.01	0.00		
Age.PF1	-0.14	0.03	-0.12	0.03
Age.Age.PF1	0.00	0.00	0.00	0.00
Age.PF2	0.01	0.00	0.03	0.00
GradeGp2.PF1	0.41	0.13		
GradeGp3.PF1	0.47	0.11		
GradeGp2.PF2	0.44	0.11		
PartTime.PF2	-0.18	0.06		

Age.Male	0.30	0.09	0.29	0.08
Age.Age.Male	-0.01	0.00	-0.01	0.00
Age.Age.Age.Male	0.00	0.00	0.00	0.00
WithPhD	0.00	0.03	-0.52	0.30
UG1.WithPhD	0.10	0.04	0.12	0.04
UG2.WithPhD	-0.01	0.05	-0.03	0.05
UG3.WithPhD	0.00	0.05	0.00	0.04
WithPhD.Age	0.05	0.01	0.07	0.01
WithPhD.Age.Age	0.00	0.00	0.00	0.00
Univ.WithPhD	-0.32	0.04	-0.29	0.03
PartTime.WithPhD	-0.23	0.04		
CommonSub	-0.04	0.02	-0.03	0.02
CommonSub.WithPhD	0.06	0.04	0.10	0.03
Senior	-0.22	0.05		
Clinical	-0.56	0.04	-0.26	0.04
Previous_Ores	-0.14	0.09	0.24	0.03
Previous_Other	-0.55	0.07	-0.41	0.03
GradeGp1.Previous_Ores	0.83	0.19		
GradeGp2.Previous_Ores	0.51	0.14		
GradeGp3.Previous_Ores	0.37	0.10		
GradeGp1.Previous_Other	0.47	0.17		
GradeGp2.Previous_Other	0.56	0.12		
GradeGp3.Previous_Other	0.24	0.08		
Black	-0.14	0.07	-0.25	0.07
Asian	0.01	0.04	-0.10	0.04
Other	-0.07	0.05	-0.13	0.05
Refused	-0.04	0.03	-0.10	0.02
Disable	-0.03	0.07	-0.01	0.07
Variation (Institution)	0.67	0.11	0.62	0.10
Variation(UOA within Institution)	0.18	0.01	0.16	0.01
Variation(Sector wide UOA)	0.09	0.02	0.07	0.02

Annex D
Bibliometrics – data extraction and linking to RAE data

1. A file was prepared by Evidence Ltd with bibliometric data appended to output records from the 2001 RAE. This technical annex was prepared from documents provided by Evidence.

Bibliometric data and the RAE – an overview

2. Research quality can be indexed for comparative purposes by using citation records as a proxy for quality. Journal articles report research work and refer to or 'cite' earlier work relevant to the work being reported. Papers that accumulate more citations are thought of as having greater significance or influence in their field.

3. Data comparable to journal citations are not readily available for books or chapters, and citation data on conference proceedings are very patchy. Citation counts are therefore largely based on articles published in journals.

4. For UK work, citation data can be conveniently extracted from the Thomson Scientific (formerly Thomson ISI ®) National Citation Report (NCR) database of UK journal article publications which is available for the period to the end of 2003 (NCR2003).

5. To be included in RAE2001, an article had to be published during the period 1996-2000 (or 1994-2000 for some arts-based UOAs). Citations accumulate over time so older papers have, on average, more citations than more recent work. At the census date for assessment there would have been many articles with little or no time to demonstrate their bibliometric significance. Now, a few years later, the papers submitted to RAE2001 have had a reasonable time to accumulate citations and there is between three and seven (or nine) years of citation data for each RAE article, depending on when it was published.

6. Citations per article vary between fields and between journals within a field so it is necessary to assess citation rates in context, usually relative to field or relative to journal. Because citations accumulate over time, the year of publication must also be taken into account when rebasing citation counts to journal or field averages.

7. It is assumed that researchers submit to the RAE those pieces of work that best demonstrate their research quality. For RAE2001 each researcher submitted up to four items which were classified by publication type. Journal articles (Publication Type D) form a varying percentage of RAE submissions across UOAs.

Selection of UOAs

8. The use by researchers of research articles as a preferred output mode varied between UOAs. Because of the nature of the Thomson Scientific databases, the proportion of research articles that might be reconciled to the NCR also varies between UOAs. On the whole, articles are more common and more frequently reconciled in the sciences and less so in the humanities.

9. UOAs initially selected in our study were those where:

a. Journal articles were a common publication type among the records submitted to the RAE, and were therefore deemed by the community to be a good reflection of their high quality research output.

b. Previous work had shown that many or most journals widely used by the UK research community were catalogued on the databases assembled by Thomson Scientific. (The Thomson data has greater richness in the bio-medical area and has a North American predominance in some disciplines.)

c. Previous work had established a significant correlation between bibliometric measures of research performance and other absolute and relative (per FTE) measures associated with research funding, postgraduate output, and so on.

d. The UOA included enough researchers to support a reasonable level of comparison.

10. Overall, nine UOAs (01, 02, 13, 14, 18, 21, 22, 28 and 30) were identified as ones where bibliometrics based on journal publication would be a good measure of research quality. This set also provided a broad subject spread.

11. To these UOAs which best met these criteria four other UOAs were included to explore whether the approach could be extended. These were:

a. UOA 16 (Food Science), which, though smaller than some UOAs, had more women researchers than most science subjects.

b. UOA 35 (Geography) and UOA 42 (Sociology) to try and extend coverage in the social sciences.

c. UOA 59 (History) to see if at least one humanities subject could be analysed

After further investigation, Sociology and History were excluded from the main analysis.

Linking RAE and Thomson Scientific databases

12. The aim was to determine citation counts for RAE RA2 source records by linking them with articles on the Thomson Scientific NCR database.

Data cleaning

13. In order to link the RA2 data to the NCR, the data in corresponding fields had to be standardised. For example, formatting and style of fields denoting year of publication needed to be the same so as to enable a simple match. Each article in the RA2 dataset was assigned to a known journal on an Evidence reference list. This was achieved by extensive electronic and manual processing of the data held in the RAE fields for ISSN (International Standard Serial Number) and OutputLoc ('output location', ie journal title for type D records) fields.

14. In addition to creating a match for OutputLoc against Journal Title, four other RA2 fields were used to identify and verify unique matches against NCR records. These were:

a. Year (cleaning usually not needed).

b. Volume (cleared of non-numeric characters).

c. Pagination (article start page extracted into derived field EditStartPage).

d. Plain title (converted to a 255 character text field; leading and trailing quotes removed where both were present; spurious characters and unnecessary punctuation removed; converted to uppercase).

Electronic matching

15. RA2 records were electronically matched to NCR records by linking on Journal, Year, Volume and StartPage.

16. These matches were expected to be from an RAE record to a unique NCR record. An NCR record may match several RAE records if the same article was submitted by multiple authors. Where the match to the NCR was not unique, the data record was extracted, checked and re-matched. If no unique match were possible then the record would have been set aside.

17. Where records matched on all four of these fields, the first 63 characters of the RA2 PlainTitle were then compared with the first 63 characters of the NCR ARTL_TITLE. If these character strings were identical, the articles were deemed to be successfully matched.

Visual inspections

18. A sample of records where the initial title strings were not identical was subjected to visual review. After a detailed review for the reasons for these differences all such records were included.

19. In addition, a further set of records that matched on three of the four key fields were identified. All records that matched either on all key fields except volume or on all key fields except start page were subjected to visual review. These were deemed to be not matched, and were excluded unless this visual inspection verified the match.

Results of linking

Error rates in matched records

20. RAE records which were uniquely matched to the NCR on Journal, Year, Volume and Start Page, but which did not match on the first 63 characters of the Plain Text title, were sub-sampled for visual checking as noted above. A randomised 10 per cent were checked by this route, and 1 per cent of these were found to have a possibility of mismatch (in other words, incompleteness or inconsistencies in the RA2 data meant that a unique match could not be absolutely verified). This remains a possibility, however, rather than a certainty. Such records were reviewed by a second operator, and it was agreed that the original author would be required to enable a final conclusion. On the whole, it was felt that these represented no more than 'uncertainty'.

21. This translates to an expected 'uncertainty' level of 0.3 per cent in the full dataset. Within this grey area, no systematic bias is expected, nor was there evidence of significant variation in uncertainty or error rates between UOAs.

Non-matched records

22. Twenty-one per cent of the RA2 output type D articles in the 13 selected UOAs could not be matched to NCR records. About half of these were not matched because they were in journals that have not been catalogued by Thomson Scientific. The remaining 10 per cent were unmatched because of such reasons as:

a. RA2 data are incomplete or erroneous.
b. The journal uses a non-standard format for pagination or volume.
c. The journal was not indexed by Thomson when the article was published.
d. None of the article's authors had UK addresses at the time of publication.

These factors apply equally across UOAs and there is no reason to suppose that they would affect articles submitted by women differently from articles submitted by men.

Matches achieved

23. Table D1 shows the numbers of RAE output records that were matched to bibliometric data.

Table D1: Proportion of records matched

UOA code	Subject	RA2 published records	RA2 type D published records	Unique NCR Matches	% RA2 records matched	% type D records matched
01	Clinical Laboratory Sciences	4,955	4,921	4,373	88%	89%
02	Community Based Clinical Subjects	5,502	5,327	4,610	84%	87%
13	Psychology	5,136	4,752	3,670	71%	77%
14	Biological Sciences	9,931	9,726	8,403	85%	86%
16	Food Science and Technology	476	447	378	79%	85%
18	Chemistry	5,428	5,344	4,500	83%	84%
21	Environmental Sciences	2,376	2,214	1,845	78%	83%
22	Pure Mathematics	2,095	1,840	1,250	60%	68%
28	Civil Engineering	2,163	1,813	1,331	62%	73%
30	Mech, Aero & Man Engineering	4,353	4,039	3,271	75%	81%
35	Geography	4,928	3,863	2,774	56%	72%
42	Sociology	3,530	1,846	885	25%	48%
59	History	7,265	2,413	916	13%	38%

24. Table D2 shows the numbers staff associated with the outputs in Table D1.

Table D2: Proportion of staff with records matched

UOA code	Subject	Staff with RA2 published records	Staff with RA2 type D published records	Staff with at least one unique NCR Match	% Staff with at least one unique NCR Match
01	Clinical Laboratory Sciences	1,054	1,054	1,015	96%
02	Community Based Clinical Subjects	1,166	1,163	1,141	98%
13	Psychology	1,177	1,166	1,081	92%
14	Biological Sciences	2,327	2,323	2,233	96%
16	Food Science and Technology	115	113	107	93%
18	Chemistry	1,269	1,269	1,230	97%
21	Environmental Sciences	542	539	516	95%
22	Pure Mathematics	498	493	427	86%
28	Civil Engineering	516	504	450	87%
30	Mech, Aero & Man Engineering	1,023	1,016	971	95%
35	Geography	1,129	1,094	969	86%
42	Sociology	816	739	491	60%
59	History	1,668	1216	625	37%

Note that the total number of staff with at least one unique NCR match (excluding UOAs 42 and 59) in Table D2 was 10,140 ten more than the 10,130 figure in Table 13. This is because for 10 records data on the sex of the researcher was missing and was excluded.

Calculation of bibliometric measures

25. For each article, published in journal J, in year Y ascribed to field F, the following data were extracted:

a. Citations to the article.

b. Mean number of citations per article published in field F in year Y.

c. Mean number of citations per article published in journal J in year Y.

26. Simple citation rates are likely to be misleading since citation practice varies by field, and articles published earlier would be expected, all other things being equal, to have more citations. Year- specific field and article average citation counts are provided so that the citation count for the article in question can be compared to one of these two baselines.

27. For most of the RAE articles the relevant 'field' is the Thomson/ISI Current Contents category, which usually equates to sub-division within a UOA representing a reasonably well defined field of study, for example, Cardiovascular Sciences, Plant Science and Chemical Physics. Where the journal has not been mapped by Thomson to a specific category, which happens in a small number of cases, the average citation rate for the whole UOA for the relevant year of publication was used.

28. Using the journal average avoids the judgements as to what constitutes a 'field', but it is likely to flatter papers which are published in less prestigious journals, as these tend to have lower citation rates than more prestigious journals covering the same field.

Bibliometric measure investigated

29. The three measures were calculated as follows:

a. Average citation count relative to a field baseline

This is calculated by firstly calculating the field baseline, given by: number of citations to articles in journals in the field in stated year divided by number of papers published in journals in the field in stated year. Then each paper is given a citation count relative to this field baseline, given by: the citation count for the article divided by the field average citation count. The average of this measure for the papers submitted by an individual produces the average citation count.

b. Best citation count relative to a field baseline

The same calculation as ‘a’ but rather than taking an average of the paper citation counts for each individual, the best paper citation count is taken.

c. Proportion of papers over field baseline

For each paper an individual has submitted, the citation count of the paper is compared to the field baseline (as in ‘a’). The proportion of an individual’s papers that exceed that field baseline is used as the measure.

Annex E
Bibliometrics – linking to HESA data

1. The bibliometric data is based on data published on the HERO web-site. These data do not include personal details of the submitted researchers save for their names. The staff identifiers used within the published data differ from those used within the actual submissions as those used in submissions may have meaning outside the actual submission. Therefore, it was necessary to link back the bibliographic records to the original RAE submissions by institution, unit of assessment, multiple submission, forename and surname. This combination is not unique in all cases and some records were lost at this stage.

2. There is no direct link between the 2001 RAE database and the HESA staff record. Therefore in order to add demographic details (such as grade) to the bibliometric data it was necessary to link the two data sources using fuzzy matching techniques. The linking was done in seven passes, with the linking at each pass using only those records not matched in a previous pass. To test the reliability of the matching at each stage we repeated the process using a set of false records obtained by aging all staff born between June and December by one year and reducing the age by one year for those staff born between January and May. We would not expect any staff in the false record set to match the bibliometric data, thus any matches we do find provide an estimate of the error rate at each stage.

3. The first pass of matching assumed that institutions had returned the HESA staff identifier as the identifier on the RAE in practice this resolved the matching fully for two HEIs.

4. We then performed a second match by identifying groups of staff on the HESA record that were identical in respect of the following dimensions: Institution, UOA, year of birth, gender, grade, mode of employment, primary employment function, terms of employment, ethnicity and disability. Where for a combination of institution, UOA and year of birth only one of these groups existed and the number of staff matched those on the bibliographic data we randomly assigned one entry on the bibliographic database to each entry on the HESA record, as the precise assignment would have no impact on onward analysis.

5. The remaining five linking processes were all similar and required institution, UOA and gender to match in addition to the following variables at each stage:

   a. Multiple submission, year of birth, and date started at institution.
   b. Year of birth, and date started at institution.
   c. Multiple submission, year of birth, and principal source of salary.
   d. Multiple submission, year of birth, and mode of employment.
   e. Multiple submission.

The number of matches and the estimate of the number of false matches is given in Table E1.

Table E1: Number of matches and estimate of false matches

Type of match	Total matches	Estimate of false matches
Staff identifier	886	0
EO-neutral groups	594	84
Institution, UOA and gender		
Multiple submission, year of birth, and date started at institution	5,977	968
Year of birth, and date started at institution	213	98
Multiple submission, year of birth, and principal source of salary	525	125
Multiple submission, year of birth, and mode of employment	238	117
Multiple submission	24	24

Annex F
Bibliometrics – estimation of within-submission differences

1. To estimate within-submission differences for a particular measure, a simple multi-level model was used, as shown in Figure C1. A separate model was used for each unit of assessment.

$$\text{Measure}_{ij} \sim \text{N}(XB,\ \Omega)$$
$$\text{Measure}_{ij} = \beta_{0j}\text{Cons} + \beta_1 \text{Category}_{ij}$$
$$\beta_{0j} = \beta_0 + u_{0j}$$

$$\begin{bmatrix} u_{0j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u)\ :\ \Omega_u = \begin{bmatrix} \sigma^2_{u0} \end{bmatrix}$$

where subscript *i* represents an individual and *j* represents the institution.

- Measure_{ij} is one of the three bibliometric outcomes for individual i in institution j: average citation count relative to a field baseline; best citation count relative to a field baseline; or proportion of papers over field baseline.

2. Category_{ij} is whether individual i in institution j is either: male or female; or a professor or non-professor, depending on which of these categories is being examined. In both cases, Category_{ij} is a 0/1 dummy variable.

3. U_{oj} is a random effect for institution j.

Annex G
Bibliometrics – full results

Average citation count relative to a field baseline

Table G1: Citation count relative to a field baseline (average by UOA and RAE2001 rating)

	2001 RAE rating						
Unit of assessment	**2**	**3b**	**3a**	**4**	**5**	**5***	**Total**
Clinical Laboratory Sciences			1.8	1.8	2.4	3.1	**2.5**
Community-based Clinical Subjects			1.6	2.0	2.4	3.1	**2.3**
Psychology		1.0	1.0	1.1	1.6	2.5	**1.6**
Biological Sciences			1.4	2.2	2.4	2.6	**2.3**
Food Science and Technology							**1.4**
Chemistry			1.1	1.7	2.2	2.6	**2.0**
Environmental Sciences	1.1	1.2		2.0	1.8	2.5	**1.6**
Pure Mathematics					1.1		**1.2**
Civil Engineering				0.8	0.9	1.3	**1.0**
Mech, Aero & Man Engineering				1.0	1.0	1.2	**1.0**
Geography		0.9	1.2	1.8	1.8	2.3	**1.7**
Sociology			1.2	1.6	1.8	1.5	**1.4**
History				1.4	2.0	1.5	**1.7**

Table G1 notes:
Based on dataset 1 (plus Sociology and History UOAs) as listed in Table 13.

Best citation count relative to a field baseline

Table G2: Best citation count relative to a field baseline (average by UOA and RAE2001 rating)

	2001 RAE rating						
Unit of assessment	**2**	**3b**	**3a**	**4**	**5**	**5***	**Total**
Clinical Laboratory Sciences			3.6	3.9	4.6	6.0	**5.0**
Community-based Clinical Subjects			3.0	3.9	4.4	6.3	**4.4**
Psychology		1.6	1.5	1.9	3.0	4.2	**2.8**
Biological Sciences			2.6	4.2	4.6	4.7	**4.3**
Food Science and Technology							**3.1**
Chemistry			2.0	3.2	4.4	4.8	**3.9**
Environmental Sciences	1.9	2.2		3.6	3.4	4.5	**2.8**
Pure Mathematics					1.8		**1.9**
Civil Engineering				1.4	1.5	2.3	**1.7**
Mech, Aero & Man Engineering				1.9	1.9	2.3	**2.0**
Geography		1.4	1.9	2.8	2.9	3.6	**2.7**
Sociology			1.5	2.0	2.3	2.0	**1.9**
History				1.8	2.4	1.7	**2.0**

Table G2 notes:
Based on dataset 1 (plus Sociology and History UOAs) as listed in Table 13.

Table G3: Best citation count relative to a field baseline – Professors and other staff

Subject	Professor	Non-Professor	Overall difference	Within-submission difference
Clinical Laboratory Sciences	6.91	4.41	2.50	** 2.52
Community Based Clinical Subjects	5.35	4.15	1.20	** 1.55
Psychology	3.61	2.53	1.08	** 1.02
Biological Sciences	5.10	4.07	1.04	** 1.16
Food Science and Technology	6.36	1.90	4.45	* 4.41
Chemistry	4.88	3.66	1.22	** 1.23
Environmental Sciences	3.11	2.48	0.63	0.50
Pure Mathematics	2.41	1.68	0.73	** 0.62
Civil Engineering	1.93	1.63	0.30	0.22
Mech, Aero & Man Engineering	2.01	1.89	0.12	0.11
Geography	3.55	2.43	1.11	** 1.08
All 11 subjects	**4.30**	**3.19**	**n/a**	**** 1.17**

Table G3 notes

Based on dataset 2 as listed in Table 13.

* indicates significantly different from 1.00 at 5% level

** indicates significantly different from 1.00 at the 1% level

Table G4: Best citation count relative to a field baseline – Men and women

Subject	Male	Female	Overall	Within submissions
Clinical Laboratory Sciences	5.39	3.71	1.68	** 1.69
Community Based Clinical Subjects	4.79	3.78	1.00	* 0.94
Psychology	2.98	2.45	0.53	0.44
Biological Sciences	4.32	4.09	0.23	0.42
Food Science and Technology	3.46	2.01	1.45	1.52
Chemistry	3.78	4.52	-0.73	-0.70
Environmental Sciences	2.97	1.97	1.00	0.86
Pure Mathematics	1.96	1.40	0.56	0.54
Civil Engineering	1.70	1.54	0.16	0.07
Mech, Aero & Man Engineering	1.99	1.64	0.36	0.30
Geography	2.69	2.70	-0.02	-0.04
All 11 subjects	**3.50**	**3.32**	**n/a**	**** 0.56**

Table G4 notes:

Based on dataset 1 as listed in Table 13.

* indicates significantly different from 1.00 at 5% level

** indicates significantly different from 1.00 at the 1% level

Table G5: Best citation count relative to a field baseline – Men and women (Staff in the lower 50% of citation rates within submissions)

Subject	Male	Female	Overall difference	Within-submission difference
Clinical Laboratory Sciences	1.80	1.76	0.03	0.02
Community Based Clinical Subjects	1.62	1.50	0.12	0.07
Psychology	1.21	1.07	0.14	0.09
Biological Sciences	1.55	1.63	-0.08	-0.03
Food Science and Technology	1.24	1.22	0.02	-0.06
Chemistry	1.48	1.47	0.01	-0.05
Environmental Sciences	1.20	0.81	0.39	* 0.25
Pure Mathematics	0.86	0.85	0.01	0.09
Civil Engineering	0.64	0.60	0.04	-0.09
Mech, Aero & Man Engineering	0.81	0.75	0.06	0.07
Geography	1.03	0.88	0.16	0.12
All 11 subjects	**1.30**	**1.34**	**n/a**	**0.05**

Table G5 notes:

Based on dataset 1 as listed in Table 13.

* indicates significantly different from 1.00 at 5% level

** indicates significantly different from 1.00 at the 1% level

Table G6: Best citation count relative to a field baseline – Ethnicity (all 11 subjects)

Population	White	Non-white	Overall difference	Within-submission difference
All staff	3.42	2.97	n/a	0.11
Staff in the lower 50% of citation rates within submissions	1.30	1.11	n/a	0.01

Table G6 notes:

Based on dataset 2 as listed in Table 13.

* indicates significantly different from 1.00 at 5% level

** indicates significantly different from 1.00 at the 1% level

Proportion of papers over field baseline

Table G7: Proportion of papers over field baseline (average by UOA and RAE2001 rating)

	2001 RAE rating						
Unit of assessment	**2**	**3b**	**3a**	**4**	**5**	**5***	**Total**
Clinical Laboratory Sciences			51%	53%	61%	71%	**63%**
Community-based Clinical Subjects			48%	55%	63%	67%	**58%**
Psychology		32%	34%	38%	52%	61%	**47%**
Biological Sciences			42%	58%	61%	68%	**59%**
Food Science and Technology							**43%**
Chemistry			37%	51%	61%	67%	**57%**
Environmental Sciences	36%	39%		55%	53%	73%	**48%**
Pure Mathematics					41%		**42%**
Civil Engineering				26%	28%	43%	**32%**
Mechanical, Aeronautical and Manufacturing Engineering				31%	34%	39%	**34%**
Geography		33%	38%	48%	61%	57%	**50%**
Sociology			34%	44%	50%	45%	**42%**
History				47%	53%	49%	**50%**

Table G7 notes:

Based on dataset 1 (plus Sociology and History UOAs) as listed in Table 13.

Table G8: Proportion of papers over field baseline – Professors and other staff

Subject	**Professor**	**Non-professor**	**Overall difference**	**Within-submission difference**
Clinical Laboratory Sciences	68%	64%	5%	* 6%
Community Based Clinical Subjects	64%	56%	8%	** 9%
Psychology	53%	46%	6%	5%
Biological Sciences	64%	59%	6%	* 6%
Food Science and Technology	46%	42%	4%	4%
Chemistry	62%	55%	7%	** 8%
Environmental Sciences	55%	45%	10%	* 9%
Pure Mathematics	48%	39%	9%	9%
Civil Engineering	35%	32%	3%	4%
Mech, Aero & Man Engineering	33%	34%	-1%	-1%
Geography	58%	49%	9%	* 8%
All 11 subjects	**57%**	**51%**	**n/a**	**** 6%**

Table G8 notes

Based on dataset 2 as listed in Table 13.

* indicates significantly different from 1.00 at 5% level

** indicates significantly different from 1.00 at the 1% level

Table G9: Proportion of papers over field baseline – Men and women

Subject	Male	Female	Overall difference	Within-submission difference
Clinical Laboratory Sciences	65%	59%	6%	* 5%
Community Based Clinical Subjects	59%	57%	1%	1%
Psychology	49%	46%	3%	1%
Biological Sciences	60%	58%	2%	3%
Food Science and Technology	44%	40%	3%	3%
Chemistry	56%	58%	-2%	-3%
Environmental Sciences	50%	38%	12%	* 10%
Pure Mathematics	42%	36%	6%	6%
Civil Engineering	32%	31%	2%	0%
Mech, Aero & Man Engineering	34%	29%	6%	6%
Geography	50%	51%	-1%	-2%
All 11 subjects	**52%**	**53%**	**n/a**	*** 2%**

Table G9 notes:

Based on dataset 1 as listed in Table 13.

* indicates significantly different from 1.00 at 5% level

** indicates significantly different from 1.00 at the 1% level

Table G10: Differences in citation rates relative to field between men and women (Staff in the lower 50% of citation rates within submissions)

Subject	Male	Female	Overall difference	Within-submission difference
Clinical Laboratory Sciences	50%	44%	6%	5%
Community Based Clinical Subjects	39%	35%	3%	3%
Psychology	27%	25%	1%	-1%
Biological Sciences	37%	39%	-2%	0%
Food Science and Technology	28%	27%	0%	0%
Chemistry	34%	42%	-8%	** -9%
Environmental Sciences	25%	18%	8%	3%
Pure Mathematics	18%	11%	7%	8%
Civil Engineering	8%	6%	2%	-1%
Mech, Aero & Man Engineering	12%	10%	3%	3%
Geography	25%	20%	5%	4%
All 11 subjects	**30%**	**32%**	**n/a**	**1%**

Table G10 notes:

Based on dataset 1 as listed in Table 13.

* indicates significantly different from 1.00 at 5% level

** indicates significantly different from 1.00 at the 1% level

Table G11: Ethnicity – all 11 subjects

Population	White	Non-white	Overall difference	Within-submission difference
All staff	52%	46%	n/a	0%
Staff in the lower 50% of citation rates within submissions	30%	23%	n/a	1%

Table G11 notes:

Based on dataset 2 as listed in Table 13.

* indicates significantly different from 1.00 at 5% level

** indicates significantly different from 1.00 at the 1% level