

# The Consistency of RAE Outcomes in the Social Sciences: A Quantitative Comparison of Social Policy and Cognate Subjects in the RAE 2001

**Stephen McKay**

Personal Finance Research Centre, University of Bristol  
E-mail: s.mckay@bristol.ac.uk

---

*In the RAE 2001, widely differing proportions of staff were judged internationally excellent in different academic subjects. This has important financial and other implications, and in 2008 a hierarchical structure will aim to increase consistency. We look at those subjects that will be part of a super-unit with social policy in 2008, and consider what objective rationale may explain the wide differences in ratings in 2001. A quantitative analysis found no such rationale, indeed some variations appear perverse. This highlights the need for greater consistency of judgements, perhaps steered by a 'super-panel'.*

## Introduction

This paper analyses data from the 2001 Research Assessment Exercise (RAE) to compare the rating of subjects that in 2008 will form the Panel that includes social policy. As may be well-known, there were large differences in the proportions of staff and submissions rated 'internationally excellent' among this group.

### *Background*

University funding in the UK is partly based on an assessment of research quality, the RAE. Taking place in various years, generally 4–5 years apart, the RAEs have been used to allocate money to institutions, selectively, based on their research quality. Better research receives more money, with a gradient intended to concentrate funding. The most recent exercise, in 2001, was responsible for the allocation of some £5 billion of funding. In addition to the financial implications, the RAE reinforces differences in the reputations of different departments. 'Good' ratings are used by departments to try to improve recruitment of both staff and students. Those with higher ratings will enjoy higher funding, which may be used to attract the best staff – a virtuous spiral for those well-funded. The RAE data are also used in the various 'league tables' that are produced to rank universities, both for particular subjects and across the board.

As many readers will be aware, funding formulas often factor in RAE scores. This may mean that budgets in lower-rated departments come under more pressure than merely the quality-related (QR) funding directly linked to the RAE. Not only may QR money be allocated selectively, but in some institutions there are multipliers that do this still further. The provision of funding selectively risks putting some departments into a vicious circle of having to find funds to pay the bills in ways that reduce their time for dedicated

pure research (such as applying for research grants and larger student bodies, including overseas students). This may therefore make it more difficult to find the time to produce the kinds of 'RAE-able' output needed to improve that rating. Alternatively departments may be closed.

The 2001 variant of the research rating used a seven-point scale with the points (1, 2, 3b, 3a, 4, 5, 5\*), shown here in ascending order. Among the higher (i.e. more desirable) points on the scale the descriptions were:

- 5\* Quality that equates to attainable levels of international excellence in more than half of the research activity submitted and to attainable levels of national excellence in the remainder.
- 5... international excellence in up to half of the research... national excellence in virtually all the remainder.
- 4... national excellence in virtually all of the research, showing some evidence of international excellence.
- 3a... national excellence in over two-thirds of the research... possibly showing some evidence of international excellence

Work regarded as being of international excellence may therefore appear in any four of the seven available rating points. Even so, only the top two are generally regarded as relating to 'international quality'.

The last set of results is from the 2001 Research Assessment Exercise. This saw considerable increases in assessed research quality over its predecessor in 1996. The highest ratings (5/5\* or internationally excellent) were awarded for Departments constituting 31 per cent of staff in 1996, but 55 per cent in 2001. More importantly for the analysis here, the 2001 RAE also saw very wide disparities across disciplines in the proportion of submissions ranked as 'internationally excellent'. For instance, 70 per cent of submissions in anthropology but only 21 per cent in social policy reached this level. Sharp (2004) looked across the last three such exercises, and at differences across subjects, and commented that: 'Panels seem to be developing different 'cultures' in respect of the proportion of 5 and 5\* ratings awarded... This is a matter which needs to be addressed if the scale is to have credibility as a measure of research quality' (Sharp, 2004: 216). Given the importance attached to ratings within universities, it would be of concern if the ratings came to reflect subject-specific custom or an emerging tradition rather than quality.

#### *The 2008 RAE*

The next RAE (2008) will have a hierarchical form, with 15 main panels covering a range of subjects and each overseeing 3–8 sub-panels devoted to single subjects. The sub-panels correspond to the panels of 2001. This reform is designed to increase transparency and consistency: it is part of the role of main panels 'to ensure consistent application across the exercise of the overall quality standards' (HEFCE, 2004: paragraph 28). The issue of consistency across panels was raised in the Roberts review of research funding (Roberts, 2003). Even so, sub-panels are expected to establish their own interpretations of levels of excellence.

We do not investigate patterns across all of higher education, we focus on the groups forming the main panel with social policy. Social policy finds itself located in 'Main Panel J', which contains sub-panels devoted to the following subjects which had their

own panels in 2001:

- anthropology;
- law;
- politics and international studies;
- social policy and administration;
- social work, and;
- sociology.

It will also include development studies, a sub-panel of geography in 2001. Given the distinct status of development studies in 2001 it is not included in the remainder of this analysis.

Since all the six subjects have a track record from 2001 (and indeed earlier research assessment exercises), it is possible to compare the kinds of decisions made against the publicly available data on submissions (HERO, 2001). In 2001, the RAE received 2,598 submissions in total. For the six subjects above there were 274 separate submissions, so this group represents a little over 10 per cent of the 2001 RAE process.

#### *Social policy in the RAE*

At a time when social policy appears to be beset by general pessimism, the low proportions ranked as internationally excellent in 2001 did little to bolster morale. Subsequent policy decisions by universities and certainly their approach to recruitment, do appear to be driven at least partly by RAE outcomes. With the funding levels attached to merely national research (grade '4') having been cut, the award of international status (grade '5' or '5\*') provides at least a degree of security and stability. The credibility of the process relies on a degree of consistency across subjects in the allocation of rankings.

There has been lively discussion of the results for social policy and administration in the 2001 RAE. An external assessor has criticised some of the methods of working (Gambrill, 2002, 2003), with vigorous defence from the then panel chair (MacGregor, 2003). Hillyard (2002) directly compared the RAE outcomes for social policy and administration and law, in the context of socio-legal studies. He noted some of the effects of significant differences in subject ratings. McKay (2003) found that the outcomes of the panel for social policy were predictable on the basis of quantitative data, albeit with new universities appearing to lose out. Craig (2002) documented some of the changing submission strategies and outcomes within social policy and administration. Alcock (2004) provides further discussion of these points.

The Social Policy Association has argued that the concentration of research funding does not have the same justification in social science funding that it may have in the expensive 'hard' sciences (Alcock, 2004). In some sciences the high cost of capital infrastructure may argue for concentration. In the social sciences often little equipment is needed, and critical mass must be considered alongside the need for diversity.

#### *Data*

The submissions universities made to the 2001 RAE are available, either from the funding councils or on-line. Any interested party may undertake analysis of many different kinds. The interested analyst may check on numbers of Ph.D. completions in various

Table 1 Distribution of ratings by subject, submissions basis (Column percentages)

Rating	Law	Anthropology	Politics and International Studies	SPA	Social Work	Sociology	Total
2	1.7	–	1.5	4.3	3.3	8.3	3.3
3b	3.3	–	2.9	8.5	<b>26.7</b>	8.3	7.3
3a	16.7	15.0	<b>34.8</b>	31.9	<b>26.7</b>	<b>25.0</b>	26.3
4	15.0	15.0	26.1	<b>34.0</b>	16.7	20.8	22.3
5	<b>50.0</b>	<b>60.0</b>	27.5	17.0	23.3	<b>25.0</b>	<b>32.1</b>
5*	13.3	10.0	7.3	4.3	3.3	12.5	8.8
5/5*	63.3	70	34.8	21.3	26.7	37.5	40.9
Number of ratings	60	20	69	47	30	48	274

Note: Mode(s) in bold.

organisations, or simply enquire what various people believed to be their ‘best’ publications (where results are often quite surprising).

Data are available either in the Microsoft access format, or as ‘comma-separated ASCII files’. The database format makes it possible to efficiently store all the various information, with the data tables appropriately set up in a relational form, but it is not well suited to any kinds of statistical analysis. Therefore even quite cursory investigations must be postponed until the data have been converted to a format suitable for statistical analysis, such as in SPSS, Stata, SAS and the like.<sup>1</sup> Analysts must also amalgamate data that are stored separately for each panel. At least one organisation will supply a range of custom reports using these data – for a fee.

The article now compares the RAE outcomes across the six social sciences subjects and attempts the adjudicated outcomes to the information included in the submissions.

### An overview of the 2001 RAE outcomes

The distribution of RAE ratings differed substantially between the different units of assessment (‘subjects’). Among this group of subjects around four submissions in every ten were classified as ‘international’ standard (5 or 5\*). This was much higher in anthropology (seven in ten) and considerably lower in social policy and administration (SPA, two in ten), as shown in Table 1. The table shows in boldface type the predominant ratings (the mode) for each subject. Clearly in law and anthropology the most common award, half or more of submissions, was a ‘5’. In the other subjects the modal department tended to be somewhere in the 3a/4 range.

Departments vary greatly in size, and often larger departments are rated as higher quality. As a result, a higher proportion of staff work in higher rated submissions than the figures in Table 1 might suggest, so Table 2 shows the proportion of (submitted) staff within different categories. What is particularly stark is that for all subjects, *apart from SPA*, the modal (most common) category was ‘5’. In social policy the largest single group of staff (37 per cent) worked in units rated ‘4’. Moreover, only in SPA (and social work which was reviewed by the same panel) did a minority rather than a majority of staff submit work in units rated of international quality.

Table 2 Distribution of staff working in different rated departments by subject (Column percentages)

Rating	Law	Anthropology	Politics and International Studies	SPA	Social Work	Sociology	Total
2	0.2	–	1.0	2.2	0.7	4.1	1.5
3b	0.3	–	0.8	5.0	19.1	5.2	3.6
3a	5.2	9.8	19.3	25.1	29.1	18.7	16.8
4	10.3	8.0	25.5	<b>37.2</b>	12.9	21.1	21.0
5	<b>58.7</b>	<b>67.0</b>	<b>38.2</b>	22.9	<b>32.8</b>	<b>29.0</b>	<b>40.4</b>
5*	25.3	15.1	15.2	7.6	5.3	22.0	16.8
5/5*	83.9	82.1	53.4	30.5	38.2	50.9	57.2 <sup>a</sup>
Number of staff	1,198	271	1,022	880	345	774	4,493

Notes: (a) By comparison, across the whole of RAE 2001 55% of staff were in units rated 5 or 5\*. Based on Category A staff, FTE. Mode(s) in bold.

As mentioned earlier, the social policy profession has set its stall against concentrating research money. Instead it has argued that good research exists in many departments, and that the costs of conducting research do not argue for a high degree of selectivity in funding. Nevertheless, if this were to be interpreted as a strategy it would seem that law and anthropology were the ones following that kind of approach. In fact the concentration of research money into relatively few departments is stronger in social policy than the other disciplines considered here. The ratings awarded to these subjects did not, in 2001, equate to identical levels of funding. The amounts available for each subject were different, and so the financial consequences of each rating could also be quite different.

### Potential causes of the differences

The issue to address is whether the higher results in law and anthropology are because research in those subjects really is better than in the others, or because a similar (or lower) standard of research was judged differently and indeed more generously. It is certainly possible that some subjects conduct research of higher quality than others, when judged against the standard of international excellence. It is also possible that standards are similar across subjects. The RAE outcomes suggest there *is* a large disparity in research quality in 2001, and we investigate if such a difference may be detected on a consistent and logical basis using the data available. We begin by first checking what kinds of definitions of quality were used.

#### *Alternative definitions of quality?*

The working methods of each panel were set out in advance of the closing date for submissions – though still some way into the period over which work was published. We may investigate if certain difference of understanding might be affecting the results. Basse (2002) has demonstrated the different ways that panels sought to define international excellence. Moreover, in a thorough analysis of the RAE process, McNay (2003: 50)

has highlighted that: 'Panels differed in: the wording in relation to criteria for national excellence; the elements included in criteria; how they constructed a final grade; use of judgements on output; and the account taken of the balance of quality across the four outputs.'

Despite this, a quick observation shows a surprisingly high degree of similarity of how high quality was to be defined. Research may be regarded as of international excellence if it is the best in its field, and/or if it may be regarded as a primary point of reference in its field.

<b>Unit of assessment</b>	<b>High quality</b>
Law	'Work can be regarded as of international excellence if it is a primary reference point in its field (in the sense that it is, or in the opinion of the Panel is likely to be, recognised as amongst the best in its field)'
Social Policy & Administration/ Social Work	'Work will be regarded as of international excellence if it is, or is likely to become, a primary reference point in its field (in the sense that it is amongst the best in its field)'
Politics and International Studies	'International excellence equates to work of highest quality. It denotes work with which any researcher in the field or sub-field ought to engage.'
Anthropology	'International excellence in this context . . . [is that which] has become, or is likely to become, a primary point of reference in its sub-field. Such research is at a level that is at, or very close to, the highest level of excellence and impact that is achievable in this sub-field'

These judgements and definitions are far from ambiguous, of course, and are open to question. Some 'points of reference' in social science disciplines are exemplars of what others believe to be mistaken (some analysts would consider this to be true of work on the underclass, or Bowlby's research on the maternal deprivation hypothesis). Second, it is arguable that being a primary reference point is a fairly objective criterion (perhaps amenable to citation analysis), whilst being 'best in field' implies a judgement *and* that other work is clearly not the best in its field (a relative judgement in other words). Time also has a role; it is difficult to argue that a recently published piece is a primary reference point (hence the above emphasis on potential which must be exceedingly difficult to predict).

#### *Quantitative data*

In this section we look at a variety of quality indicators, to consider how far the differences in ratings reflected clear differences in the data supplied with submissions. These include data on staff numbers, publications, students and external fund-raising. The list is by no means exhaustive, but represents information that was consistently supplied as part of the RAE process.

*Size of submissions.* Past research studies (McKay, 2003; Taylor, 1995) have found a strong positive correlation between the size of a submission and its rating – bigger units do better. This may be cause or effect; successful departments may grow more quickly, but critical mass may make for successful departments. This link was certainly found

Table 3 Distribution of staff by subject and ratings (Median number of staff submitted (FTE))

Rating	2/3b/3a	4	5/5*	All
Law	5	14	26	17
SPA	13	17	25	16
Sociology	10	14	21	14
Anthropology	9	7	15	14
Politics and International Studies	8	15	17	13
Social Work	10	9	14	11
Total (these subjects)	8	14	21	14

Table 4 Proportion of staff submitted (to total staff) by subject and ratings (Mean ratio of selected to total staff)

Rating	2/3b/3a	4	5/5*	All
Law	28	56	86	69%
Social Work	57	80	90	69%
SPA	78	80	92	81%
Politics and International Studies	73	86	93	84%
Sociology	71	89	96	84%
Anthropology	98	100	93	95%
Total (these subjects)	66%	81%	91%	79%

among all the units of assessment studied here. Departments rated of international quality were always somewhat larger or much larger than other departments within the same subject (see Table 3). In SPA, somewhat unusually, some of the lower-rated departments were quite large. But there is no evidence that any differences in department sizes in different subjects explains the substantial ratings differences. The SPA panel seemed the most prepared to give lower ratings to relatively large departments.

However, law and social work seemed the most selective in their decisions about submitting staff. Some of the lower-rated submissions were highly selective. In most cases the top-rated departments were not (or did not need to be) quite so selective. Again this is not really an explanation of the differences in ratings – since anthropology was the least selective overall (see Table 4).

*External research funding.* A simple measure of research income was created, adding up all sources of external funding over the relevant RAE period (financial years 1996–2000), without any uprating. The average amounts raised within each department are shown in Table 5, the upper half relating to departments, and the lower half expressing figures on a simple per capita basis (by number of selected staff). In terms of money raising across whole subjects, social policy leads the field by some margin. Expressed per member of staff, social work is nearly as successful at raising funds.

Success at external fund-raising was quite closely correlated with the final rating, at least when looking *within* each of these subject disciplines. But even the most lowly rated social policy departments raised external income in excess of top-rated law

Table 5 Total external research income (1995/6–1999/00) by subject and ratings (£ median by submission, and per capita)

Rating	2/3b/3a	4	5/5*	All
<i>By unit</i>				
SPA	£467,148	£1,190,757	£3,340,139	£1,013,660
Social Work	£284,921	£1,455,802	£1,817,146	£822,010
Sociology	£184,038	£1,048,763	£1,859,756	£733,908
Anthropology	£585,863	£70,142	£612,799	£567,904
Politics and International Studies	£38,000	£421,127	£580,672	£281,310
Law	£23,000	£128,262	£360,724	£162,494
Total (six subjects)	£175,586	£674,551	£745,319	£451,185
<i>Per capita</i>				
SPA	£44,800	£82,300	£135,000	£72,300
Social Work	£29,200	£141,200	£109,300	£67,600
Sociology	£22,800	£56,800	£90,200	£50,600
Anthropology	£47,000	£10,300	£37,100	£30,100
Politics and International Studies	£4,500	£30,000	£30,300	£19,422
Law	£5,800	£7,400	£15,800	£10,000
Total (six subjects)	£19,200	£51,200	£39,400	£32,000

departments – a median over £3.3 million for the 5/5\* social policy submissions, or nearly four times as much as the average across all these subjects. With no apparent trace of irony the Law Panel later told us that within their subject: ‘the overall level of research income was impressive’ (Law Panel overview report, p. 2)

Interestingly, the results are fairly similar whether one looks at total amounts raised or on a per capita basis.

In Annex 1, a separate table shows funding derived only from the research councils (or that of the then AHRB). This kind of funding is sometimes seen as indicative of higher academic quality than funding of other kinds – the process is peer-reviewed by academics and many proposals are rejected. The main change to the figures is that Social Work appears much lower funded than before, reflecting a more diverse funding base than the other disciplines. Overall there was a strong correlation between total funding and this kind of funding ( $r = 0.7$ ). Those with high overall funding tend to have high research council funding.

*Post-graduate students and research degrees.* In Table 6 we show the total number of doctorates awarded during the recording period. Anthropology departments awarded the most, across the board, but within 5/5\* submissions SPA and Politics awarded more than the others. Social Work departments had the fewest completions of this kind, reflecting a different emphasis in this subject.

Unsurprisingly, the total numbers of research students were positively correlated with doctoral completions, subject to time lags. Social Work again had the fewest research students, with Law again somewhat behind the other social sciences, and all trailing Anthropology. The 5/5\* departments working hardest at postgraduate training appeared to be in SPA, Sociology and Politics and International Studies (See Table 7).



Table 6 Number of doctorates, awarded by subject and ratings (Median number of doctorates awarded in department)

Rating	2/3b/3a	4	5/5*	All
Anthropology	7	7	21	15
SPA	5	11	32	10
Politics and International Studies	3	11	33	10
Law	0	3	14	9
Sociology	5	9	27	8
Social Work	3	2	7	3
Total (these subjects)	3	8	20	8

Table 7 Numbers of research students (FTE) on courses by subject and ratings (Median number of doctorates awarded)

Rating	2/3b/3a	4	5/5*	All
Anthropology	31	29	111	92
SPA	45	79	152	68
Sociology	29	62	154	67
Politics and International Studies	18	72	153	63
Law	13	35	96	53
Social Work	28	26	61	32
Total (these subjects)	28	63	114	62

Table 8 Distribution of publication types by subject (Column percentages)

Rating	Law	Anthropology	Politics and International Studies	SPA	Social Work	Sociology	Total
Authored book	17	14	20	15	14	16	17
Edited book	1	6	4	4	3	4	3
Chapter	27	32	23	23	22	24	25
Article	52	44	51	49	54	51	51
Software	2	1	1	5	5	2	2
Others	1	3	1	4	2	3	2
Number of pubs	5,328	1,200	4,410	3,952	1,650	3,530	20,070

Source: Taken from RA2 data, RAE 2001.

*Publications.* One of the most difficult – but crucial – tasks of comparison relates to the treatment of publications. Clearly, one would expect ratings to be closely identified with the quality of publications submitted, and panels generally made this clear in their guidance. Under the new rules proposed for RAE-2008, *at least* 50 per cent of the overall profile will be based on the publications profile, and this could be as high as 90 per cent if panels so choose. So, are differences in the 2001 ratings to be found in different modes of publication? The results shown in Table 8 do not provide support for any strong differences. In Anthropology there were somewhat fewer articles and more book chapters

Table 9 The representation of 'high quality' journals within all journal articles submitted, by subject (Cell percentages)

Rating	2/3b/3a	4	5/5*	All	Ratio of intern'l:rest
Anthropology	8.1	14.3	14.1	13.6	1.32
Sociology	5.6	9.1	10.1	8.7	1.42
Social Work	9.9	11.9	16.8	12.7	1.62
Law	2.9	9.9	11.3	10.4	1.65
SPA	3.1	6.5	9.4	6.4	1.88
Politics and International Studies	3.6	6.8	11.1	8.5	2.07

Note: 'High quality' journals were those three journals appearing most often among 5/5\* submissions in that subject.

Source: Taken from RA2 data, RAE 2001.

Table 10 Rankings correlated across the six subjects

Characteristics of submissions	Rank correlation coefficients with the percentage of staff rated internationally excellent within the subject	
	Kendall's tau	Spearman's rho
Size	+0.28	+0.32
Selectivity	-0.14	-0.06
Money raised	-0.87*	-0.94**
Doctorates awarded	-0.20	-0.20
No. of students	-0.20	-0.20
Selectivity of journals cited	-0.20	-0.14

(and, logically enough, more edited books) than was typical for this group; in Law the profile was very close to that represented overall (two percentage points fewer edited books, the reverse for chapters). What differences there are must be regarded as very small; perhaps with Anthropology the least like the other subjects, owing to the relative importance of book chapters compared with articles. This is likely to be related to the availability of relevant journals. And it is to journal publications that we now turn.

A simple means of comparing within subjects is to identify those journals that the high rated submissions were themselves more likely to use. We identified for each subject those *three* journals which were the most commonly included in the submissions rated 5 or 5\* (only these, and not the others). These may be taken, for our purposes, as providing some indication that these outlets are regarded as good quality. Of course this measure is based on internal consistency, and is not an external objective yardstick.<sup>2</sup> We may then compare how often these journals were part of the 'best' submissions compared with the others, and this is done in Table 9.

The three journals cited most often in the international-quality submissions accounted for anywhere between 6 per cent and 14 per cent of all articles, and from 9 per cent to 17 per cent of articles submitted by the international quality departments. Clearly, the concentration of journals is not becoming particularly strong overall.

As would be expected, the journals included most often in the 5/5\* submissions were less frequently cited among lower-rated submissions lowly – but by how much? In fact, the journals cited by the 5/5\* submissions were 32 per cent more frequent for Anthropology and up to 107 per cent more common within Politics and International Studies than for lower-rated submissions. This suggests, that even the most lowly rated departments were nonetheless publishing in the same outlets as the ‘best’. The extent to which this happened varied, but did not appear to be related to how many departments or staff were regarded as the best.

*Putting it all together.* The subjects had a clear ranking with regard to the proportion of staff rated as internationally excellent. It is also possible to use the analyses above to rank the subjects on a range of criteria related to funding, staff numbers, students and publications. We may then investigate how far those ranks correspond with the overall quality ranking. And, unsurprisingly by now, there was little correlation (see Table 10). The most helpful piece of information is the funding profile. There is a very strong *inverse* correlation between subjects raising the most external funds, and subjects most likely to rate their staff as internationally excellent. Compare McNay’s view (2003: 51) that ‘professional areas fare worse than cognate academic areas’. That aside, having larger units was weakly correlated with a higher rate of success. None of the other information had much of a systematic relationship with overall quality, taken across the six subjects.

## Conclusions

The 2001 RAE ratings for the six subjects analysed here were distributed very unevenly. In some the badge of international excellence was awarded to over *eight* staff in every ten; in others to barely *three* in ten. Previous research has tended to find that, *within* particular subjects, systematic differences occur between different ratings in a predictable manner. Our aim has been to investigate whether a solid quantitative basis may be found for these differences that occur *across* subjects.

The analysis can provide no objective, quantitative justification for these differences, nor any differences in the explicit criteria used for rating. If it is true that differences in research quality between subjects determine these results, it has not been possible to identify what such differences are. These results affirm that improving the consistency of judgements will be a key task in 2008.

Overall, it is difficult to find any straightforward statistical explanation for the differences. Information that has proved useful in explaining differences within a subject do not appear to be quite so powerful when looking across subjects. Indeed some input measures, such as the amount of external research funding, were negatively related to a subject’s own opinions of its quality *vis-à-vis* other subjects.

It is, of course, possible to argue that analysis of this kind can only go so far, and cannot accurately represent the deliberations of the panels. However, with such large differences in outcomes one might have suspected considerable differences to be evident. They are not. The next RAE, due in 2008, is being established on a new basis, with a hierarchical structure. One aim is to ensure consistency of judgements. For the small group of subjects analysed here, it seems that greater attention to consistency of ratings is overdue, and the new structure should help to enhance the credibility of the process.

## Acknowledgements

I would like to thank the higher education funding institutions, covering England (HEFCE), Northern Ireland (DEL), Scotland (SHEFC) and Wales (HEFCW), for making the RAE data both publicly available and readily accessible. I also thank Alan Deacon (Leeds), Martin Powell (Bath), Karen Rowlingson (Bath) and George Smith (Oxford) who read the draft and made many helpful suggestions that improved the analysis. Referees also provided helpful comments. However, any mistakes left are mine.

## Notes

1 The author downloaded the relevant data in MS-Access format, and converted the required data tables into Stata format. All analysis was then conducted by the author using Stata v8.2.

2 There are academic disciplines which have tried, or are trying, to develop guides to which journals 'count' as four-star, three-star, etc. One example is Business and Management Studies, and one list here (<http://www.uwe.ac.uk/bbs/research/journalph.pdf>) draws together the different views of six different UK universities in this field. The list is edited by Charles Harvey and Huw Morris of Bristol Business School.

## References

- Adams, J.** (2000), 'The role of selectivity and the characteristics of excellence', Report to the Higher Education Funding Council for England, a consultancy study within the Fundamental Review of Research Policy and Funding Leeds, Evidence Ltd (available at <http://www.evidenceuk.com/downloads/selectivity-report.pdf>, accessed 10.2.2005).
- Alcock, P.** (2004), 'Social policy and the next research assessment exercise', *Social Policy News*, March, pp. 1–4.
- Bassey, M.** (2002), 'Emperor's new clothes now fill a cupboard', *Research Intelligence*, 78: 12–14 (included as an annex to <http://www.ra-review.ac.uk/invite/responses/subj/20.pdf>, accessed 10.3.2005).
- Craig, G.** (2002), 'What does it all mean?', Social policy and the 2002 Research Assessment Exercise Memorandum submitted to the Select Committee on Science and Technology.
- Gambrill, E.** (2002), '“I am not a rubber stamp”: my experience as a non-UK RAE advisor', *Journal of Social Work*, 2, 2, 169–185.
- Gambrill, E.** (2003), 'Reply to Susanne MacGregor's response, "We Did Not Simply Mimic Received Opinion"', *Journal of Social Work*, 3, 2, 253–256.
- HEFCE** (2004), 'Units of assessment and recruitment of panel members', July, Ref RAE 03/2004.
- Higher Education and Research Opportunities in the UK (HERO)** (2001), 'Research assessment exercise 2001: the results', available online at: [www.hero.ac.uk/rae/index.htm](http://www.hero.ac.uk/rae/index.htm)
- Hillyard, P.** (2002), 'Invoking indignation: reflections on future directions of socio-legal research', *Journal of Law and Society*, 29, 4, December, 645–656.
- Law Panel** (2002), *RAE 2001: General overview*, available online at: [www.hero.ac.uk/rae](http://www.hero.ac.uk/rae)
- MacGregor, S.** (2003), '“We did not simply mimic received opinion”: response from the chair of the social policy and administration and social work RAE panel to Eileen Gambrill's "I am not a rubber stamp"', *Journal of Social Work*, 3, 1, 119–124.
- McKay, S.** (2003), 'Quantifying quality: can quantitative data ("metrics") explain the 2001 RAE ratings for social policy and administration?', *Social Policy and Administration*, 37, 5 (October), 444–467.
- McNay, I.** (2003), 'Assessing the assessment: an analysis of the UK Research Assessment Exercise, (2001), and its outcomes, with special reference to research in education', *Science and Public Policy*, 30, 1, 47–54.
- Roberts, G.** (2003), 'Review of research assessment', report by Sir Gareth Roberts to the UK funding bodies.
- Sharp, S.** (2004), 'The Research Assessment Exercises 1992–2001: patterns across time and subjects', *Studies in Higher Education*, 29, April, 201–218.
- Taylor, J.** (1995), 'A statistical analysis of the 1992 research assessment exercise', *Journal of the Royal Statistical Society, Series A*, 158, 2, 261–261.

**Annex**

Table A1 Total external research income from research councils and AHRB (1995/6–1999/00) by subject and ratings £median by submission

Rating	2/3b/3a	4	5/5*	All
<i>By unit</i>				
Sociology	£17,195	£235,666	£708,323	£192,684
SPA	£10,815	£196,735	£608,451	£131,580
Anthropology	£37,017	£49,459	£203,642	£101,506
Politics and International Studies	£4,300	£54,969	£265,934	£38,074
Social Work	£0	£20,204	£113,606	£13,630
Law	£0	£16,399	£44,401	£10,813
Total (six subjects)	£4,616	£80,141	£171,316	£51,433