

Comparing alternatives to the *Web of Science* for coverage of the social sciences' literature

Michael Norris, Charles Oppenheim*

Department of Information Science, Loughborough University, Loughborough, Leicestershire LE11 3TU, UK

Received 17 October 2006; received in revised form 5 December 2006; accepted 7 December 2006

Abstract

The *Web of Science* is no longer the only database which offers citation indexing of the social sciences. *Scopus*, *CSA Illumina* and *Google Scholar* are new entrants in this market. The holdings and citation records of these four databases were assessed against two sets of data one drawn from the 2001 Research Assessment Exercise and the other from the *International bibliography of the Social Sciences*. Initially, *CSA Illumina's* coverage at journal title level appeared to be the most comprehensive. But when recall and average citation count was tested at article level and rankings extrapolated by submission frequency to individual journal titles, *Scopus* was ranked first. When issues of functionality, the quality of record processing and depth of coverage are taken into account, *Scopus* and *Web of Science* have a significant advantage over the other two databases. From this analysis, *Scopus* offers the best coverage from amongst these databases and could be used as an alternative to the *Web of Science* as a tool to evaluate the research impact in the social sciences.

© 2007 Charles Oppenheim. Published by Elsevier Ltd. All rights reserved.

Keywords: RAE; *Web of Science*; *Scopus*; *CSA Illumina* *Google Scholar* research impact

1. Introduction

Since 1986, university research funding decisions in the UK have been based on a series of Research Assessment Exercises (RAE). These have been concerned with assessing the quality of academic research by the peer review of the scholarly output from university departments. Government funding decisions have then been taken based on the ranking of these departments after the review process. This process of assessing academic research by peer review is unlikely to be used after the forthcoming 2008 RAE. It is expected that after this, assessment will be based, in part at least on other means, notably by the use of some form of metrics, possibly including bibliometric measures of research impact. Using such measures relies heavily on having adequate bibliometric records of the subject being assessed. Up until now, this has generally been achieved by using the citation indexes provided by the *Web of Science*. In the case of the social sciences, despite some criticisms of its coverage of the field, the *Web of Science*, until fairly recently was the only credible database which had coverage of the subject and provided citation indexing, thus allowing the possibility of making some measurement of a department's impact in a subject. Recently, however, a number of other providers have entered this market and also offer such a service (Roth, 2005, pp. 1531–1536). *Scopus*, a multidisciplinary database,

* Corresponding author. Tel.: +44 1509 223065; fax: +44 1509 223053.

E-mail addresses: M.Norris2@lboro.ac.uk (M. Norris), C.Oppenheim@lboro.ac.uk (C. Oppenheim).

was launched by Reed Elsevier in 2004 with citation indexing. Likewise, *CSA Illumina* has added this feature to some of its databases and *Google* has added *Google Scholar* to its family of free services as a database of electronic scholarly sources.

It is generally recognised, however, that a good proportion of the scholarly output in the social sciences other than those subjects related to medicine and health are less well covered than the natural sciences in the *Web of Science* (Moed, 2005, pp. 125–126). This lack of coverage is usually attributed to the publication patterns in the social sciences, which to a certain extent favours monographs (Hicks, 2004). Hicks (2004, p. 477) notes that 85% of natural scientists' output is found in journal and conference papers, whilst for social scientists this can range between 42 and 61%. Added to this, there is evidence that many citations in the social sciences are to books rather than journal articles (Hicks, 2004, pp. 480–484). Nederhof (2006, p. 83), in his review of research performance in the social sciences, confirms a regional or national orientation to the publications patterns for some of the fields in the social sciences rather than publishing research on an international basis. However, the *Social Science Citation Index* in the *Web of Science* indexes material that is predominantly in English (93–95%) with the remainder being shared between German (2–3%), French (1%) and other languages (Nederhof, 2006, p. 84). There does, however, appear to be a shift towards social scientists publishing more of their work in journals. Larivière et al. (2006, pp. 1002–1003) compared the referencing practices of those in the natural science and engineering fields with those in the humanities and the social sciences and concluded that in the social sciences, there appears to be a steady increase in the share of the number of citations to journals.

Assessing which database to use to measure the scholarly output and impact of the social sciences is an issue of growing importance, given that the *Web of Science* is now not the only choice and that the results of such measurements are likely to assume greater importance in the apportionment of research funding. Ideally, any database which offers coverage of the social sciences would incorporate to a greater degree the scholarly output found in monographs, reports, articles and articles appearing in non-English language journals.

Now that it is possible to assess the coverage of the social sciences by evaluating other databases it seems appropriate to make some assessment of their coverage and quality compared to the *Web of Science*. Usefully, the journal article submissions made to the last RAE in 2001 across 13 Units of Assessment (UoA), which broadly cover the social sciences in the UK, can be used as a benchmarking tool to assess the coverage of these databases (see Appendix A). In particular, using these article submissions, an assessment can be made of the depth of coverage of the:

- Journals in which UK social scientists publish.
- A number of European foreign language journals in the social sciences, notably German, French, Italian and Spanish.

Finding credible sources which could be used to benchmark the coverage of monographs or reports is difficult and hence these have been excluded from the analysis.

2. Database selection

2.1. *CSA Illumina*

CSA Illumina, formerly *Cambridge Scientific Abstracts*, is a collection of about a 100 bibliographic databases, some of which it hosts with other database partners. The social sciences have a distinct group of databases which includes, for example, *Social Services Abstracts*, *Sociological Abstracts*, and *PsycINFO*. *Sociological Abstracts* is available from 1952, and indexes about 1800 serials, a number of conference proceedings and books. A selection policy classifies journal titles and their contents into one of three categories—'core', 'priority' or 'selective'. Core journals are those journals which are key publications in the discipline and almost all articles are indexed, priority journals are related to the discipline and over 50% of the articles are indexed. Journals which are classed as selective have fewer than 50% of their articles are indexed. Within these databases citation indexing is currently only available in some of them and sometimes this is limited to only core journals. For example, all journal articles contained within the *Social Services Abstracts* have had their cited references indexed from 2004, whilst *Sociological Abstracts* has cited references from core journals only.

2.2. Google Scholar

Google Scholar, launched in its beta form in 2004, provides at no cost multidisciplinary access to scholarly information. It is not clear from which sources *Google Scholar* has built its database, or how large it is, but it is evident that a number of publishers have allowed their electronic journal records to be indexed by them. *Google Scholar* offers basic and (fairly crude) advanced searches, but it is not possible to email, save or manipulate these records in any meaningful way. Citation indexing is available and results are presented roughly in order of the number of times they have been cited.

2.3. Scopus

Scopus, launched in 2004, is a multidisciplinary database with citation indexing. It indexes about 14,000 journal titles, of which about 2850 are from the social sciences ([Scopus Content Coverage, 2006](#), p. 8). Content coverage varies dependant on subject, but for the social sciences this goes back to 1996. From this date, *Scopus* has cover-to-cover indexing of contents, subject to minor exclusions. *Scopus* has basic, author and advanced searches with tools available to manipulate the search results. Cited references can be counted, followed and tracked in the conventional manner with links across the whole database.

2.4. Web of Science

The *Web of Science*, a multidisciplinary database is made up of three citation indexes: *Science Citation Index Expanded*, *Social Sciences Citation Index* and *Arts & Humanities Citation Index*. *Social Sciences Citation Index* started in 1973 and has retrospective coverage going back to 1956. This database indexes the contents of about 1900 journals on a cover-to-cover basis, but it also indexes selectively from over 3000 other journal titles. Additionally, a limited number of conference proceedings and monographs are indexed. The *Web of Science* has general, cited reference and advanced search features, with an extensive range of tools with which to manipulate search results. Cited references can be followed, tracked, counted, processed and analysed across all three of the databases.

2.5. Other source issues

To benchmark foreign language holdings, the *International Bibliography of the Social Sciences*, a bibliography managed by the London School of Economics and Political Science was used. It regularly indexes over 2800 journals which broadly cover the social sciences. A substantial part of these holdings are published in German, French, Italian and Spanish.

3. Methods

The submission records for the 13 UoAs these were extracted from the Higher Education & Research Opportunities (HERO) website which holds the submission records for the last (2001) RAE ([HERO, 2006](#)). The 2800 journal titles contained within the IBSS were also extracted ([About IBSS, 2006](#)). With the exception of *Google Scholar*, each of the selected databases' journal holdings were also extracted by their respective International Standard Serial Number (ISSN).

3.1. Record processing

Journal articles (which comprised 66.2% of the submissions to the selected UoAs) were checked for their accuracy using their ISSN and journal title against [Ulrich's Periodicals Directory \(2006\)](#). In this process, 720 (2.1%) unverifiable article records were found and discarded. Journal titles were weighted by the frequency with which articles had been submitted to them. Thus, a journal which had one submission to it would have a weighting of one, two submissions would have a weighting of two and so on. Where journals had made clear unambiguous title changes after 2001 and had three or more articles submitted to them in their earlier title, the new journal title and its ISSN were added to the listing with the same journal weighting. Overall 4594 unique journal titles with a total of 33,533 associated article records

were collected. No account was taken of the frequency with which multiple authors of the same article submitted the same article independently of each other to the 2001 RAE.

Turning to the non-English journal titles, these journal records from the IBSS database were checked against *Ulrich's Periodicals Directory* (2006) to verify their country of origin. This ensured that French titles were in fact published in France rather than, say, Canada or that a Spanish title was published in Spain rather than South America. From this analysis, 581 journal titles were identified as being published in the required countries, split between France 318, Germany 186, Italy 26 and Spain 81.

For the analysis of *Google Scholar's* journal holdings, a statistically valid sample of 380 journal titles, randomly selected, was taken from the complete listing of journal titles submitted to the UoAs. On a similar basis, a sample of 229 records was selected from the 581 IBSS journal titles which cover the four countries given above.

Whilst the selected databases may hold a particular journal title it is important to verify whether they have indexed its content or not. To do this, a statistically valid random sample of 306 articles drawn from different journals was used to check the holdings of the four databases. This process established whether the specific article could be located and the number, if any, of citations to the article.

3.2. Evaluation procedures

The ISSNs from the processed listings were compared record by record to the holdings of *CSA Illumina*, *Scopus* and *WoS*. From this process, a count and percentage coverage was obtained for each database. A weighted holding for each database was calculated based on the matching journal records found and the frequency with which articles had been submitted to them. This, for example, weighted the *British Journal of Sociology* at 73, given that it had had 73 of its articles submitted to the 2001 RAE. Using the same method, the ISSNs from the IBSS foreign journal title listing was compared.

Using *Google Scholar's* advanced search feature, the 380 records selected through the sampling process were entered by journal title. The matching records from these searches were ranked into the following categories:

- Zero hits;
- citation only records;
- multiple websites listing the journal and various articles from them; and
- consistent hits leading to a single credible website.

Zero hits indicate no matching records were found. Citation only records are where matching records to the journal article have been cited but the original article record could not be found. Where multiple websites reported articles from the journal, these derived from publishers, aggregators or open access sources. Hits showing a high degree of uniformity usually led to the publisher's own website.

4. Results

4.1. Overall coverage

Table 1 shows the number of matching records found and the percentage coverage for each database against the 4594 unique journals titles identified in the process described above. The weighted UoA takes the matching journal titles found for each database and gives the sum of the number of articles submitted to them and the percentage reported is of all the articles submitted.

Table 1
Record count and percentage coverage for each database

Database	All UoAs (titles)	Weighted UoA (articles)
<i>Web of Science</i>	1994: 43.4%	20,265: 60.4%
<i>Scopus</i>	2324: 50.6%	22,996: 68.6%
<i>CSA Illumina</i>	2678: 58.3%	24,436: 72.9%

Table 2
Foreign journal coverage

	<i>Web of Science</i>	<i>Scopus</i>	<i>CSA</i>
IBSS (581 titles)	71: 12.2%	61: 10.5%	140: 24.1%

Table 3
Coverage by *Google Scholar*

	All UoAs	IBSS
Zero hits	15: 3.9%	211: 92.1%
Citations only	38: 10.0%	3: 1.3%
Multiple websites	111: 29.2%	8: 3.5%
Single website	216: 56.8%	7: 3.1%

Taking the journal holdings from the IBSS for France, Germany, Spain and Italy and comparing this to the holdings of each of the three databases in the same manner as Table 1 above, then coverage details are as given in Table 2.

Of the 581 IBSS journals identified for benchmarking, 92 (15.8%) were used by authors in their submissions to the 2001 RAE.

4.2. *Google Scholar*

Extrapolated coverage of the 4594 UoA journal titles and the 581 foreign journal titles by *Google Scholar* are given in Table 3 below. These are ranked by the nature of the hits, as defined above. Sample sizes were, respectively, 380 from the 4594 UoA titles, and 229 from the 581 foreign language titles.

The 216 records representing 56.9% of the UoA sample, and the 7 records from IBSS which could be attributed to a single website, can be compared to those matches found by the record by record comparison for the other databases.

4.3. *Article and citation coverage*

The coverage test at journal, article and citation level showed that all the databases, with the exception of one title in *CSA Illumina*, had the journal title records that they claimed to hold, albeit, at differing levels. *Google Scholar's* holdings are of course unknown so it is not possible to say what may be missing. Table 4 shows the coverage.

Analysis of 15 records from the 65 article records not found by *CSA Illumina* showed that 11 of these had not been indexed, even though other articles from the same year or journal issue had been. In the case of the other four articles, the year in which they were published the journal did not appear to have been indexed at all. Further analysis showed that 14 of these journals had the minimal 'selective' category for indexing their content, and the remaining journal had 'priority' status only.

Using the 'Journal and article found percentage coverage' from Table 4 with the earlier results from Table 1 for the 'Weighted UoA', an overall estimate of the coverage was calculated. These are shown in column 4 in Table 5 below. This result is based upon taking the journal and article found percentage and assuming that this find rate can be used

Table 4
Coverage at article level

	<i>Web of Science</i>	<i>Scopus</i>	<i>CSA Illumina</i>	<i>Google Scholar</i>
Journal not found	0	0	1	12
Citation only record (GS only)	–	–	–	29
Journal found but not the specific article record searched for	37	15	65	0
Journal and article found and percentage coverage	269, 87.9%	291, 95.1%	240, 78.4%	265, 86.6%
Articles found but no citations noted	31	27	140	25
Average citation count per article found	13.7	14.5	3.1	17.7

Table 5
Record count and percentage coverage for each database

Database	All UoAs titles	Weighted UoA by article frequency	Re-weighted by percentage found
<i>Web of Science</i>	1994: 43.4%	20,265: 60.4%	17,782: 53.1%
<i>Scopus</i>	2324: 50.6%	22,996: 68.6%	21,869: 65.2%
<i>CSA</i>	2678: 58.3%	24,436: 72.9%	19,158: 57.1%

Table 6
Paired coverage of submissions to the 2001 RAE

Additions	<i>CSA Illumina</i>		<i>Scopus</i>		<i>Web of Science</i>	
	<i>Scopus</i>	<i>WoS</i>	<i>CSA</i>	<i>WoS</i>	<i>CSA</i>	<i>Scopus</i>
Titles	497	361	851	170	1045	500
Weighted article count	3423	2415	4863	707	6586	3438
Article recall adjustment	3255	2123	3813	621	5163	3270

Table 7
Combined coverage

	<i>CSA Illumina</i>		<i>Scopus</i>		<i>Web of Science</i>	
	<i>Scopus</i>	<i>WoS</i>	<i>CSA</i>	<i>WoS</i>	<i>CSA</i>	<i>Scopus</i>
Total titles	3,175	3,039	3,175	2,494	3,039	2,494
Total weighted articles	27,859	26,851	27,859	23,703	26,851	23,703
Article recall adjustment	22,413	21,281	25,682	22,490	22,945	21,052
Overall article coverage %	66.8	63.5	76.6	67.1	68.4	62.8

as a satisfactory measure of recall rate. Multiplying this recall rate against the original ‘Weighted UoA’ score gives the expected coverage from the articles submitted to the 2001 RAE.

4.4. Overlap analysis

There is a common set of 1516 journals that each of the three databases share with each other, and these account for 17,253 (51.4%) of the 33,533 articles. Table 6 below shows the overlap of the databases in relation to each other. The relationship shows the effect of combining each of the databases on a paired basis. For example, *CSA Illumina* covers 2678 of the journal titles submitted to the 2001 RAE and is shown as the primary source of coverage. *Scopus* covers a further 497 journal titles that *CSA Illumina* does not cover and their value is shown in terms of the 3423 articles that those additional titles would bring. This additional article count is then adjusted by the percentage coverage factor from Table 4; hence the article coverage is adjusted by a factor 0.951.

Table 7 combines the results of Tables 5 and 6. The original 2678 titles covered by *CSA Illumina* shown in Table 5 have been combined with the 497 additional titles from *Scopus* shown in Table 6 to give a total of 3175 journal titles which the two databases cover from the 2001 RAE submissions. Likewise, the article counts have been combined along with adjustments for the percentage coverage factor. The ‘Overall article coverage %’ varies between each pairing because each of the databases has a different percentage coverage factor. Finally, an overall coverage figure is given in terms of the number of articles covered compared to the total of 33,533 articles that were identified.

5. Discussion

5.1. Coverage

From the initial analysis (see Tables 1 and 2), it appears that *CSA Illumina* has the greater coverage in terms of journal titles for all the UoAs taken together and the IBSS database collection. In a simple ranking by total jour-

nal coverage, *CSA Illumina* is followed by *Google Scholar*, *Scopus* and then by *Web of Science*. *Google Scholar* was, however, the least successful database in the way that it presented the results it did find. Whilst the other databases presented their results chronologically, in an orderly manner, by volume and then by issue, the results from *Google Scholar* were generally ranked by citation count irrespective of volume or issue sequence. *CSA Illumina* provided the best coverage of the foreign journals selected from the IBSS database, by a factor of almost 2 when compared to the *Web of Science* with *Google Scholar* giving the worst coverage, finding only 7 of 229 titles tested.

Whilst *CSA Illumina*'s success at journal title level appears to be good, this good result was not maintained when individual articles submitted to the 2001 RAE were presented to it. From the sample of 306 articles presented to each of the databases, *CSA Illumina* was the least successful, scoring only 240 hits out of the 306 possible. *Scopus* was best at 291 hits with *Web of Science* finding 269 and *Google Scholar* finding 265. The *Web of Science* and *Scopus* appear to have a broader and more robust 'cover to cover' indexing policy, *CSA Illumina*'s policy appears to be more variable and is dependent on how close the journal's contents is to the main discipline of the database in which it is held. Examination of the selection policy of the 15 sample journals where articles could not be found showed that their selection was not 'core', so hence not all articles would have been indexed.

Given the newness of *Scopus* and *Google Scholar*, these two databases have been frequently reviewed and compared and in several cases they have been compared to the *Web of Science*. In a number of papers, Jacso (2005a, pp. 208–214, 2005b, pp. 1537–1547) has discussed the limitations of *Google Scholar*. He has concluded that it is unreliable and unpredictable in the results it returns, both in its links to the sources it has found and in its coverage. This view of *Google Scholar* is also shared, generally, by others who have also found significant omissions in the coverage and recall from this database (Myhill, 2005; Notess, 2005). It is evident, however, that most reviewers feel that *Google Scholar* has the potential to become a useful source of scholarly information provided its shortcomings are addressed.

Like *Google Scholar*, *Scopus* has been subject to close scrutiny, Deis and Goodman (2005) have compared the *Web of Science* and *Scopus* as has LaGuardia (2005). Others (Burnham, 2006; Dess, 2006; Jacso, 2004) have also reviewed *Scopus* extensively. These reviewers generally acknowledge that both of these databases are primarily concerned with the sciences. The number of science based records held by each greatly outweighs their holdings in the social sciences and the arts. Nevertheless, both databases still have large holdings of records in the social sciences and these are comprehensively indexed, although *Scopus* has only done this from 1996. This lack of coverage by *Scopus* is a noted concern, Dess (2006) recommends the *Web of Science* as the only plausible option for searching prior to 1996 for interdisciplinary subjects such as the social sciences. When, however, Dess (2006) carried out citation searches for post 1996 records, *Scopus* had a small advantage of 1.2% in the number of citations that it found over the *Web of Science*. The difference found in the citation count test here, in average citations per article found was a 5.4% advantage in favour of *Scopus*.

5.2. Citation coverage

The use of citation counts to measure research performance in the social sciences and humanities is more problematic than in the sciences. Nederhof (2006, pp. 81–100) acknowledges the shortcomings of the *Web of Science* to cover adequately the social sciences and in particular the humanities, given the propensity of academics in the latter to publish more in monographs than in serials.

Despite this tendency to publish in non-journal sources in the humanities, Nederhof (2006, p. 86) in his review considers that "In most disciplines in the social sciences and humanities, journals were found to be the single most important medium for publication. [and] In the behavioural sciences and economics, journal articles account for the majority of citations". There is a recognised trend: by social scientists publish more of their work in journals than was formerly the case. Adams (n.d.) has analysed the increase in journal submissions between the 1996 and 2001 RAE, and found that these submissions increased as a percentage of all submissions from 42.4% in 1996 to 54.1% in 2001 for the social sciences. In the UoAs included for analysis in this work, journal articles made up 66.2% of all submissions.

The citation count results from Table 4 show that there is significant variation in the number of citations that have been indexed between the different databases. In many cases, individual article counts from the *Web of Science* and *Scopus* were very closely matched. *CSA Illumina* reported the lowest citation counts and had the greatest number

of articles without any citations. This is not surprising since where *CSA Illumina* does index cited references, it has done so only fairly recently and for a limited range of its databases (*Cited Reference Linking*, n.d.). It also has very limited citation links between citing records from different, but related databases. Whilst *Google Scholar* has the highest citation counts, when these are examined individually, it is clear that the results have not, at the very least, been de-duplicated. This makes the citation counts from *Google Scholar* highly suspect. *Jacso (2006)* in an examination of how *Google Scholar* counts and reports citation counts, found that it was consistently unreliable.

Clearly, citation coverage varies, depending on the database being used. *Web of Science* has the greater coverage historically and its functionality and sophistication exceeds the other databases considered. *Scopus* has the second largest coverage, in terms of citations and the closest functionality to that offered by *Web of Science*. Whilst the depth of cited reference coverage is not as great as *Web of Science*, cited references in *Scopus* can be searched, analysed and processed in a very similar way, although this may take a little longer than for the *Web of Science*. Citation indexing in *CSA Illumina* is currently only available in some of the databases and is often limited to only core journals. All of the databases other than *Google Scholar* returned search results by volume and issue number or in an order selected by the user.

6. Conclusions and recommendations

All of the databases examined give significant coverage of the social sciences in terms of journals and articles indexed as well as the indexing of cited references. There are, however, noticeable differences between the databases. In its current form, *Google Scholar* cannot seriously be thought of as a database from which metrics could be used to measure scholarly activity.

Whilst *CSA Illumina* has the greatest journal coverage overall, its coverage at article level is disappointing as is the number of journals from which it indexes cited references. However, when the results are weighted by article submission frequency, *CSA Illumina* does rank second in its coverage. This ranking is improved to first position when the database has its coverage combined with *Scopus* or the *Web of Science* as major partners; this is, respectively, 76.6 and 68.4%. Despite this good coverage in combination, when compared to the other databases, however, the citation count from *CSA Illumina* per article appears to be significantly out of step, suggesting that the base from which it is collecting these records is too small. The analysis of cited references is limited, users can identify citing and cited authors; but analysis and record processing are less sophisticated than *Scopus* or the *Web of Science*. *CSA Illumina* falls too far below the standards set by *Web of Science* and *Scopus*, even though it has good article coverage, to be considered a serious option to cover the social science literature and in particular, to be used to measure scholarly activity and impact in this field.

Taking an overall view, both *Web of Science* and *Scopus* offer the best coverage at journal, article and cited reference level. Both index their journal holdings on a cover to cover basis. They would together give an article coverage rate of 67.1%. Both, on the other hand, are rather weak on their coverage of foreign journals and *Scopus* does not currently go further back than 1996 for the social sciences. Given that coverage by *Scopus* is considered good, and its tools to analyse citation counts are sufficient, then arguably it offers the best choice from amongst the multidisciplinary databases reviewed here. On this basis, notwithstanding its poor foreign journal title coverage, it is suggested that *Scopus* could be used as an alternative to the *WoS* as a tool to evaluate research impact in the social sciences.

As indicated earlier, finding an adequate set of book and report records with which to credibly benchmark database holdings from the 2001 RAE submissions has not been successful. Examination of the submissions to the 2001 RAE show that reports make up less than 2% of the total and *Nederhof (2006, p. 95)* in his review suggests that for 'grey' publications such as unpublished reports, that their impact is "rather disappointing". Although monographs remain an important social sciences scholarly output, recent research has shown that there is a steady movement towards publication in journals and away from monographs. It is expected that this trend will continue. Therefore, the significance of this lack of coverage will decline over time.

Acknowledgements

We are grateful for the ESRC for funding this research and to Dr. J. Cooper and R. Cornish for technical help and statistical advice.

Appendix A

Units of Assessment (UoA)

UoA number	Name
13	Psychology
34	Town and Country Planning
35	Geography
37	Anthropology
38	Economics and Econometrics
39	Politics and International Studies
40	Social Policy and Administration
41	Social Work
42	Sociology
43	Business and Management Studies
44	Accounting and Finance
56	Linguistics
68	Education

References

- About IBSS. (2006). <<http://www.lse.ac.uk/collections/IBSS/about/alphabeticalJournals.htm>> Accessed 15.05.06.
- Adams, J. (n.d.). *Research Assessment and UK publication patterns*. <www.uksg.org/presentations8/adams.pps> Accessed 22.05.06.
- Burnham, J. (2006). Scopus database: A review. *Biomedical Digital Libraries*, 3(1). <<http://www.bio-diglib.com/content/3/1/1>> Accessed 30.05.06.
- Cited Reference Linking. (n.d.). <http://www.csa.com/help/Advanced_Search/cited_reference.html> Accessed 05.05.06.
- Deis, L. & Goodman, D. (2005). Web of Science (2004 version) and Scopus. *The Charleston Advisor*, 6(3). <<http://www.charlestonco.com/comp.cfm?id=43>> Accessed 03.05.06.
- Dess, H. (2006). Database reviews and reports: Scopus. *Issues in Science and Technology Librarianship*. <<http://www.istl.org/06-winter/databases4.html>> Accessed 03.05.06.
- HERO Higher Education Research Opportunities. (2006). <<http://www.hero.ac.uk/rae/Results/index.htm>> Accessed 08.05.06.
- Hicks, D. (2004). The four literatures of the social science. In H. K. Moed, W. Glanzel, & U. Schmoch (Eds.), *Handbook of Quantitative Science and Technology Research* (pp. 473–495). Dordrecht: Springer.
- Jacso, P. (2004). *Péter's Digital Reference Shelf. Scopus*. <<http://www.galegroup.com/reference/archive/200409/scopus.html>> Accessed 20.09.05.
- Jacso, P. (2005a). Google Scholar: The pros and the cons. *Online Information Review*, 29(2), 208–214.
- Jacso, P. (2005b). As we may search—Comparison of major features of *Web of Science*, *Scopus* and *Google Scholar* citation-based and citation-enhanced databases. *Current Science*, 89(9), 1537–1547.
- Jacso, P. (2006). Deflated, inflated and phantom citation counts. *Online Information Review*, 30(3), 297–330.
- LaGuardia, C. (2005). *E-Views and Reviews: Scopus vs Web of Science*. *Library Journal.com*. <<http://www.libraryjournal.com/article/CA491154.html%22>> Accessed 03.05.06.
- Larivière, V., et al. (2006). The place of serials in referencing practices: Comparing natural sciences and engineering with social sciences and humanities. *Journal of the American Society for Information Science and Technology*, 57(8), 997–1004.
- Moed, H. K. (2005). *Citation analysis in research evaluation*. Dordrecht: Springer. Chapter 7: ISI Coverage by discipline
- Myhill, M. (2005). *Google Scholar*. <<http://www.charlestonco.com/review.cfm?id=225>> Accessed 03.05.06.
- Nederhof, A. (2006). Bibliometric monitoring of research performance in the social sciences and the humanities: A review. *Scientometrics*, 66(1), 81–100.
- Notess, G. (2005). *Scholarly web searching: Google Scholar and Scirus*. <<http://www.infoday.com/Online/jul05/OnTheNet.shtml>> Accessed 03.05.06.
- Roth, D. (2005). The emergence of competitors to the *Science Citation Index* and the *Web of Science*. *Current Science*, 89(9), 1531–1536.
- Scopus Content Coverage*. (2006). <http://www.info.scopus.com/docs/content_coverage.pdf> Accessed 19.06.06.
- Ulrich's Periodicals Directory*. (2006). <<http://www.ulrichsweb.com/ulrichsweb/>> Accessed 08.05.06.