2011

Coverage and overlap of the new social science and humanities journal lists

Diana Hicks, *Georgia Institute of Technology - Main Campus*
Jian Wang, *Georgia Institute of Technology - Main Campus*

Coverage and overlap of the new social science and humanities journal lists

Diana Hicks* & Jian Wang
School of Public Policy
Georgia Institute of Technology
685 Cherry Street
Atlanta, GA, 30332-0345
Phone: 404 385 6015
Fax: 404 385 0504

dhicks@gatech.edu

jianwang@gatech.edu

* Corresponding author

October 2010

Abstract

This is a study of coverage and overlap in second generation social sciences and humanities journal lists with attention paid to curation and the judgment of scholarliness. We identify four factors underpinning coverage shortfalls: journal language, country, publisher size and age. Analysing these factors turns our attention to the process of assessing a journal as scholarly, which is a necessary foundation for every list of scholarly journals. Although scholarliness should be a quality inherent in the journal, coverage falls short because groups assessing scholarliness have different perspectives on the social science and humanities literature. That the four factors shape perspectives on the literature points to a deeper problem of fragmentation within the scholarly community. We propose reducing this fragmentation as the best method to reduce coverage shortfalls.

Introduction

In the social sciences and humanities it is largely impossible to substantiate statements on excellence in scholarship with reliable indicators for international benchmarking of fields and institutions. A central problem in conducting useful, large scale evaluations in social science and humanities has been limited coverage of social science and humanities journals in the large databases. Useful evaluation requires an adequate bibliometric infrastructure, and this must have some claim to be comprehensive.

Many studies have documented the inadequacy of Social Science Citation Index (SSCI), and lately Google Scholar and Scopus coverage of social science and humanities (SSH) literature. These studies use several types of methodology. In one type of study a bibliography is compiled from sources such as an institution's annual report, end of award reports, submissions to the RAE, etc. and the share of this material also found in a database is calculated (Burnhill & Tubby-Hille, 1994; Royle & Over, 1994; Pestana, Gomez, Fernandez, Zulueta, & Mendez, 1995; Villagra Rubio, 1992; Norris & Oppenheim, 2007; Walters, 2007). A second method uses one database as a source of references to assess coverage of another database; we might call this database overlap analysis (Winclawska, 1996; Webster, 1998; C. Neuhaus, E. Neuhaus, Asher, & Wrede, 2006; Gavel & Iselid, 2008; Frandsen & Nicolaisen, 2008). A third methodology compares database coverage to a canonical source recognized to be an almost complete journal list. Ulrich's is currently used for this purpose (Archambault, Vignola-Gagné, Côté, Lariviere, & Gingras, 2006; de Moya-Anegón et al., 2007), and a UNESCO list has been used in the past (Schoepflin, 1992). The literature demonstrates that the coverage of SSH journals in SSCI, Scopus and Google Scholar is inadequate and so evaluations based only on analysing these databases would also be inadequate.

In response to this consensus, new evaluation methods have been developed and existing resources have been augmented. Bibliometricians have developed new methods such as analysis of non-indexed cited material and library catalogs (Butler & Visser, 2006; Torres-Salinas & Moed, 2009). Metric-based evaluation systems have been designed that are based on university submission of bibliographies. In 2009 Web of Science (WoS) and Scopus added a large number of SSH journals, increasing the size of the SSH list in WoS by 22% and in Scopus by 39%. And the European Research Index for the Humanities (ERIH)

was developed to showcase high quality European humanities research. These efforts have produced new, larger lists of social science and humanities journals, offering an opportunity to reassess the issue of coverage of social science and humanities literature. Here we compare the augmented WoS and Scopus, ERIH and journal lists developed as part of metric-based evaluation systems. We apply methods of coverage and overlap analyses to these post-consensus resources and ask whether the problem has been solved. Our method combines the canonical source approach, using Ulrich's, with a comparison of five journal lists. We find that the traditional coverage problem persists, even in augmented databases and in simple lists designed specifically to overcome the problem. To understand why this is the case, we identify four factors underpinning coverage shortfalls (language, country, publisher size and time) and assess their relative importance. We then suggest how coverage shortfalls finally might be overcome.

Qualifying as a journal - curation

In this study we use Ulrich's as a canonical source of the complete social science and humanities literature. Ulrich's is the authoritative source of bibliographic and publisher information on more than 300,000 periodicals of all types from around the world. It includes academic and scholarly journals, open access publications, peer-reviewed titles, popular magazines, newspapers, newsletters, and more. Ulrich's has been used in bibliometric studies as the benchmark against which WoS and Scopus coverage is measured (see for example: Archambault et al., 2006; de Moya-Anegón et al., 2007). Studies have found only very small numbers of journals that are not yet indexed in Ulrich's. We found 30-40 journals, all newer, that were not yet indexed. We told Ulrich's about these journals and they have been incorporated.

Using Ulrich's, we assess five lists of SSH journals: ERIH, the Norwegian reference list, the Australian ERA list, Web of Science (WoS) and Scopus. The European Reference Index for the Humanities, or ERIH, is a project of the European Science Foundation aimed initially to identify, and gain more visibility for top-quality European humanities research published in academic journals, potentially in all European languages (European Science Foundation, 2010). The Norwegian and ERA lists are the reference lists of journals whose papers are acceptable submissions to the Norwegian and Australian university research evaluation systems. WoS is Thomson-Reuters' Web of Science incorporating the Science Citation Index (SCI), Social Science Citation Index (SSCI) and Arts and Humanities Citation Index (A&HCI). Scopus is an Elsevier journal article and citation database. WoS and Scopus are really not journal lists, rather they are databases indexing the articles in a delineated set of journals, and we analyse those journal sets. All except ERIH are comprehensive across fields. We only analyse the SSH journals in them. We bought Ulrich's data. Ulrich's flags journals indexed in WoS and Scopus, and we used this to obtain the WoS and Scopus lists. We obtained ERIH and ERA lists from their websites and obtained the Norwegian list from Gunnar Sivertsen.

Table 1 compares these lists on several key dimensions. Note that the lists are built using two types of processes. Commercial products use an editorial process. In addition Scopus has a review board including scientists and librarians. ERIH and ERA use peer committee-based processes; the Norwegian list screens journals in an editorial fashion with difficult cases referred to a group of scholars. Several of the lists classify journals into levels,

recognizing that broadly distinguishing levels of quality is a necessity because the literature is vast and variable.

TABLE 1. Lists of scholarly journals in the social sciences and humanities.

Name	Process to choose journals	Estimated size of SSH Journal list[*]	Journal classification	Type
Ulrich's	Comprehensive, no filtering	25,195	refereed & academic	Comprehensive commercial database of periodicals
ERA	peer	9,854	4 levels	National evaluation system base journal list
Norwegian	editorial & peer	7,009	2 levels	National evaluation system base journal list
Scopus	editorial & peer	6,829	no	Commercial citation and journal article database
ERIH	peer	3,878	3 categories	Index of scholarly humanities journals
WoS	editorial	3,159	no, considered to be selective	Commercial citation and journal article database

[*]Size estimated as of October 2009. Only includes active, regularly appearing journals whose existence is confirmed by a match in Ulrich's.

The figure for the size of the list provided in Table 1 will not match the figure given by the list's source. As Gavel and Iselid describe in some detail, close inspection reveals errors in every journal list and database (Gavel & Iselid, 2008). The peer lists suffer from a rather high rate of error. The ERIH list we obtained in January 2009 had not been cleaned or checked for errors. It contained duplicate records with slight differences in title or typos in ISSN, as well as erroneous ISSN numbers and titles. Journal publishing is dynamic: journals merge, change names and evolve; both ERIH and the Norwegian list contained old ISSNs. We cleaned up the lists to remove these errors, thus our lists are shorter than the originals.

A second way in which our version of the lists diverges from the originals is that we define the scholarly literature to include only active, current journals. ERIH and the Norwegian list contain journals that have ceased publication, are suspended, are published irregularly, and journals whose status is unknown. Excluding such titles produces a "level playing field" for comparison with WoS and Scopus who exclude such journals. This issue has not always been recognized in previous studies of WoS and Scopus coverage (for exceptions see: de Moya-Anegón et al., 2007; Gavel & Iselid, 2008). We would argue that an evaluation infrastructure should, like the databases, cover active regularly appearing journals. This is because the world of publishing is vast and many vehicles of dubious status come and go. It is not unfair to ask SSH researchers to focus on, and support, outlets with quality standards and some on-going existence. There is in addition the problem that it is impossible to guarantee consistent coverage of a set of transient material unless resources would be infinite.

Unfortunately, periodical publication is not the tightly controlled world portrayed in the Web of Science. Rather it is a vast, ever shifting and heterogeneous enterprise, with many pretenders to scholarly status. In consequence, careful curation is essential to produce a sound journal list. Curation is in general an invisible and undervalued activity. Peer groups

constructing lists of scholarly journals and analysts comparing lists could usefully place more value on seemingly mundane considerations of accuracy and journal status in order to enhance the quality of their lists and analyses.

Scholarliness

Of course, regular appearance alone is not enough to qualify a periodical for a list of academic journals. A journal also must be scholarly. But what does this mean, exactly? How is scholarly defined? Do people agree on what is and is not scholarly? Are there certain characteristics of journals that make them more likely to be certified as scholarly? The lists themselves provide worked out answers to these questions which we explore here.

To begin, we examine the accession criteria articulated by the list-makers. First, mundane criteria enter into the assessment of scholarliness. Geographical diversity of authorship is important for the Norwegian list, ERIH, WoS and Scopus. Formal, editorial qualities such as regular appearance and correct formatting are important for WoS and Scopus (Elsevier, 2010; Thomson Reuters, 2010). Though formal criteria tend to be neglected by the scholarly community, Gimenez-Toledo and Roman-Roman argue against this because formal criteria are related to parameters that scholars value, such as quality of editors (Gimenez-Toledo & Roman-Roman, 2009). As we described above, this analysis also places value on these qualities by including only journals known to be active.

Of more interest here is the more highly esteemed criterion of scholarly quality. Every list claims to include only peer reviewed journals. The lists differ in their processes for identifying peer reviewed journals. ERIH and ERA were assembled by groups of scholars convened for the purpose of list construction. The Norwegian list comprises all journals used by Norwegian scholars that meet the criteria of presenting new insights in a form that allows the research findings to be verified and/or used in new research activity in a language and with a distribution that makes the publication accessible for a relevant audience in a publication channel with peer review that is not limited to the output of one institution (Sivertsen, 2010, p. 24). Norwegian scholars request that journals be added to the list and a candidate journal is assessed by administrators who confirm the peer review status of the journal with the publisher if necessary. If there is doubt, the candidate journal goes to a National Review Board for decision. WoS is compiled by editors who use indicators such as citations to the journal, its editors and authors. Scopus uses editors as well as a committee of scientists and librarians who score journals on criteria including: convincing editorial concept/policy, level of peer review, academic contribution to field, and citations to journal and editors (Elsevier, 2010, p. 21). Ulrich's differs because it aims for maximum title coverage of serials. However, it identifies peer reviewed journals:

> *The Ulrich's editorial team assigns the "refereed" status to a journal that is designated by its publisher as a refereed or peer-reviewed journal. Often, this designation comes to us in electronic data feeds from publishers. In other cases Ulrich's editors phone publishers directly for this information, or research the journal's information posted on the publisher's website* (SerialsSolutions, 2010).

Unfortunately, Ulrich's simply tags journals as refereed. If a journal is not tagged, we do not know if it has been confirmed to be not refereed, or if its status is unknown. Comparison with the lists suggests that Ulrich's refereed status seems to be incomplete, particularly for non-English language journals.

Lists differ in the processes used to assess journal scholarliness, and they seem to come to different conclusions, as evidenced by the differing lengths of the lists. Given the variability in accession criteria between the lists, it is useful to apply a single criterion to all lists to assess the overall scholarliness of their content. All lists claim to be restricted to scholarly material. However, lists are found to contain material assessed as non-academic by Ulrich's, such as consumer/magazines or trade journals. For example, in history ERIH includes coin collecting magazines. We would argue that the stated intent of ERIH to cover quality, peer reviewed journals is correct; publishing in non-scholarly journals is important for reaching the general public, but should be dealt with separately as enlightenment rather than scholarly literature. If the first priority is advancing evaluation of scholarly publishing; enlightenment literature should be clearly differentiated (Hicks, 2004).

To assess list coverage against the universe in Ulrich's, we needed to narrow down Ulrich's list to academic journals only. Analysis suggested that Ulrich's "refereed" status was incomplete. Ulrich's "academic/scholarly" status was better, though too broad in including newspapers for the university market, such as the *Chronicle of Higher Education*, and too narrow in classifying some journals as "trade" (*Energy Economics* was classified as trade rather than academic/scholarly). Therefore, we devised the following definition. All periodicals classified as "academic/scholarly" by Ulrich's were labelled academic by us as well except newspapers, newsletters, bulletins and magazines – which were only labelled academic if they were also on two of the other lists. In addition, any periodical on two of the other lists was labelled academic if Ulrich's had not classified the periodical's type or if Ulrich's had classified the periodical as "trade." Finally, the serials of four publishers with "academic/scholarly" status in Ulrich's were excluded because they were middle school curriculum guides, test study guides, compilations of articles for use in the classroom etc.[1] Using this definition, we analysed the overall academic content of the lists by calculating the share of non-academic material in them, see Table 2. We can see that WoS (2.1% non-academic) has the most credible claim to being a purely academic database. Next are ERA (2.9%), the Norwegian list (3.3%), ERIH (5%) and finally Scopus (9.3%).

TABLE 2. Share of non-academic journals.

List	Journals	% non-academic	% non-academic Also in WoS	Not in WoS
Scopus	6,829	9.3	1.7	14.8
English	5,686	8.1	1.5	13.7
Non-English	1,143	17.8	4.2	18.8
ERIH	3,878	5.5	0.6	8.9
English	2,538	2.5	0.5	4.9
Non-English	1,340	13.0	1.4	13.0
Norwegian	7,009	3.3	1.9	4.2
English	5,838	2.4	1.6	3.0
Non-English	1,171	8.2	4.3	8.7
ERA	9,854	2.9	0.4	3.9
English	8,402	2.7	0.4	3.7
Non-English	1,452	4.3	0.4	5.0
WoS	3,159	2.1		
English	2,794	1.7		
Non-English	365	5.2		

The table includes a breakdown by language of the journal which shows that the share of non-academic material is much higher for non-English language journals. Academic status is clearly contested with the distinction between international and national literatures pivotal. Taking English language as defining international literature (which is handy but not entirely true), there is much more agreement between the lists and Ulrich's definitions of academic for internationally oriented journals. Identifying the academic part of national literatures seems to be far more difficult because the share of non-academic material is much higher in the non-English portion of the lists. It is likely very difficult to devise and consistently apply criteria of academic quality across a range of languages. Indeed, WoS has only recently taken on this challenge with its campaign to extend coverage to "regional" journals. Given the importance of national language publishing in SSH (Hicks, 2004), solving the problem of consistent, evidence-based criteria for journal scholarly quality that can be applied impartially and without favouritism across the range of European languages will be crucial to building a respected bibliometric infrastructure for SSH.

The table splits the contents of the lists into journals also indexed in WoS, and the rest. The material not indexed in WoS has a considerably higher share of journals whose academic status is open to question. Thus, the only thing all parties seem to agree on is that journals indexed in WoS are academic. As we will see below, the other lists basically incorporate WoS and build out from there. Thus, WoS, which was the first to attempt to identify and index academic journals, has become the de facto standard to define the

scholarly. This was first noticeable when evaluation systems, such as the Australian Composite Index simply allowed submissions of WoS indexed material and China and Korea started rewarding scholars for WoS indexed papers. Of course, WoS's definition of the scholarly has been criticised, but not for including junk, rather it is attacked for being too narrow, particularly in its coverage of non-English language, non-Anglo-Saxon material (Archambault et al., 2006; Pestana et al., 1995; Villagra Rubio, 1992; Winclawska, 1996; Webster, 1998; Schoepflin, 1992). Table 2 suggests that though the criticism may be fair, there is little consensus on how to extend the journal list beyond WoS.

List Coverage

All the lists analysed here respond in some way to the finding that the SSH literature is larger than has been indexed in the past, but how much progress has been made in adequately identifying the SSH literature? To answer this question, we analysed list coverage. We define a list's coverage as the share of academic journals listed in Ulrich's that are also found on the list. As suggested above, carefully defining the field of legitimate publication will be crucial to the quality of the coverage analysis. Because we recognize the importance of formal parameters of journal quality, only active regularly appearing journals are analysed. Because it has been established that pre-qualifying journals as scholarly substantially raises coverage figures, (Burnhill & Tubby-Hille, 1994; Nederhof & Zwaan, 1991; Schoepflin, 1992), we limit the analysis to academic journals, using the definition of academic journals devised above. In addition we restrict this analysis to journals published in a European country or in the United States.[2] Finally, because the field coverage of ERIH, ERA and the Norwegian list varies slightly, we constructed a thesaurus of field names matched to Ulrich's field names and used this to restrict each comparison to fields covered by the list. The results of the coverage analysis are reported in Table 3.

TABLE 3. Journal coverage: Share of Ulrich's academic journals found on list

	All journals	English	Non-English	English language		
				Old journals	From US, UK or Neth.	Large publisher
ERA	39	54	18	58	59	77
Norwegian	28	37	15	42	43	65
Scopus	25	34	12	39	40	61
ERIH*	24	27	23	30	28	49
WoS	12	18	5	21	22	35

* Calculated on a smaller group of fields, humanities only.

The first column shows that no resource covers more than 40% of the available SSH academic literature. Not surprisingly, we see that the simpler lists of journals - ERA, ERIH and the Norwegian list - are larger than the more complex databases of articles – Scopus and WoS. Disappointingly, no list is adequate if the goal is to provide a comprehensive guide to SSH academic journals.

Coverage is better if we restrict ourselves to English language journals. Columns three and four demonstrate that English language coverage is higher than non-English language coverage. This is not surprising in light of previous studies. ERA achieves the

highest coverage at 54% of Ulrich's English language academic journals. At 37% and 27% the Norwegian list and ERIH are notably lower. Coverage of non-English language journals is lower in every list. ERIH is unique in its emphasis on non-English language journals. Its coverage of non-English language journals is almost as strong as its coverage of English language journals. In addition, ERIH has the best coverage of non-English language material, at 23%.

Other factors in addition to language influence the chances of a journal being listed. The last three columns in Table 3 report coverage figures for English language journals only. Ulrich's captures even the newest journals, while lists and databases lag by some years. We can raise the English language coverage figures for ERA and the Norwegian list to 58% and 42% respectively by considering only journals founded before about 2006.[3] Journals published in the United States, United Kingdom or the Netherlands are also more likely to be found on lists. Another relevant factor is size of publisher. Journals published by large publishers, defined as publishing 15 or more academic SSH journal titles, are much more likely to be covered.

Clearly coverage is incomplete. Well established journals published by large publishers, that appear to be scholarly but are not included in any list except Ulrich's include: Journal of Reformed Theology (Brill), Equality, Diversity and Inclusion (Emerald), Baha'i Studies Review (Intellect), Sikh Formations (Routledge), Wege zum Menschen (Vandenhoeck und Ruprecht) and so on. Other factors that appear to put journals at risk for being ignored by the lists include being about non-Christian religion, being of purely regional American local interest or being an Inderscience journal with a title beginning: "International Journal of . . ."

To summarize, we analysed the share of Ulrich's academic journals found on each list, or list coverage. We found that coverage varied a great deal depending on a host of factors including whether a journal was published in English or not, whether a journal was published in the US/UK/Netherlands or not, whether the journal was new or not, and the size of the journal's publisher. Since each list claims to be a comprehensive representation of the scholarly literature in SSH, one might conclude that scholarliness of journals depends on language or country of publication, age of journal and publisher size. Yet none of these factors is articulated in accession criteria.

Overlap and consensus

The discussion of coverage analysed each list's relationship to Ulrich's. But what about the lists' relationships with each other? To understand this we need to analyse overlap. Overlap analysis can be quite complex because of the many dimensions to analyse. Venn diagrams are often used (Gavel & Iselid, 2008; Gluck, 1990), and because they are so accessible, we use them here. This required that we simplified a 5-way comparison that would require MDS into three, three way comparisons with a fourth list, Ulrich's easily added since by definition it includes the rest. The results of the analysis are shown in Figure 1. Each Venn diagram reports the overlap between a list (ERA, ERIH, Norwegian) and Scopus, WoS and Ulrich's. For each list there are two diagrams, one for English language journals and one for other languages. The Venn diagrams are all drawn to the same scale, thus the Ulrich's circle for foreign language journals is smaller than the Ulrich's circle for

English language journals because Ulrich's contains more English language journals. The Ulrich's circle varies in size for different lists because the field coverage of each list differs. The area of intersection for WoS, Scopus and each list is shaded in grey.

The Venn diagrams reprise the coverage results which are displayed as the ratio of the areas of list and Ulrich's circles. This ratio equals the percentage coverage reported above. Thus the smaller coverage of non-English language journals equates to a smaller part of the Ulrich's circle covered in the non-English language Venns.

Examining the overlap between the circles can tell us more. Overall, the lists do not just differ in size, and therefore coverage, they also choose different journals. The set of journals shaded in grey represents maximum consensus; in each case this area is notably smaller than the union of all lists would be. On ERA and the Norwegian Venns, WoS is most completely shaded grey, indicating that those lists as well as Scopus incorporate almost all of WoS. This substantiates the point made above that WoS indexing has come to signify acceptance as a scholarly journal. In addition non-English language journal sets overlap less, signalling greater disagreement over which journals are scholarly. The Venn diagrams demonstrate that the lists and databases overlap a great deal, but each contains journals not indexed by anybody else except Ulrich's.

Figure 1 - Journal list overlap by language.
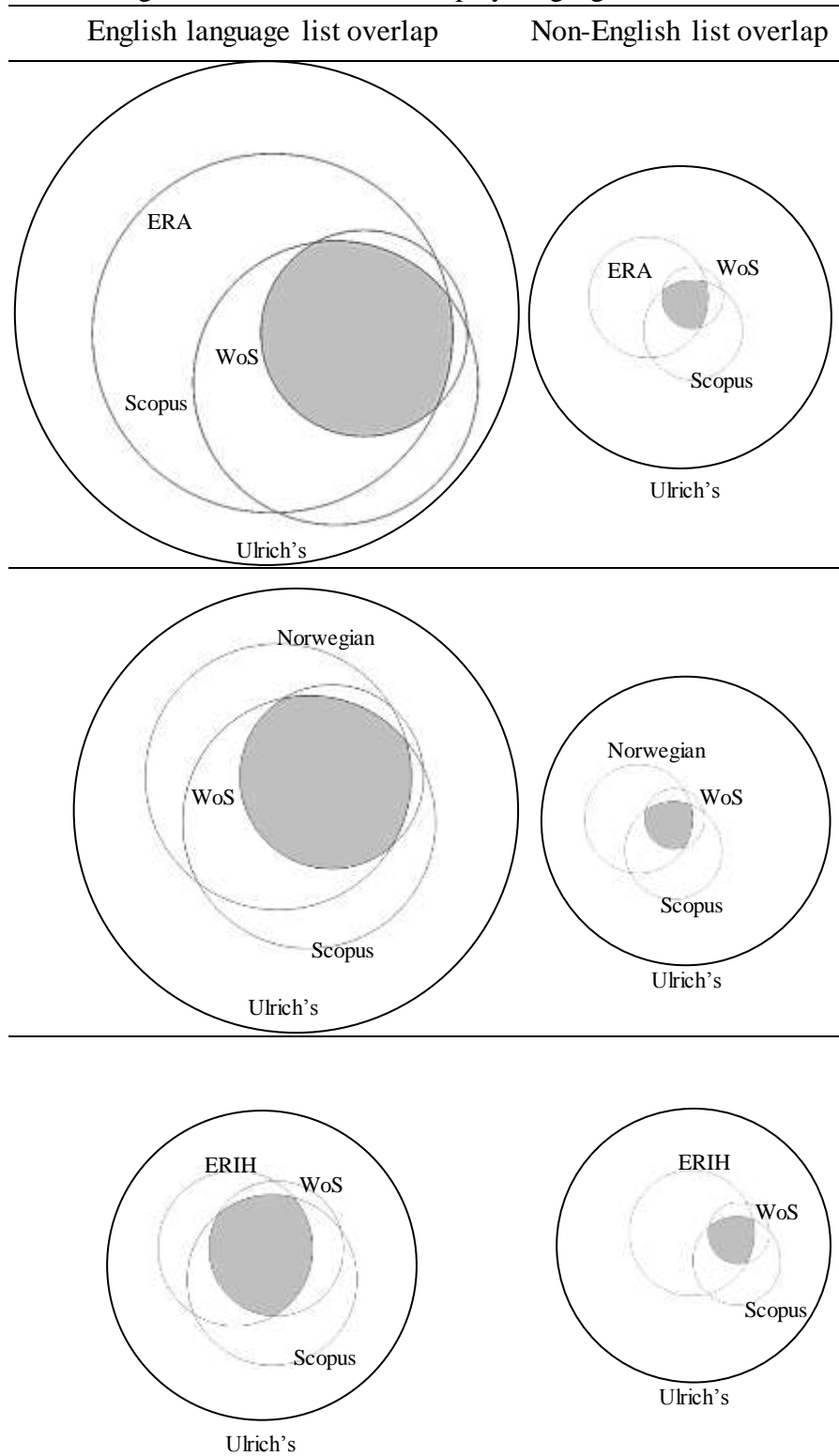
Figure 2- Consensus difficult outside large publishers, English language and leading publisher countries.
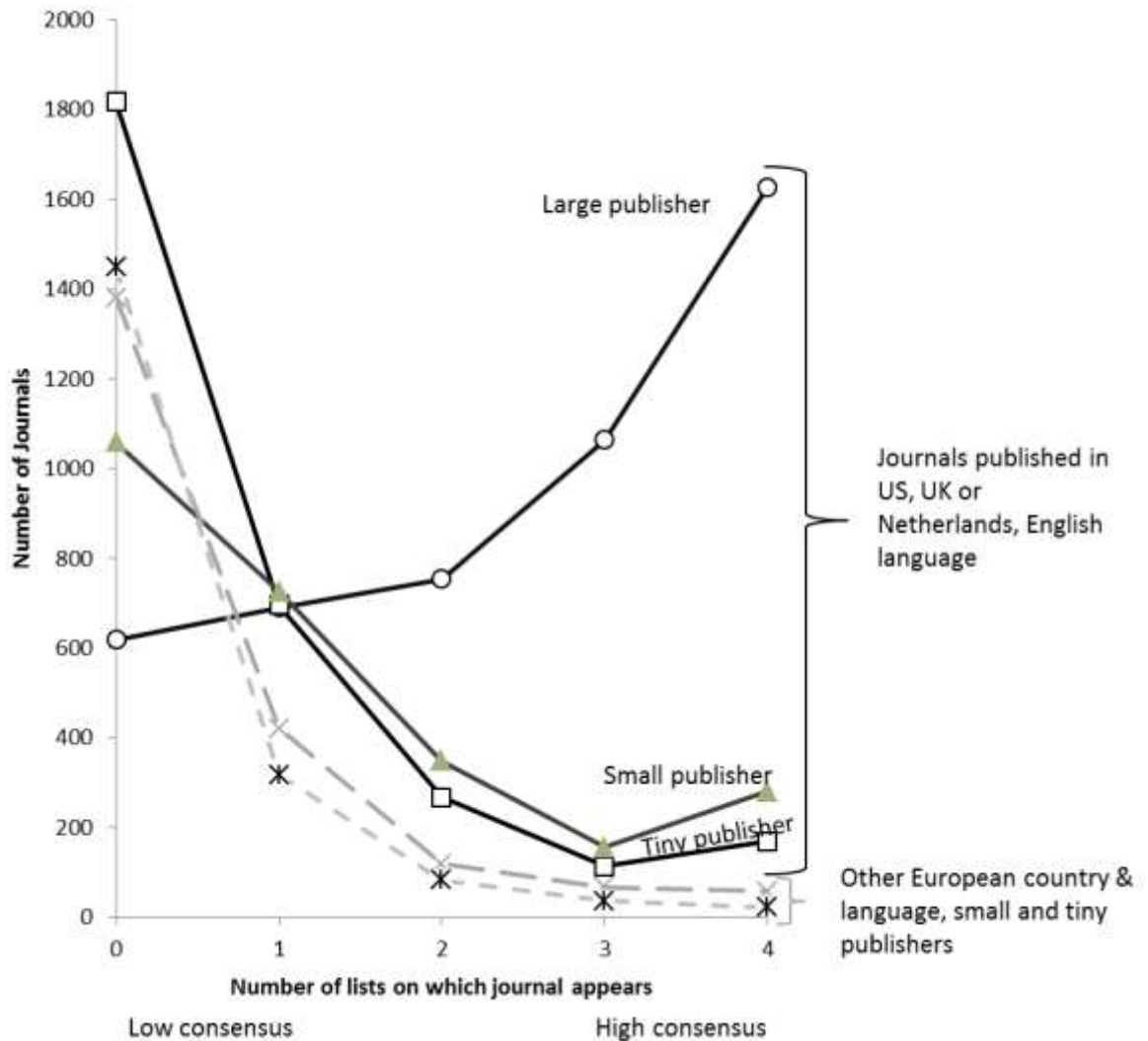
The less than 100% overlap in these diagrams is disappointing as it suggests difficulty reaching consensus on what in addition to WoS constitutes the scholarly literature in SSH. In fact, those seeking to extend the definition of scholarly seem influenced by a host of factors not usually considered germane to delineating the scholarly – language and country of publication, size of publisher and age of journal. Figure 2 illustrates this. The x-axis displays a measure of consensus - the number of lists containing a journal - and the y-axis plots number of journals. There is a line for each category of journals. For example one line plots the number of journals published in the US/UK or Netherlands, in English, by large publishers. The five largest journal sets are displayed. The difference between the favoured line (US/UK or Netherlands, English, large publishers) and the rest is dramatic. A larger number of favoured journals are indexed on all lists than are found on no list. For every other journal set, the opposite is true: journals are most likely to be on no list and very few are on all lists.[4] Consensus on the scholarly status of a journal is clearly influenced by language, country and publisher size.

12

These factors are correlated, most obviously language and country. Also, European language journals published outside the UK or Netherlands tend to be published by small publishers while English language journals published in the US, UK or Netherlands tend to be produced by large companies. We can use statistical technique to deepen our understanding of each factor individually by assessing its influence on consensus while controlling for the other factors. To do this we conducted a multivariate logit regression. As before, we defined consensus on scholarly status as being recognized as scholarly by at least two lists. Four lists were considered: WoS, Scopus, ERA and the Norwegian.[5] Thus, consensus on the scholarly status of a journal was the dependent variable, coded as 1 if a journal is indexed by at least 2 lists and 0 otherwise. The independent variables were age of journal (proxied using left two digits of ISSN), publisher size, country of publication and language. The reference group was older, English language journals published by large publishers in US, UK, or Netherlands. The independent variables were coded as follows:

- Small - 1 if the publisher produces 2-14 SSH journals in total, 0 otherwise;
- Tiny - 1 if the publisher produces 1 SSH journal only, 0 otherwise;
- C-European - 1 if the journal is published in a European country (except the UK or Netherlands), 0 otherwise,
- C-Other - 1 if the journal is not published by in the US or Europe, 0 otherwise;
- L-European - 1 if the journal is published in European languages (except English), 0 otherwise;
- L-Other - 1 if the journal is not published in English or other European languages, 0 otherwise.
- Time – 0 if the left two digits of the ISSN are between 0 and 16, 1 if left two digits of the ISSN are between 17 and 21. 37 journals with ISSN beginning 87 are excluded from the analysis.

Table 4 presents results of the logistic regression. The second column shows that each independent variable significantly reduces the chances of achieving consensus that a journal is scholarly when all else is held constant. The strongest effect is seen when the language of a journal is non-European. Journals that are the only journal produced by a publisher or are new have very much reduced chances of being recognized as scholarly. Being published in a European language other than English or being one of a small group of journals published by a small publisher have a smaller, though still substantial effect. Next comes being published outside Europe and finally being published in a European country other than the UK or Netherlands, though the country effect is much weaker than the language effect. The third column expresses the size of the effect as percentage change information. Take "small" as an example, if the publisher changes from large to small, the probability of consensus (being listed on more than one list) drops by 35.2 percentage points, holding other variables fixed. The consensus probability is 76% for the reference group. That is, older English language journals published in the US, UK or Netherlands by large publishers have a 76% chance of being listed on more than one list.

TABLE 4. Multivariate logit analysis of factors reducing consensus on scholarly.

	log odds of Consensus	Consensus Probability Change
Small	-1.53*** (-36.54)	-0.352
Tiny	-2.17*** (-48.99)	-0.495
C-European	-0.45*** (-8.34)	-0.092
C-Other	-0.8*** (-15.84)	-0.173
L-European	-1.47*** (-25.35)	-0.338
L-Other	-2.78*** (-17.02)	-0.596
Time	-1.92*** (-31.67)	-0.443

Note. Observations: 24,569, absolute value of z statistics in brackets, *** Significant at 1%

Discussion

How are we to interpret the findings that a process supposedly based on elite judgements of such an ineffable quality as scholarliness are in fact subject to mundane considerations such as language, country of origin, newness of journal and even worse - size of publishing company? Ideally, such characteristics should be irrelevant to judging the scholarly quality of a journal.

Stephen Cole in his book *Making Science* provides a starting point for the discussion. Cole argues that in deciding what good science is, we base only part of our judgments on our own direct reading and analysis of ideas. To a larger extent, our opinions of what good science is and who has done good work are based on judgments made by other people, especially elites who dominate the evaluation systems within fields (Cole, 1992, p. 195). In line with Cole's insight, we note that the lists of scholarly journals are not the work of any single person; groups are assembled. In fact people needing a list assemble groups of *other* people to create lists, relying on the judgment of others, as Cole notes. ERIH and ERA are assembled by committees of elite scholars, i.e. the very people to whom we delegate judgments of good science. These groups are qualified to assess journals for inclusion.[6] Scopus has convened a group of scholars and librarians to assess journals. WoS is compiled by editors; people outside the scholarly community who are thereby unqualified to render a scholarly judgement themselves. Therefore they rely on indicators of the judgement of groups of scholars such as citations to the journal, its editors and authors.

Peer review signifies the scholarly because peer review is the quintessential process marshalling the "judgement of others." But the peer review criterion cannot be implemented algorithmically, so list construction itself is a process of judging the scholarly, or marshalling the judgement of others. Because groups produce lists, consensus is implied in the

14

production of a valid assessment of scholarly status. When comparing lists, we found a lack of consensus between list-making groups. Since scholarliness should be an invariant quality of the journal, this is a bit puzzling.

Cole provides some insight into how legitimate differences might arise in assessing scholarliness. Relying on the work of Hargens, Cole explains that some fields take a broad approach and accept all work that is not obviously in error so that they do not risk missing potentially important contributions. By accepting all work that is apparently valid, they rely on time to correct errors. Physics takes this approach. The alternative is to reject work unless it is a significant contribution to knowledge, even though that elevates the risk of missing important contributions. Sociology journals in the United States use this principle. If applied to the construction of journal lists, these two fundamentally different orientations will result in longer and shorter lists. Both lists will be judged scholarly by their own communities, but each community would find fault with the other's list. The "longer list" people would see the shorter list (for example, WoS) as incomplete. The "shorter list" people would see the longer list (for example Scopus) as including non-scholarly junk. Thus we see that there is room for legitimate disagreement over the criteria groups use to assess scholarly status.

Although Cole discussed intellectual disagreements on criteria, our results suggest more mundane sources of disagreement. To understand our results we need to introduce the concept of a fragmented academic community. Clearly, everyone is capable of judging only what they are aware of. Factors that inhibit awareness will fragment the academic community and so compromise academic judgement. It is well accepted that specialization fragments awareness. This is why academics only offer judgments within the domain of their expertise. Therefore, ERA and ERIH are built not by one large group but by a collection of subject area specialty groups. Our analysis finds that the level of consensus does vary by specialty. Overall, 49% of academic journals (as defined above) are on a list and 31% are on more than one list. In classical studies, a small field, 62% of the journals are on more than two lists. At the other extreme, 20% of law journals see some level of consensus on their scholarly quality, which makes sense since law is likely organized into what are in effect national subfields.

Our analysis suggests that there are other factors in addition to specialization that fragment the scholarly community – journal age, language, country, and publisher size. That it takes a few years for awareness of journals to spread and for journals to establish their quality is not a surprise. That journals not in English struggle is not a surprise either because many studies have pointed out inadequacies in coverage of non-English language material. Though the tendency in the literature has been to blame the databases, this study finds the same problem in lists constructed by scholars themselves, suggesting that the inadequacies of non-English language journal coverage in databases may originate with fragmentation inside the scholarly community. After all, the databases use metrics based upon scholarly behaviour in their accession process.

Country of publication is a factor perhaps related to deeper issues. Others have noted before that awareness of scholarly works in countries such as Poland is limited (Webster, 1998; Winclawska, 1996). This paper suggests that Europeans are even less aware of work outside Europe and the United States. Glaser also established the continuing existence of differentiated national communities in social sciences, even in an English speaking country,

Australia (Glaser, 2004). Clearly the existence of what are in effect nationally defined SSH subfields fragments the larger SSH community. However, this tends not to be recognized, or rather respected. Rather than constituting nationally defined subfield committees to build journal lists, the tendency is to privilege international, that is English language journals as higher quality.

Publisher size is more influential than country, yet almost unrecognized. Our analysis established that journals from small publishers face severe disadvantages in being recognized as scholarly by the broader community. This means that in practice, though not in theory, the judgments of scholarliness are influenced by market dynamics. Publishers grow large by acquiring and starting journals. Only journals with large and rich markets will be attractive acquisition or start-up targets. Compare social science and humanities research with medical research. Governments spend vastly more money on medical research, in addition lots of firms conduct medical research. Therefore, there are many more places that need subscriptions to medical journals, and institutions have the money to pay for the subscriptions. Because there is money to be made from medical journals, large publishers will buy journals, consolidating the publishing industry. Databases no doubt find it easier to deal with large publishers, who can send them all their meta data electronically and put their journals on-line. Small, obscure publishers, impoverished because they serve the impoverished SSH community, will not have the resources to go electronic and their visibility will suffer.

The government resources invested in medical research extend to information infrastructure. PubMed has long made the medical literature broadly available. PubMed is free at the point of use because the extremely well-funded US National Institutes of Health spends $115 million a year on it (National Library of Medicine, 2010). Somewhat less ostentatious, in physics there is the ArXiv preprint server, again free at the point of use but not free to run. ArXiv requires $400,000 per year, currently supplied by Cornell University library (Cornell University Library, 2010).[7] Social sciences and humanities, because they are relatively impoverished, have not developed such resources. The absence of a PubMed-type infrastructure carries over into less database coverage. Neuhaus et al. demonstrate that Google Scholar replicates the weak SSH coverage found in studies of WoS and Scopus and wonder whether Google Scholar's comparatively weak SSH coverage is "simply the by-product of a preponderance of freely accessible records of scientific and medical research" (C. Neuhaus et al., 2006, p. 138).

Conclusions

Assessing the scholarliness of journals is a step required to build a list of scholarly journals and is a community judgment. As such it requires agreement among a group of people. Each person in the group will be less aware of new journals in a different language, produced by unknown publishers, dramatically reducing the chances for consensus on such a journal's scholarly status. The factors identified here basically fragment the SSH scholarly community. Fragmentation unnecessarily reduces the community size for broad swathes of SSH, which will serve to lower standards and inhibit the development of knowledge. Reducing fragmentation would expand the horizons of scholars and so enhance scholarship as well as aiding efforts to build an SSH evaluation infrastructure. To overcome these problems

and build a well-founded evaluation infrastructure, the fragmentation in the community needs to be reduced.

De-facto de-fragmentation is underway. Government pressure for "international", English language publication and the higher weighting afforded such papers in metrics systems are serving to increase English language publication and decrease national language publication in SSH. Although this neglects the possibly invaluable role national literatures play in SSH scholarship (Hicks, 2004; Li & Flowerdew, 2009), abandonment of national literatures in favour of English language publication will serve to reduce the fragmentation noted here. Similarly, scholars could make more explicit their preference for big publishing houses over small ones and simply abandon journals produced by small players. In this fashion, SSH scholarship could be reshaped to become more integrated.

Alternatives to this vision exist. Technological means to overcome fragmentation have become feasible. A public infrastructure, like PubMed, could overcome fragmentation in the SSH scholarly field. Such an infrastructure would provide full text indexing of SSH journals not indexed in WoS or Scopus. The infrastructure would be expensive to create because it would require finding and interacting with a large number of very small publishers. However, once the flow of incoming material was established, the infrastructure could create clean meta-data, needed by WoS and Scopus. As a relatively large entity, the infrastructure could establish relationships with WoS and Scopus and make it easy for them to index the journals. The infrastructure would provide on-line full text indexing. This would enable articles to be found using Google Scholar and to be roughly translated using Google translate, at no cost to anyone. This findability and accessibility would help integrate the SSH scholarly community around the world. The infrastructure could also financially support the small journals by making it easy to buy an article. The infrastructure could allow viewing of one page at a time, and purchase of a full article at a small charge, which would be returned to the publisher.

Other groups excluded from WoS have used this model. In Latin America there is SciELO, Scientific Electronic Library Online, a federation of electronic journal infrastructures that meet a centrally defined standard of excellence in journal publishing (scielo.org). SciELO's site not only provides access to 250,000 articles from 660 journals, but also offers basic bibliometric statistics. Similarly, in Africa there is African Journals Online (ajol.info) hosting 46,000 articles from 396 peer reviewed journals.

Social science and humanities scholarship are changing. There is interest in reducing fragmentation both from governments keen to see their scholars integrate into an international community and by scholars, such as the group that produced ERIH. This paper argues that explicit attention should be devoted to understanding the fragmentation issue and resources invested in overcoming it in a way that preserves diversity yet facilitates flow of information and knowledge between communities.

Acknowledgements

Notes

1 The four publishers were Alberta Education, Barron's Educational Series, McGraw Hill Contemporary Learning Series, and Princeton Review Publishing.

2 Coverage of SSH literature outside Europe and the Anglo-Saxon countries is abysmal, only strengthening the conclusions we draw by analyzing European literature.

3 We do not have founding year of journal in our data, we use ISSN number to proxy journal age.

4 If we were to plot other journal sets, their shapes would all be the same as the four descending lines in this graph.

5 ERIH was excluded from this analysis because its field coverage differs substantially.

6 Beyond certifying journals as scholarly, the list-making groups of elite scholars were also tasked with stratifying journals into ranks, that is, identifying the premier (super-scholarly?) journals.

7 However, that funding is ending and donations are now being sought.

References

Archambault, É., Vignola-Gagné, É., Côté, G., Lariviere, V., & Gingras, Y. (2006). Benchmarking scientific output in the social sciences and humanities: The limits of existing databases. Scientometrics, 68(3), 329-342.

Burnhill, P., & Tubby-Hille, M. (1994). On measuring the relation between social science research activity and research publication. Research Evaluation, 4(3), 130-152.

Butler, L., & Visser, M. S. (2006). Extending citation analysis to non-source items. Scientometrics, 66(2), 327-343.

Cole, S. (1992). Making science: Between nature and society. Cambridge Massachusetts: Harvard University Press.

Cornell University Library. (2010, January 21). arXiv.org help - arXiv Support FAQ. Cornell University Library. Retrieved July 14, 2010, from http://arxiv.org/help/support/faq

Elsevier. (2010). Scopus Content Coverage Guide: Complete version (No. V9/2010-05-07) (p. 31). Elsevier. Retrieved from http://info.scopus.com/documents/files/scopus-training/resourcelibrary/pdf/sccg0510.pdf

European Science Foundation. (2010, May 5). ERIH - European Reference Index for the Humanities: European Science Foundation. European Science Foundation. Retrieved July 23, 2010, from http://www.esf.org/research-areas/humanities/erih-european-reference-index-for-the-humanities.html

Frandsen, T. F., & Nicolaisen, J. (2008). Intradisciplinary differences in database coverage

and the consequences for bibliometric research. Journal of the American Society for Information Science and Technology, 59(10), 1570-1581.

Gavel, Y., & Iselid, L. (2008). Web of Science and Scopus: A journal title overlap study. Online Information Review, 32(1), 8-21.

Gimenez-Toledo, E., & Roman-Roman, A. (2009). Assessment of humanities and social sciences monographs through their publishers: A review and a study towards a model of evaluation. Research Evaluation, 18(3), 201-213.

Glaser, J. (2004). Why are the most influential books in Australian sociology not necessarily the most highly cited ones? Journal of Sociology, 40(3), 261-282.

Gluck, M. (1990). A review of journal coverage overlap with an extension to the definition of overlap. Journal of the American Society for Information Science, 41(1), 43-60.

Hicks, D. (2004). The four literatures of social science. In H. Moed, W. Glänzel & U. Schmoch (Eds), Handbook of quantitative science and technology studies (pp. 473-496). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Li, Y., & Flowerdew, J. (2009). International engagement versus local commitment: Hong Kong academics in the humanities and social sciences writing for publication. Journal of English for Academic Purposes, 8, 279-293.

de Moya-Anegón, F., Chinchilla-Rodríguez, Z., Vargas-Quesada, B., Corera-Álvarez, E., Muñoz-Fernández, F. J., González-Molina, A., & Herrero-Solana, V. (2007). Coverage analysis of Scopus: A journal metric approach. Scientometrics, 73(1), 53-78.

National Library of Medicine. (2010, February 1). Congressional Justification FY2011 - National Library of Medicine (NLM). United States National Library of Medicine, National Institutes of Health. Retrieved July 14, 2010, from http://www.nlm.nih.gov/about/2011CJ.html#budget_auth

Nederhof, A. J., & Zwaan, R. (1991). Quality judgments of journals as indicators of research performance in the humanities and the social and behavioral sciences. Journal of the American Society for Information Science, 42(5), 332-340.

Neuhaus, C., Neuhaus, E., Asher, A., & Wrede, C. (2006). The depth and breadth of Google Scholar: An empirical study. Portal-Libraries and the Academy, 6(2), 127-142.

Norris, M., & Oppenheim, C. (2007). Comparing alternatives to the Web of Science for coverage of the social sciences' literature. Journal of Informetrics, 1(2), 161-169.

Pestana, A., Gomez, I., Fernandez, M., Zulueta, M., & Mendez, A. (1995). Scientometric evaluation of R&D activities in medium-size institutions: A case study based on the Spanish Scientific Research Council (CSIC). In M. Koening & A. Bookstein (Eds.), The Proceedings of the Fifth International Conference of the International Society for Scientometrics and Informetrics (pp. 425-434). River Forest IL, United States.

Royle, P., & Over, R. (1994). The use of bibliometric indicators to measure the research productivity of Australian academics. Australian Academic & Research Libraries, 25(2), 77-88.

Schoepflin, U. (1992), Problems of representativity in the social sciences citation index. In P. Weingart, R. Sehringer & M. Winterhager (Eds.), Representations of Science and Technology, Proceedings of the International Conference on Science and Technology Indicators (pp. 177-188). Bielefeld, Germany: DSWO Press.

SerialsSolutions. (2010). Frequently Asked Questions. UlrichsWeb. Retrieved July 23, 2010, from http://www.ulrichsweb.com/ulrichsweb/faqs.asp#About_Ulrichs

Sivertsen, G. (2010). A performance indicator based on complete data for the scientific publication output at research institutions. ISSI Newsletter, 6(1), 22-28.

Thomson Reuters. (2010). The Thomson Reuters Journal Selection Process. Thomson Reuters. Retrieved July 23, 2010, from http://thomsonreuters.com/products_services/science/free/essays/journal_selection_process/

Torres-Salinas, D., & Moed, H. F. (2009). Library catalog analysis as a tool in studies of social sciences and humanities: An exploratory study of published book titles in economics. Journal of Informetrics, 3(1), 9-26.

Villagra Rubio, A. (1992). Scientific production of Spanish universities in the fields of social sciences and language. Scientometrics, 24(1), 3-19.

Walters, W. H. (2007). Google Scholar coverage of a multidisciplinary field. Information Processing & Management, 43(4), 1121-1132.

Webster, B. (1998). Polish Sociology Citation Index as an example of usage of national citation indexes in scientometric analysis of social science. Journal of Information Science, 24(1), 19-32.

Winclawska, B. (1996). Polish Sociology Citation Index (principles for creation and the first results). Scientometrics, 35(3), 387-391.