Citation analysis and peer ranking of Australian social science journals

GABY HADDOW

Department of Information Studies, Curtin University of Technology

PAUL GENONI

Department of Information Studies, Curtin University of Technology

Address for correspondence:

GABY HADDOW

Department of Information Studies, Curtin University of Technology

PO Box U1987, Perth, 6845 Western Australia

E-mail: g.haddow@curtin.edu.au

Abstract

Citation analyses were performed for Australian social science journals to determine the differences between data drawn from Web of Science and Scopus. These data were compared with the tier rankings assigned by disciplinary groups to the journals for the purposes of a new research assessment model, Excellence in Research for Australia (ERA), due to be implemented in 2010. In addition, citation-based indicators including an extended journal impact factor, the *h*-index, and a modified journal diffusion factor, were calculated to assess whether subsequent analyses influence the ranking of journals. The findings suggest that the Scopus database provides higher number of citations for more of the journals. However, there appears to be very little association between the assigned tier ranking of journals and their rank derived from citations data. The implications for Australian social science researchers are discussed in relation to the use of citation analysis in the ERA.

KEYWORDS: Citation analysis; Social science journals; Research Assessment; Citation sources; Australia; Journal ranking

Introduction

From 2010 Australian research output will be assessed using a new model, the Excellence in Research for Australia (ERA), developed by the Australian Research Council (ARC). An important component of ERA is a list of journals ranked in four tiers, A*, A, B and C,[1] which will be used to indicate the 'quality' of articles published in those journals. The list of approximately 20,000 titles was created by scholarly academies and disciplinary groups who were asked to submit lists of peer reviewed titles relevant to their field and to assign a tier rank to each. Australian and non-Australian journals across all disciplines are represented in the list. At the time of writing, disciplinary groups are in the process of reviewing the journals listed in their corresponding Field of Research (FoR) code, a numeric indicator for subject areas (AUSTRALIAN BUREAU OF STATISTICS 2008). The ARC provided very few guidelines to direct the initial journal ranking process and the methods used by different disciplinary groups varied widely (GENONI & HADDOW 2009), with some utilising metrics such as citation-based indicators, while others ranked titles solely on peer judgments.

Information currently available suggests that, in *most* discipline areas, journal articles in ranked titles will be subjected to some form of citation analysis. The ARC has stated in this regard "ERA will use the most appropriate citation data supplier for each discipline" (2008, p. 6). However, the Humanities and the Creative Arts cluster of research fields, which was involved in an ERA trial in 2009, has been exempted from citation analysis. This exemption is an encouraging recognition that citation indicators if applied to research output in these fields were unlikely to produce useful data. As will be explored in this paper, however, there are reasons to believe that citation analyses will be equally unhelpful for many social science fields.

This paper reports on a study of the efficacy of citation measures for determining the quality of Australian social science journals. The study achieves this by examining the differences between citations data drawn from Web of Science and Scopus and comparing these findings with the tier rankings assigned for ERA. In addition, a number of citation-based indicators (an extended journal impact factor, the *h*-index, and a modified journal diffusion factor) were calculated to assess whether subsequent analyses influence the relative ranking of journals. The research contributes to the discourse on the value and utility of bibliometric indicators, and has particular relevance to assessment of social science research.

The paper will also address several important and related issues that arise when considering the implications of ERA for Australian social science researchers. Firstly, journals in these fields are generally not well indexed by the main citation sources (Web of Science and Scopus) and this coverage is further reduced for Australian journals. Secondly, the lack of consistency and transparency in the ranking process raises questions of comparability between titles in the tiers. On the assumption that publishing in journals in higher tiers will attract greater reward in the ERA process, any subsequent citation analyses being applied to individual articles may significantly alter results.

Citations data and the assessment of social science research

For many years the Institute for Scientific Information (ISI) citation indexes were the sole source of citations data readily available and relatively easily retrieved. These printed, CD ROM, and online (Web of Science) indexes facilitated the development of increasingly sophisticated bibliometric research methods and measures of research impact. While ISI held a monopoly position in citation index production, its annually calculated Journal Impact Factor (JIF) held a similar role as the foremost indicator of a journal's impact. By association, the JIF has been used as a proxy to indicate the 'value' of a researcher's work within a journal. Competitors to ISI (now owned by Thomson Corporation) emerged in 2004 when Elsevier launched the Scopus database and Google introduced the Google Scholar website. Along with these new sources of citations data, a range of alternatives to the JIF have been proposed, including the widely cited *h*-index (HIRSCH 2005), the *g*-index (EGGHE 2006), the Discounted Cumulated Impact Index (JARVELIN & PERSSON 2008), the Article-Count Impact Factor (MARKPIN et al. 2008), and Journal Diffusion Factors (FRANDSEN 2004; ROWLANDS 2002).

This extraordinary growth in alternative methods of evaluating research impact is in part driven by the ease with which citations data can be accessed and manipulated when it is available in digital form. However, another important impetus is the increased interest (and activity) of governments in tying research funding to assessments of research quality in the higher education sector. The Research Assessment Exercise (RAE) of the UK has been operating in various forms since 1986, using peer review to evaluate research outputs. This "burdensome and costly system" of research assessment will be replaced with a "metrics-based" model after 2008 (HM TREASURY 2006). Metrics are also a major component of the Australian ERA model

(AUSTRALIAN RESEARCH COUNCIL 2008, p. 6). There is no shortage of discussion relating to the relative merits, or otherwise, of using citations as a measure of research impact or quality (see for example, BORNMANN, MUTZ, NEUHAUS, & DANIEL 2008; BUTLER 2008; HADDOW 2008; HAYES 1983; MOED 2005; WARNER 2000), and several studies have been conducted to test this association for RAE results (BUTLER 2006; HOLMES & OPPENHEIM 2001; OPPENHEIM 1995, 1997; OPPENHEIM & SUMMERS 2008), but to date few studies have investigated the use of citations as a measure of research impact in the Australian context.

The proposal that ERA will use citation data from different sources according to discipline raises an important issue, acknowledged by the ERA Indicators Group. If "valid quantitative indicators" are not available, the document proposes peer review may be a more appropriate method of assessing research quality (AUSTRALIAN RESEARCH COUNCIL 2008); a proposal tested in the Humanities and Creative Arts (HCA) cluster trial. A Consultation Paper released in September 2009 (AUSTRALIAN RESEARCH COUNCIL 2009a) suggests the preferred quantitative indicator for all social science fields (in the Social, Behavioural and Economic Sciences cluster) is citation analysis.

At the centre of the argument presented here and discussed below—that is, bibliometric indicators are likely to be inappropriate for the assessment of social science research in Australia—are three issues. The first relates to the problematic usefulness of citation-based measures as an indication of research quality or impact; the second to the differences between the scholarly communication practices of social science researchers and those of science researchers (for whom citation indexes were initially developed); and thirdly, the adequacy of the two main citation indexes in providing citations data for Australian social science journals.

Editors of the excellent Theme Section 'The use and misuse of bibliometric indices in evaluating scholarly performance' in the journal *Ethics in Science and Environmental Politics* (BROWMAN & STERGIOU 2008, p. 1) preface the papers with a 1992 quote from Per Seglen:

> Citations represent a measure of utility rather than of quality – and a limited kind of utility at that.

Seglen's statement epitomises the tension between the attraction of easily obtainable numeric values to assess research and the meaning ascribed to those values. Understandably, however, the relatively effortless and inexpensive process of identifying citations to research articles appeals to policy makers involved in the development of research assessment models. Indeed the extent to which the volume of citations indicates quality has been the subject of much discussion (for example BORNMANN et al. 2008; BOURKE 1994; BUTLER 2008; CAMERON 2005; MOED 2005; RESEARCH EVALUATION & POLICY PROJECT 2005; STEELE, BUTLER, & KINGSLEY 2006; WARNER 2000), and it is the inclination to equate citations with quality in research assessment that leads to assertions such as BROWMAN and STERGIOU's that "the consequences of an uninformed over-reliance on these metrics are insidious" (2008, p. 3).

Equating quality with citations is only one part of the problem. Most bibliometrics scholars would concur that citations reflect the communication behaviour of scholars in a particular field, and may have limited utility when compared to similar data from other fields. Derek DE SOLLA PRICE pioneered the research into disciplinary differences in scholarly communication (1970), finding that scholars in the hard sciences are likely to give more citations in their papers and that these citations were to more recently published works. The conclusion that the time lag between publication and citation was shorter in the hard sciences than it is in other disciplines has been supported by later studies, including EARLE & VICKERY (1969), LINE (1981), and HICKS (1999). Commentators also note that "non-journal publishing is significant in the social sciences" (HICKS 2004, p. 476), whereas in science a much high proportion of publishing is in journals (MOED 2005).

In addition to the types of publications in which social science researchers publish, HICKS (1999) and MOED (2005) also note the tendency for these researchers to publish in journals with a national focus. This is explained by HICKS, who observes that "because social sciences investigate society they are oriented to their social context and are inherently more national" (p. 202). If social science researchers choose to publish (and are more likely to have their articles accepted) in journals that focus on national issues then the availability of citations data for these publications will be limited due to the restricted coverage of the main citation indexes. For example, MOED rates the coverage of humanities and some fields in social science by the Thomson index as 'moderate' with less than 40% of the discipline's citations going to journals indexed by the database (2005, p. 137). This problem is exacerbated in a country such as Australia, where the index's coverage has always been lower than that provided for North America and Western Europe.

Australian social science research and publications

The nature of scholarly communication in the social sciences, the national orientation of social science journals, and the coverage of citation indexes, all suggest that Australian academics working in these fields will not be well served by citation analyses associated with research assessment. While relatively few in number, previous studies that have examined Australian research in the social sciences and humanities support this argument.

One of the first studies to examine citations to Australian science and social science journals found "Australian journals do not rank highly compared to overseas journals on the basis of impact factor or citations received" (ROYLE 1994, p. 170). ROYLE proposed that the national focus of Australian journals, as well as "lower circulation and poorer coverage by major abstracting and indexing services", might explain this conclusion. A related study (ROYLE & OVER 1994) identified the characteristics of journals in which Australian researchers (science and social science) published most frequently. Social scientists predominantly published in Australian journals, whereas science researchers published in a greater number of journals published outside of Australia—the proportion of Australian journals used by social science researchers was 73% compared with 22% for science researchers. The authors also note the different degree of coverage of these journals by ISI sources – 27% of the social science journals and 87% of the science journals were indexed by ISI. Commenting in *Campus Review* in the same year, Paul BOURKE (1994, p. 9) wrote that the use of impact factors as research indicators in the Australian context "would be indefensible in the social sciences and humanities".

The Research Evaluation and Policy Project (REPP) at the Australian National University has contributed greatly to discussion regarding the assessment of Australian research. In a literature review from 2005, REPP described the use of quantitative indicators for social science and humanities as a "thorny issue precisely due to the limitation of indexed database coverage", and that "[T]hese concerns are heightened with respect to the relative international 'periphery' of Australian research in the social sciences and humanities" (RESEARCH EVALUATION & POLICY PROJECT 2005, p 27). REPP Director, Linda Butler, examined the Thomson coverage of articles published by Australian researchers, illustrating the limitations of using the index to collect citations data for humanities and social science (HSS) disciplines in Australia. With the exception of philosophy, economics, and politics and policy articles (with 30-40% coverage), Thomson indexes between 6% and 28% of the articles in the full range of HSS fields—and this from a sample that includes Australian and non-Australian journals (BUTLER & VISSER 2006).

In a study focusing on the citations in five major Australian economics journals, SMYTH noted that only one was indexed in Web of Science (1999). His findings provide evidence of “the importance to Australian economists of a number of economics journals published in Australia which do not rank world wide” (p. 131).

Australian humanities journals were the focus of research undertaken by John EAST, in which a number of indicators—including library holdings, indexing by databases, and citations from Web of Science source journals—were examined (2006). Over a period of ten years, the citations given to Australian history journals by source journals ranged between 0 and 218, which, at most produces an average of around 22 citations each year. Similarly low numbers of citations to Australian humanities and social science journals were found two recent studies (HADDOW 2008; HADDOW & GENONI 2009). These studies identified citations to over 300 Australian journals over a six year period, reflecting the ERA research assessment time frame, and found 84% of the journals attracted less than 50 citations for the entire period. Only 17 (5.5%) of the journals were indexed by Web of Science.

A hypothesis that might be drawn from this previous research and commentary on the nature of Australian HSS publication is that citation based indicators may not be suitable for the purposes of research assessment. However, the alternative, peer ranking of journals, is not without its own set of problems (GENONI & HADDOW 2009).

Excellence in Research for Australia (ERA)

There is very little known about how different disciplinary groups ranked journals for ERA. With scant direction from the ARC other than an approximate number of titles to rank, the percentage of journals to comprise each tier, and an outline of the characteristics of a journal in each tier (GENONI & HADDOW 2009), the ranking processes remain somewhat opaque. Information has been found on websites and in articles for several disciplinary groups, including education, library and information science, and computer science, which indicates the approach taken differs widely in order to reflect the nature of the field. For example, the education sector conducted a large online survey and asked participants to rank titles according to importance to the professional and academic communities, in separate lists. The final ranking was achieved through combining the survey results with impact factor calculations. (Unfortunately, the webpage providing details of this process is no longer available). In library and information science ranking was performed by academics in the field through peer judgments only (SMITH & MIDDLETON 2009).

With so little information about the ranking processes, but indications that they varied greatly, it is impossible to know the degree of equivalence between journals assigned to the same tier from different disciplinary groups. There is a real possibility that Australian social science journals, valued by those undertaking the journal assessment due to the national focus typical of the disciplines, have been assigned to an unjustifiably high tier. Citations to these journals may not be useful as indicators of quality, but they may provide a sense of equivalence in quantitative terms. It is this aspect of journal ranking and the intention of the ARC to use "the most appropriate citation data supplier for each discipline" that was the impetus for this study.

Methods

Australian humanities and social science journals were identified by searching *Ulrich's Periodicals Directory* using limits that ensured only refereed (scholarly), active journals (from at least 2001) with 'Australia' listed as place of publication were located. Irregular publications and those with a subject focus outside of social science were excluded. A number of additional titles that were not retrieved in the first *Ulrich's* search were located by searching for titles containing the term 'Australia*'. A total of 244 journals comprised the first sample for analysis after excluding titles that were not listed as an ERA ranked journal. Using *Ulrich's* in the first instance ensured that all Australian titles with the potential to be classed as humanities and social science were included in the sample. As will be seen in the results, not all of the titles located in *Ulrich's* came within the social science cluster.

Citations data are now available from two major subscription databases, Web of Science and Scopus, as well as other databases and websites, notably Google Scholar. A number of studies have been undertaken with a view to comparing these sources in terms of their coverage and functionality (see for example BAKKALBASI, BAUER, GLOVER, & WANG 2006; BAR-ILAN 2008; BAUER & BAKKALBASI 2005; BOSMAN, VAN MOURIK, RASCH, SIEVERTS, & VERHOEFF 2006; FALAGAS, PITSOUNI, MALIETZIS, & PAPPAS

2008; GAVEL & ISELID 2008; GENONI & HADDOW 2009; P. JACSO 2005; MEHO & YANG 2007; NORRIS & OPPENHEIM 2007; VAUGHAN & SHAW 2008). The types of samples included in the studies vary widely and hinders any overall comparison between the sources, with the exception of findings for Google Scholar. JACSO (2008) is particularly critical of the reliability of citation data retrieved from Google Scholar, which in most cases retrieved many more citations than the other sources. JACSO's concerns about using Google Scholar are supported by others, who discuss the appropriateness of using potentially unreliable data in subsequent analyses or to assess research activity (FALAGAS et al. 2008; NORRIS & OPPENHEIM 2007; A. G. SMITH 2008). Web of Science and Scopus, however, are more difficult to separate in terms of preferred source. It would appear from the previous research that neither source is best for all citation needs and that their usefulness is dependent upon subject areas and the age of publications.

The three sources, Web of Science, Scopus and Google Scholar, were searched for citations to the Australian journals. The different functionalities of the sources required a range of approaches to locate citations data, including; the 'cited reference search' in Web of Science; citation tracking and 'more' tab searches in Scopus; and using the Publish or Perish website to retrieve Google Scholar citations. These searches identified all citations given between 2001 and 2007 to articles published in the years 2001 to 2006. This period was selected to reflect the 6 year period for which research outputs are being assessed by ERA. In addition, the citing period (2001-2007) supports the notion that the time lag between publication and citation is longer in the humanities and social science than science disciplines, for which the 2 year publication period and one year citing period of the impact factor was designed. An earlier study (HADDOW 2008) demonstrated that citation-based analyses, such as the journal diffusion factor, applied to journals with fewer than 50 citations produced anomalous results, and therefore only titles with 50 or more citations were subjected to further analysis. From the 244 titles only 44 were cited more than 50 times (in both Web of Science and Scopus) over the period.

The intention of the study was to analyse Australian humanities and social science journals. However, only two journals in the sample of 44 with more than 50 citations could be categorised as humanities journals. This, combined with the ARC's decision to exclude citation analysis from the Humanities and Creative Arts cluster of research fields, resulted in the focus of this research being on Australian social science titles.

Results

Web of Science, Scopus and Google Scholar as citation sources

The raw citation counts found for the three citation sources supported previous studies' conclusions in relation to Google Scholar. Only three of the Australian journals had a lower number of citations in Google Scholar (using the Publish or Perish software) than the other citation sources. One other journal retrieved no citations in the Google Scholar search. These results are perplexing, possibly reflecting the problems discussed above in relation to Google Scholar. For the remaining 40 titles, the citations found in Google Scholar were on average 2.7 times more than the number of citations found in Web of Science or Scopus (the highest value was used as the denominator). Due to the considerable difference between the citations found in Google Scholar and the other sources and the reservations about the reliability of these data, no further analyses were conducted using results from Google Scholar.

Overall, Scopus retrieved a higher number of citations for more journals than Web of Science—24 titles compared with 19 respectively (one title had an equal number of citations). The difference between the citations found by the two sources for each title was calculated to determine the extent of difference. On average, Scopus had 67.2 more citations across the 24 titles compared with Web of Science which had 40.6 citations more than Scopus for the set of 19 titles. Table 1 displays this data, the median number of citations difference, and range of difference. The high standard deviation for the Scopus titles suggests that the median difference in citations is possibly a better measure with which to compare the sources. A further comparison made between the sources was a calculation of the ratio between the citations located, expressed as a positive number for the source in which the higher number of citations were identified. Scopus had an average ratio of 1.45 compared with 1.26 for Web of Science.

TABLE 1

The range of difference between the citations located in the sources for each title varied greatly (see Table 1), particularly for the Scopus titles. In order to explore whether a few titles with very high citations in Scopus had skewed the findings, the differences between citations in each source were coded in ranges of difference: 1-10, 11-20, 21-50, 51-100, and >100 citations difference (again expressed as a positive number for the source in which the higher number of citations was found). Figure 1 illustrates the results of this analysis, indicating that Scopus not only ‘out-performs’ Web of Science in terms of the number of titles for which higher citations counts were found, but was also responsible for a higher number of titles with greater difference in the number of citations found.

FIGURE 1

Earlier studies have found that Web of Science and Scopus perform differently according to subject area. To test whether this is true of the Australian journals, the titles assigned the same two digit FoR code were examined, resulting in:

FoR 13: Education	8 titles
FoR 14: Economics	7 titles
FoR 15: Commerce, Management, Tourism & Services	5 titles
FoR 16: Studies in Human Society	11 titles
FoR 21: History & Archaeology	5 titles
FoR 04: Earth Sciences	3 titles

The first four of these FoR groups are co-located in the Social, Behavioural and Economic Sciences cluster, the FoR 21 group is in the Humanities and Creative Arts cluster, and the last group (04) is part of the Physical, Chemical and Earth Sciences cluster. Seven titles were assigned FoR codes that differed from all other titles in the sample and are therefore not included in this analysis. Two titles were assigned both the 16 and 04 FoR codes and are included in both groups. Table 2 presents the number of titles with a higher number of citations in either Scopus or Web of Science by FoR group.

TABLE 2

From these results, Scopus would be the preferred citation source for the first three FoR groups. Web of Science provides higher citations for more titles in the Studies in Human Society set, however the difference between Web of Science and Scopus for four of these titles is less than 50 citations. Scopus, on the other hand, was found to have more than 50 more citations difference for four of the titles for which it recorded a higher number of citations. It is notable that Web of Science appears to perform better than Scopus for the FoR 21 titles, but as part of the Humanities and Creative Arts cluster in ERA these titles will not be subject to citation analyses.

Similarly, although Web of Science performed better for the FoR 04 set, the Physics, Chemistry and Earth Sciences cluster in which it is located underwent an ERA trial in 2009 with Scopus selected as the citation source.

Citations and peer ranking

As noted previously, journals were ranked in four tiers—A*, A, B, and C—for ERA purposes, and this process was carried out by disciplinary groups using a range of methods, including peer review and metrics. To explore if tier rank was reflected in the citations a title received, the sample was organised into the four tiers and a range and mean citations were calculated for Scopus and Web of Science. Table 3 presents the findings of this analysis and indicates there is no association between tier rank and citations for the titles based on the mean. Interestingly, the mean citations in Scopus and Web of Science for the A* (particularly) and A titles are similar, while the means for the C are widely divergent.

TABLE 3

Citation-based indicators, citations and peer ranking

Three further calculations were carried out on the citations found in both sources for each title; an *h*-index value, an extended impact factor, and a modified diffusion factor. The *h*-index is calculated automatically for titles by Scopus, however for Web of Science citations the calculation was conducted manually by listing citations to a journal's articles and sorting from highest to lowest to find the *n*th article with *n* or more citations. An extended impact factor was created to allow for longer citation lag time in the social sciences and also to reflect the ERA assessment period of six years. Extended impact factors were calculated for all titles and both sources using the following equation:

$$\frac{\text{Number of citations (2001- 2007) to journal articles (2001-2006)}}{\text{Number of articles published in journal (2001-2006)}}$$

Frandsen's New Journal Diffusion factor (FRANDSEN 2004) was modified for the same reasons as described for the extended impact factor, using a six year publication period and seven year citation lag time. The following equation expresses the calculation for the modified diffusion factor:

$$\frac{\text{Number of different citing journals (2001- 2007) to journal articles (2001-2006)}}{\text{Number of articles published in journal (2001-2006)}}$$

For these analyses, the source responsible for the higher number of citations was compared with the source which produced higher values for the *h*-index, impact factor and diffusion factor. In the earlier results, Scopus was found to have the higher number of citations for 24 of the 44 titles, however only eleven of these titles were found to have the highest values across all indicators for Scopus data. None of the 19 Web of Science titles achieved the same. The title with equal number of citations in both sources produced varied results for the other indicators. That is, the *h*-index value was higher using Web of Science data, the diffusion factor was higher using Scopus data, and impact factor was identical to six decimal places. Supporting previous concerns about the reliability of subsequent analysis of citation data, the analyses found 24 of the titles had at least one higher indicator derived from different source data. For example, a title may have higher citations in Scopus, but a higher diffusion factor using the Web of Science data.

Across the three indicators, the impact factor emerged as the metric most closely associated with citations. That is, the 43 titles with higher citations from one or other of the sources also had the higher impact factor value for the same source. This finding is hardly surprising given the equation used to calculate the impact factor.

The *h*-index value is always expressed in whole numbers, leading to less differentiation between the values. There were 17 titles (39% of the sample) with higher citations and higher *h*-value using the same source data. Scopus data was responsible for the majority of these (11 titles). A strong association was found between higher citations and the diffusion factor using the Scopus data. All titles with higher citations in Scopus produced a higher diffusion factor.

The patterns that emerge are surprising when the indicators (calculated as a mean for all titles in the tier) were examined in relation to the tier rank (see Table 4). The mean impact factor for both sources generally reflects the tier ranking, with Web of Science data producing a more consistent trend. Web of Science data calculated for the *h*-index and diffusion factor also results in a closer match with the tier ranking, although the means for tier A and B titles are inverted for both indicators. Scopus, on the other hand, produces mean *h*-index values and diffusion factors which bear almost no relationship to the tier ranks.

TABLE 4

The four social science groups (FoRs 13, 14, 15, and 16) were analysed separately to determine if the ERA tier ranking within a sub-disciplinary group (and potentially by the same peer group) was reflected in the total citations and the three indicators. Tables 5-8 present these analyses for each of the FoRs. Although losing the finer detail of differences between the analyses results, the FoRs are presented in rank order to make the results easier to read. For example, a title with an impact value of 0.6613 is ranked higher than a title with an impact value of 0.3925. Note, due to the expression of the *h*-index value as a whole number and the potential for titles within the FoR groups to have the same *h*-index result, the ranking in the *h*-index column does not always include every number rank in the sequence.

TABLE 5

TABLE 6

TABLE 7

TABLE 8

When the ranking results of the sources are compared, there is some consistency evident in the FoR 13 and FoR 14 titles. However, a great deal of variation can be seen within the citation indicator rankings using a single citation source and also the rankings across citation sources. In general, the ERA tier ranks bear very little relation to the citation-based ranks. Only one A* title in the FoR 13 group is ranked highly using citation data, while a tier B journal would appear to deserve a higher tier rank based on the citation-based indicators. The citation indicator ranks for the two tier A journals in the FoR 14 group differ widely, suggesting that citations were not a factor in determining ERA tier rank by the disciplinary group. A similar result is seen for the five titles in the FoR 15 group. In FoR 16 the different citation sources produce variation in the citation indicator rankings, with almost no agreement between the sources' ranking order, and no association evident between the ERA tier rank and citation indicator rank. Remarkably, the results from the citation analyses generally agree on the only A* title, ranking it *lowest* of the 11 journals.

Discussion and conclusions

There are two important findings from this study that support the argument that citations data may not be the most appropriate method of assessing research output in Australian social science journals. Firstly, a relatively low percentage of these titles attract sufficient citations to make such an assessment meaningful on the article level. Of the 244 titles originally identified only 44 (18%) had attracted 50 or more citations over the seven year period. From these 44 titles, six titles attracted less than 100 citations (in Scopus and Web of Science) in that time. This equates to an average of around 14 citations per year for all articles in these journals, and suggests a large proportion of articles within the journals attract no citations. This observation is associated with the second peripheral finding, that the number of citations found for the titles may be associated (although this wasn't tested in the study) with the indexing coverage of the two citation sources.

The ERA will apply citation analysis to the Social, Behavioural and Economic Sciences (SBE) cluster, having noted that citation analysis will be used 'for those disciplines where at least half of the total output of the discipline (including non-journal articles) is indexed by the citation information supplier" (AUSTRALIAN RESEARCH COUNCIL 2009b, p. 5). With this in mind, the indexing of Scopus and Web of Science for the 31 titles included in the SBE cluster was examined. Scopus (just) met the Australian Research Council's (ARC) criteria by fully indexing 51.6% of the titles compared with 35.4% by Web of Science. A closer analysis of the indexing, however, revealed that Scopus achieved the ARC's 50% standard for titles in only two sub-groups, for the FoRs 14 and 16 FoR (57% and 91% respectively). Indexing of the FoR 13 and 15 titles was at best 25% for the period 2001-2006. Web of Science fully indexed 82% of the FoR 16 titles, but otherwise did not meet the 50% benchmark set by the ARC.

The variations in indexing found for titles in the SBE cluster, and the possibly related low number of citations, have important implications for Australian social science researchers who publish in national journals. As the journals in this study's sample were drawn directly from lists created by researchers in social science disciplines, it must be assumed that they are valued national scholarly communication outlets. Yet it would appear that these same researchers will find much of their published output performs poorly when the citation analysis indicator adopted by ERA is applied.

Google Scholar citations for the Australian journals were on average 2.7 times the number found in Web of Science or Scopus; findings that support previous research comparing citation sources (FALAGAS et al. 2008;

JACSO 2008; NORRIS & OPPENHEIM 2007; A. G. SMITH 2008). Reiterating the conclusions drawn by these various authors, Google Scholar clearly requires a cautious approach to reaching any conclusions based on its data.

Across the sample of 44 Australian titles, Scopus would appear to be the preferred citation source. A larger proportion of the titles was found to have higher numbers of citations in Scopus and the difference between the sources was also greater for Scopus. These results are repeated when the titles are gathered into their FoR codes within the Social, Behavioural and Economic Sciences cluster. However, Web of Science is marginally better than Scopus for the sub-group Studies in Human Society in that same cluster. Web of Science also found higher numbers of citations to more titles (three compared to two in Scopus) in the History and Archaeology sub-group; a finding of academic interest only as the cluster to which this sub-group belongs will not be subjected to citation analysis in ERA.

The tier ranking assigned to titles in the sample was conducted by at least four, but probably more, different disciplinary groups. For example, the Centre for the Study of Research Training and Impact at the University of Newcastle coordinated the ranking of education titles for the Australian Association for Research in Education. Titles relevant to the field of librarianship and information science were ranked by members of the Australian Library and Information Association. Other organisations listed on the ARC web page as contributing to the process, and associated with the fields of research included in the sample, are the Australian Academy of the Humanities, the Economic Society of Australia, and the Academy of the Social Sciences in Australia. As discussed above, very few details are available about the ranking processes, therefore leaving unanswered the question about equivalence between journals assigned the same tier within different FoRs. On the basis of citations, the descriptive statistics calculated for this study (mean citations per title) indicate no association between the tier rank and citations. However, the variation found for these analyses means that the results cannot be presented as concrete evidence in this regard.

When the titles were subjected to further citation analyses - the extended journal impact factor, the *h*-index, and a modified journal diffusion factor – the strongest association was found for impact factor and source data. That is, the citation source with higher numbers of citations will also produce a higher impact factor. Scopus produced the most consistent results when subsequent analyses were conducted, with a diffusion factor, as well as an impact factor, mirroring the findings for raw citation numbers. However, it is Web of Science data that reflects the tier ranking of titles most closely when the citation analyses are calculated. On the assumption that journals will be weighted according to tier in the ERA process, the choice of citation provider and any subsequent citation analyses being applied to individual articles will potentially alter results. For example, the selection of Scopus, a decision that would be supported by the findings for raw citation numbers in this study, will reduce the effects of weighting if further analyses are applied. In addition, the rankings that resulted from the citations and further analyses for titles within FoR groups demonstrate that the assignment of tier ranking has little, if any, relationship with citations, regardless of source.

It is important to note, as in most bibliometrics research, that the identification of citations, particularly when conducted manually using the subscription databases, is challenging and some degree of human error may occur. In addition, the different functions of the two major sources mean that the methods for identifying citations also differ. Each of these factors place limitations on the researcher's degree of confidence in relation to arriving at unqualified conclusions.

An acknowledged aspect of social science research is the importance of national focus, which, in terms of the ERA journal ranking exercise creates difficulties. If Australian researchers are to be assessed using international benchmarks, then their publications should be found in the most important journals in the field, whether international or national. In the context of overall research outputs, as submitted to ERA, the findings for indexing coverage of Australian journals may not significantly affect the results of an individual's research assessment. However, the findings do point to potential problems for individuals who have published extensively in the national journals as many attract low citation numbers.

The differences apparent in the analyses reiterate an earlier comment about the likelihood that some Australian social science journals have been assigned a relatively high rank due to the national focus typical of the disciplines. If these rankings are accepted in the final ERA model being implemented in 2010, and positive weighting is applied to higher tiers, then the number of citations will have less impact on the outcome, somewhat evening out the calculations for the citations indicator.

In general, Scopus appears to be the better citation source for the social science journals, but this is by no means a consistent finding across all the Australian titles. There is no doubt that whichever citation source is selected for the social science cluster in ERA, some journals will be negatively affected. This degree of variation could have important implications for individuals and research groups. While recognising that any research assessment model will have its shortcomings, the findings of this study indicate that applying citation analysis to the research outputs of social science researchers in Australia is not a reliable or appropriate method to determine quality.

---

[1] Typically an A* journal would be one of the best in its field or subfield in which to publish and would typically cover the entire field/subfield. Virtually all papers they publish will be of a very high quality. These are journals where most of the work is important (it will really shape the field) and where researchers boast about getting accepted. Acceptance rates would typically be low and the editorial board would be dominated by field leaders, including many from top institutions.

The majority of papers in a Tier A journal will be of very high quality. Publishing in an A journal would enhance the author's standing, showing they have real engagement with the global research community and that they have something to say about problems of some significance. Typical signs of an A journal are lowish acceptance rates and an editorial board which includes a reasonable fraction of well known researchers from top institutions.

Tier B covers journals with a solid, though not outstanding, reputation. Generally, in a Tier B journal, one would expect only a few papers of very high quality. They are often important outlets for the work of PhD students and early career researchers. Typical examples would be regional journals with high acceptance rates, and editorial boards that have few leading researchers from top international institutions.

Tier C includes quality, peer reviewed, journals that do not meet the criteria of the higher tiers.

Exact text from: *Tiers for the Australian Ranking of Journals*: http://www.arc.gov.au/era/tiers_ranking.htm (accessed February 8 2010)

References

AUSTRALIAN BUREAU OF STATISTICS (2008). *Australian and New Zealand Standard Research Classification*. Retrieved September 16 2009, from http://www.abs.gov.au/Ausstats/abs@.nsf/Latestproducts/6BB427AB9696C225CA2574180004463E?opendocument

AUSTRALIAN RESEARCH COUNCIL (2008). *ERA Indicator Principles*. Retrieved 21 December 2008, from http://www.arc.gov.au/era/indicators.htm

AUSTRALIAN RESEARCH COUNCIL (2009a). *Draft ERA submission guidelines: Physical, Chemical and Earth Sciences (PCE) and Humanities and Creative Arts (HCA) clusters*. Retrieved 29 January 2009, from http://www.arc.gov.au/pdf/Draft_ERA_Sub_Guide.pdf

AUSTRALIAN RESEARCH COUNCIL (2009b). *ERA Indicators Consultation Paper*. Retrieved October 4 2009, from http://www.arc.gov.au/pdf/ERA_indicators_consult.pdf

BAKKALBASI, N., BAUER, K., GLOVER, J., & WANG, L. (2006). Three options for citation tracking: Google Scholar, Scopus and Web of Science. *Biomedical Digital Libraries, 3*(7).

BAR-ILAN, J. (2008). Which h-index? - a comparison of WoS, Scopus and Google Scholar. *Scientometrics, 74*(2), 257-271.

BAUER, K., & BAKKALBASI, N. (2005). An examination of citation counts in a new scholarly communication environment. *D-Lib Magazine, 11*(9), [np].

BORNMANN, L., MUTZ, R., NEUHAUS, C., & DANIEL, H.-D. (2008). Citation counts for research evaluation: Standards of good practice for analyzing bibliometric data and presenting and interpreting results. *Ethics in Science and Environmental Politics, 8*, 93-102.

BOSMAN, J., VAN MOURIK, I., RASCH, M., SIEVERTS, E., & VERHOEFF, H. (2006). *Scopus Reviewed and Compared: The Coverage and Functionality of the Citation Database Scopus, Including Comparisons with Web of Science and Google Scholar*. Utrecht: Utrecht University Library, http://igitur-archive.library.uu.nl/DARLIN/2006-1220-200432/UUindex.html.

BOURKE, P. (1994, 19-25 May). Quantitative research indicators: Citations treatment 'deficient'. *Campus Review*, p. 9.

BROWMAN, H.I., & STERGIOU, K.I. (2008). Factors and indices are one thing, deciding who is scholarly, why they are scholarly, and the relative value of their scholarship is something else entirely. *Ethics in Science and Environmental Politics, 8*, 1-3.

BUTLER, L. (2006). *RQF Pilot Study Project: History and Political Science: Methodology for Citation Analysis*: REPP. Retrieved 8 September 2009,

BUTLER, L. (2008). Using a balanced approach to bibliometrics: Quantitative performance measures in the Australian Research Quality Framework. *Ethics in Science and Environmental Politics, 8*, 83-92.

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

CAMERON, B.D. (2005). Trends in the usage of ISI bibliometric data: Uses, abuses, and implications. *Portal: Libraries and the Academy, 5*(1), 105-125.

DE SOLLA PRICE, D.J. (1970). Citation measures of hard science, soft science, technology, and nonscience. In C.E. NELSON & D.K. POLLOCK (Eds.), *Communication Among Scientists and Engineers* (pp. 3-22). Lexington, MA: Heath Lexington Books.

EARLE, P., & VICKERY, B. (1969). Social science literature use in the UK as indicated by citations. *Journal of Documentation, 25*, 123-141.

EAST, J.W. (2006). Ranking journals in the humanities: An Australian case study. *AARL: Australian Academic & Research Libraries, 37*(1), 3-16.

EGGHE, L. (2006). Theory and practise of the g-index. *Scientometrics, 69*(1), 131-152.

FALAGAS, M.E., PITSOUNI, E.I., MALIETZIS, G.A., & PAPPAS, G. (2008). Comparison of PubMed, Scopus, Web of Science, and Google Scholar: strengths and weaknesses. *Faseb Journal, 22*(2), 338-342.

FRANDSEN, T.F. (2004). Journal diffusion factors: A measure of diffusion? *Aslib Proceedings, 56*(1), 5-11.

GAVEL, Y., & ISELID, L. (2008). Web of Science and Scopus: a journal title overlap study. *Online Information Review, 32*(1), 8-21.

GENONI, P., & HADDOW, G. (2009). ERA and the ranking of Australian humanities journals. *Australian Humanities Review, 46*(May), 7-26.

HADDOW, G. (2008). Quality Australian journals in the humanities and social sciences. *AARL: Australian Academic and Research Libraries, 39*(2), 79-91.

HADDOW, G., & GENONI, P. (2009). Australian education journals: Quantitative and qualitative indicators. *AARL: Australian Academic & Research Libraries, 40*(2), 88-104.

HAYES, R.M. (1983). Citation statistics as a measure of faculty research productivity. *Journal of Education for Librarianship, 23*(3), 151-172.

HICKS, D. (1999). The difficulty of achieving full coverage of international social science literature and the bibliometric consequences. *Scientometrics, 44*(2), 193-215.

HICKS, D. (2004). The four literatures of social science. In H.F. MOED, W. GLÄNZEL & U. SCHMOCH (Eds.), *Handbook of quantitative science and technology research : The use of publication and patent statistics in studies of S&T systems* (pp. 473-496). Dordrecht Kluwer Academic Publishers.

HIRSCH, J.E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences, 102,* 16569-16572. Retrieved 16 June 2007, from

HM TREASURY (2006). *Government Meeting Science Goals*. Retrieved 28 December 2008, from http://www.hm-treasury.gov.uk/press_53_06.htm

HOLMES, A., & OPPENHEIM, C. (2001). Use of citation analysis to predict the outcome of the 2001 Research Assessment Exercise for Unit of Assessment (UoA) 61: Library and Information Management. *Information Research, 6*(2).

JACSO, P. (2005). As we may search - Comparison of major features of the Web of Science, Scopus, and Google Scholar citation-based and citation-enhanced databases. *Current Science, 89*(9), 1537-1547.

JACSO, P. (2008). The pros and cons of computing the h-index using Google Scholar. *Online Information Review, 32*(3), 437-452.

JARVELIN, K., & PERSSON, O. (2008). The DCI index: discounted cumulated impact-based research evaluation. *Journal of the American Society for Information Science and Technology, 59*(9), 1433-1440.

LINE, M.B. (1981). The structure of social science literature as shown by a large-scale citation analysis. *Social Science Information Studies, 1,* 67-87.

MARKPIN, T., BOONRADSAMEE, B., RUKSINSUT, K., YOCHAI, W., PREMKAMOLNETR, N., RATCHATAHIRUN, P., et al. (2008). Article-count impact factor of materials science journals in SCI database. *Scientometrics, 75*(2), 251-261.

MEHO, L.I., & YANG, K. (2007). Impact of data sources on citation counts and rankings of LIS faculty: Web of Science versus Scopus and Google Scholar. *Journal of the American Society for Information Science and Technology, 58*(13), 2105-2125.

MOED, H.F. (2005). *Citation analysis in research evaluation*. Dordrecht: Springer.

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.

OPPENHEIM, C. (1995). The correlation between citation counts and the 1992 Research Assessment Exercise ratings for British library and information science university departments. *Journal of Documentation, 51*(1), 18-27.

OPPENHEIM, C. (1997). The correlation between citation counts and the 1992 Research Assessment Exercise ratings for British research in Genetics, Anatomy and Archaeology. *Journal of Documentation, 53*(5), 477-487.

OPPENHEIM, C., & SUMMERS, M.A.C. (2008). Citation counts and the Research Assessment Exercise, part IV: Unit of assessment 67 (music). *Information Research, 13*(2), paper 342.

RESEARCH EVALUATION & POLICY PROJECT. (2005). *Quantitative Indicators for Research Assessment: A Literature Review* Canberra: Research Evaluation and Policy Project, Australian National University. Retrieved 17 March 2007,

ROWLANDS, I. (2002). Journal diffusion factors: A new approach to measuring research influence. *Aslib Proceedings, 54*(2), 77-84.

ROYLE, P. (1994). A citation analysis of Australian science and social science journals. *Australian Academic and Research Libraries, 25*(3), 162-171.

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

SMITH, A.G. (2008). Benchmarking Google Scholar with the New Zealand PBRF research assessment exercise. *Scientometrics, 74*(2), 309-316.

SMITH, K., & MIDDLETON, M. (2009). Australian library and information studies (LIS) researchers ranking of LIS journals. *AARL: Australian Academic & Research Libraries, 40*(4), 1-20.

SMYTH, R. (1999). A citation analysis of Australian economic journals. *AARL: Australian Academic & Research Libraries, 30*(2), 119-133.

STEELE, C., BUTLER, L., & KINGSLEY, D. (2006). The publishing imperative: The pervasive influence of publication metrics. *Learned Publishing, 19*(4), 277-290.

VAUGHAN, L., & SHAW, D. (2008). A new look at evidence of scholarly citation in citation indexes and from web sources. *Scientometrics, 74*(2), 317-330.

WARNER, J. (2000). A critical review of the application of citation studies to the Research Assessment Exercise. *Journal of Information Science, 26*(6), 453-460.

Figure 1: Titles and range of difference between citations found in sources

Table 1: Mean and median for the difference between citations found for sources

Source	**Titles (n)**	**Mean**	**Median**	**Range**	**SD**
Higher cites in Scopus	24	67.2	53	3-370	76.9
Higher cites in Web of Science	19	40.6	29	7-116	29.8

Table 2: Source with higher citations for titles grouped by Field of Research code

FoR Code		**Number of titles**	
		Higher citations in Scopus	Higher citations in Web of Science
13	Education	6	2
14	Economics*	5	1
15	Commerce, Management, Tourism S.	4	1
16	Studies in Human Society	5	6
21	History & Archaeology	2	3
04	Earth Sciences	1	2

*One title had equal citations in the sources

Table 3: Citations to titles by ERA rank: Range and mean

ERA rank	**Titles (n)**	**Scopus**		**Web of Science**	
		Range of citations	Mean citations per title	Range of citations	Mean citations per title
A*	9	63-481	191.9[1]	88-450	191.8[2]
A	18	54-463	208.7[3]	54-541	203.5[4]
B	13	54-438	232.5[5]	79-472	201.2[6]
C	4	66-527	210.5[7]	93-157	125.5[8]

(Standard deviation: [1] 144.2; [2] 118.6; [3] 110.2; [4] 128.5; [5] 105.3; [6] 117.9; [7] 212.8; [8] 33.7)

Table 4: Citation-based indicators (mean) compared with ERA rank

ERA rank	***h*-index**		**Impact factor**		**Diffusion factor**	
	Scopus	WoS	Scopus	WoS	Scopus	WoS
A*	5.8	6.4	1.64	1.64	0.64	0.44
A	6	5.8	1.4	1.38	0.54	0.37
B	6	5.9	1.4	1.2	0.65	0.38
C	5	5	1.3	0.75	0.41	0.25

Table 5: FoR 13 (Education) titles and rank order by ERA rank, citations and indicators

ERA rank	Scopus				Web of Science			
	Total citations	*h*-index	Impact factor	Diff. factor	Total citations	*h*-index	Impact factor	Diff. factor
A*	**7**	5	7	6	**7**	6	7	6
A*	**5**	2	6	5	**4**	4	4	4
A*	**1**	1	1	2	**1**	1	1	1
A	**3**	2	3	4	**2**	2	2	3
A	**8**	2	8	7	**8**	8	8	7
B	**2**	7	2	1	**3**	3	3	2
B	**4**	8	4	3	**5**	6	6	5
C	**6**	5	5	8	**6**	5	5	8

Table 6: FoR 14 (Economics) titles and rank order by ERA rank, citations and indicators

ERA rank	Scopus				Web of Science			
	Total citations	*h*-index	Impact factor	Diff. factor	Total citations	*h*-index	Impact factor	Diff. factor
A*	**2**	1	1	3	**3**	1	1	4
A	**7**	7	7	6	**7**	7	7	6
A	**1**	1	2	1	**1**	2	2	1
B	**4**	3	5	4	**4**	4	4	3
B	**3**	4	4	5	**2**	2	3	5
B	**6**	6	6	7	**6**	6	6	7
B	**5**	4	3	2	**5**	4	5	2

Table 7: FoR 15 (Commerce, Management, Tourism & Services) titles and rank order by ERA rank, citations and indicators

ERA rank	Scopus				Web of Science			
	Total citations	*h*-index	Impact factor	Diff. factor	Total citations	*h*-index	Impact factor	Diff. factor
A	**1**	3	1	2	**2**	2	2	5
A	**4**	5	3	3	**5**	2	5	3
A	**2**	1	4	4	**1**	1	1	4
B	**3**	1	2	1	**3**	2	3	1
B	**5**	3	5	5	**4**	2	4	2

Table 8: FoR 16 (Studies in Human Society) titles and rank order by ERA rank, citations and indicators

ERA rank	**Scopus**				**Web of Science**			
	Total citations	*h*-index	Impact factor	Diff. factor	Total citations	*h*-index	Impact factor	Diff. factor
A*	**11**	10	11	10	**11**	11	10	7
A	**10**	5	9	11	**8**	5	7	5
A	**8**	5	4	4	**6**	5	2	2
A	**6**	5	8	9	**5**	5	6	10
A	**4**	2	6	5	**2**	2	3	4
A	**9**	5	10	8	**10**	10	11	8
B	**2**	2	3	7	**3**	2	5	9
B	**7**	4	7	3	**4**	5	4	3
B	**3**	1	2	1	**1**	1	1	1
B	**5**	10	5	6	**7**	5	8	6
C	**1**	5	1	2	**9**	4	9	11