Journal of Informetrics 4 (2010) 554–563

Contents lists available at ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi

# Ranking of library and information science researchers: Comparison of data sources for correlating citation data, and expert judgments

Jiang Li[a], Mark Sanderson[b,*], Peter Willett[b], Michael Norris[c], Charles Oppenheim[c]

[a] *Department of Information Management, Nanjing University, China*
[b] *Department of Information Studies, University of Sheffield, UK*
[c] *Department of Information Science, Loughborough University, UK*

ARTICLE INFO

*Article history:*
Received 14 April 2010
Received in revised form 16 June 2010
Accepted 16 June 2010

*Keywords:*
Expert judgments
*g*-index
*h*-index
*H*-index
Library and information science
Peer review

ABSTRACT

This paper studies the correlations between peer review and citation indicators when evaluating research quality in library and information science (LIS). Forty-two LIS experts provided judgments on a 5-point scale of the quality of research published by 101 scholars; the median rankings resulting from these judgments were then correlated with *h*-, *g*- and *H*-index values computed using three different sources of citation data: Web of Science (WoS), Scopus and Google Scholar (GS). The two variants of the basic *h*-index correlated more strongly with peer judgment than did the *h*-index itself; citation data from Scopus was more strongly correlated with the expert judgments than was data from GS, which in turn was more strongly correlated than data from WoS; correlations from a carefully cleaned version of GS data were little different from those obtained using swiftly gathered GS data; the indices from the citation databases resulted in broadly similar rankings of the LIS academics; GS disadvantaged researchers in bibliometrics compared to the other two citation database while WoS disadvantaged researchers in the more technical aspects of information retrieval; and experts from the UK and other European countries rated UK academics with higher scores than did experts from the USA.

## 1. Introduction

Pressures on the funding of higher education have resulted in an increasing focus on the development of quantitative criteria for the evaluation of research quality. In particular, the use of citation data to measure the impact of individual academics, departments, publication forums, and disciplines has seen increased interest. Such a use of citation data can be validated by showing that it correlates well with peer judgments (the classic approach to assessment in academia). The growth of interest in citation has mirrored the growth in provision of databases that enable such evaluations to take place. Large-scale services such as Web of Science (WoS) and Scopus commonly form the basis of citation studies. There has been an increased focus on the use of Google Scholar (GS) to provide citation information, however, there is concern that there is a great deal of noise in GS data.

A previous study investigated the extent to which peer judgments and citation indicators from WoS correlated with a panel of 101 of the world's leading researchers in library and information science (LIS). In this paper, we extend that study to compare correlations with Scopus and GS. The next section reviews previous work in the field. We then describe the peer

---

[*] Corresponding author. Fax: +44 114 2780300.
*E-mail address:* m.sanderson@sheffield.ac.uk (M. Sanderson).

1751-1577/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.joi.2010.06.005

judgments and citation indicators that we used, taking particular note of the three different sources of citation data that were employed, before presenting and discussing our results.

## 2. Correlating expert judgments and citation data

There are at least three variables that must be considered when reviewing the many studies that have sought to correlate peer review with citation data. The first is the source of the peer review data. In the case of RAE-like studies,[1] the judgments are provided by panels of subject experts. The second is the granularity of the study, i.e., the nature of the individual items that are being assessed, with studies reported that range across a spectrum stretching from the research outputs of entire nations down to those of specific researchers. The third is the source of the citation data: this has traditionally been the Web of Science (WoS), but the availability of the Scopus and Google Scholar databases is now providing alternative bases for the evaluation of research performance. All three of these variables have been considered in the many previous studies that have demonstrated the existence of meaningful correlations.

Patterson and Harris (2009) found a low but statistically significant correlation between the judgments of referees for the journal *Physics in Medicine and Biology* and the subsequent citations to papers published in that journal. Bornmann and Daniel (2006) found that papers published by successful candidates for an important post-doctoral fellowship attracted more citations than articles by unsuccessful candidates. Korevaar and Moed (1996) found that ratings of publications and journals by a group of expert mathematicians correlated well with citation counts, although the degree of correlation was strongest for the most highly ranked items. Studies in the Netherlands showed a fair correlation between bibliometric indicators and assessors' judgments of research programmes in condensed matter physics (Rinia, van Leeuwen, van Vuren, & van Raan, 1998) and chemistry (van Raan, 2006). Reale, Barbara, and Costantini (2007) found significant correlations between peer assessments of articles in chemistry, biology, the humanities and economics submitted for VTR (the Italian equivalent of the RAE) and the impact factors of the journals where those articles appeared. Abramo, D'Angelo, and Caprasecca (2009) found significant correlations between expert assessments of institutions and the impact factors of the journals where those institutions published for eight subject areas in VTR. There have also been studies where the extent of the correlation is much less marked: Nicol, Henadeera, and Burtler (2007) found only poor correlations between a range of bibliometric indicators and reviewer assessments of grant applications to the National Health and Medical Research Council, as did Aksnes and Taxt (2004) in a study of 34 research groups at the University of Bergen. Maier (2006) found no significant positive correlation between peer assessment of journals in regional science and the impact factors of those journals; indeed, where there was a significant correlation, it was often negative.

In this paper, we consider the extent of the correlation between peer judgments and citation counts in LIS. It is noteworthy that two of the very first such correlation studies involved LIS, with both Oppenheim (1995) and Seng and Willett (1995) correlating citation counts for UK LIS departments with their gradings from the 1992 RAE. Both of these reports used WoS data but studies are now appearing in the literature (Bar-Ilan, 2008; Jacso, 2005) that compare WoS with Scopus and GS, with some of these studies focusing on LIS. Thus, Bauer and Bakkalbasi (2005) compared citation counts provided by WoS, Scopus and GS for articles from the *Journal of the American Society for Information Science and Technology* published in 1985 and in 2000. They found that WoS provided the highest citation counts for the 1985 articles and GS provided significant higher citation counts than either WoS or Scopus for the 2000 articles. Meho and Yang (2007) compared the three databases when applied to the ranking of LIS faculty at Indiana University in Bloomington. They found that, compared to WoS, Scopus significantly altered the position of middle-ranked academics and that GS provided better coverage of conference proceedings and non-English materials. They also noted that comprehensive analyses based on GS could be extremely time-consuming when compared to the carefully edited data in WoS and Scopus; Sanderson (2008), however, pointed out that this need not be a problem if just the most cited articles are required for each of the authors under investigation.

The work reported here draws on several previous studies (some of which have already been referred to above) that have used the *h*-index (Alonso, Cabrerizo, Herrera-Viedma, & Herrera, 2009) to assess and compare the citation counts of LIS academics. Within a year of the publication of Hirsch's seminal paper (Hirsch, 2005), Cronin and Meho (2006) used the index to rank influential information scientists in the USA and found that it provided a more nuanced ranking than straight citation counts; this work was rapidly complemented by an analogous study of UK researchers (Oppenheim, 2007). The Cronin-Meho and Oppenheim studies had both used WoS as the data source. Sanderson (2008) noted that certain types of computationally oriented LIS research were under-represented (principally because WoS did not then cover the conference proceedings that provide a primary research outlet in computer science) and that GS provided a rather different picture of individuals' research impacts. Norris and Oppenheim (2010) have recently reported a study in which the WoS-derived *h*-indices and *g*-indices of leading LIS researchers from around the world were correlated with human judgments of these individuals provided by an expert panel of academics and journal editors. In the present paper, we extend the analysis of Norris and Oppenheim by comparing their WoS rankings with those obtained by using Scopus and GS to provide the citation counts that are to be correlated with the expert judgments.

---

[1] The Research Assessment Exercise (RAE) is carried out on a regular basis to direct central-government funding to those higher education institutions that demonstrate the highest levels of research excellence. Such evaluations have been carried out in the United Kingdom since 1984 with analogous exercises run in Australia, Italy, the Netherlands and New Zealand.

**Table 1**
Scale used by the 42 members of the expert panel to score the quality of the research produced by the 101 LIS academics.

| Score | Meaning |
|---|---|
| 4 | Work that is world-leading in terms of originality, significance and rigour |
| 3 | Work that is internationally excellent in terms of originality, significance and rigour but which nonetheless falls short of the highest standards of excellence |
| 2 | Work that is recognised internationally in terms of originality, significance and rigour |
| 1 | Work that is recognised nationally in terms of originality, significance and rigour |
| 0 | Work that does not merit national recognition |

## 3. Methods

The set of researchers and expert judgments used here was that developed in the study by Norris and Oppenheim (2010). The generation of the data is spelled out in detail in their article, and we hence provide only a brief summary. A total of 101 active LIS researchers were identified for analysis from two sources. The first was the individuals discussed in the studies of Cronin and Meho (2006), Oppenheim (2007) and Sanderson (2008) (*vide supra*). The second was a random selection of individuals who had written 15 or more articles that had been published in LIS journals listed in the Thomson Reuters Journal Citation Reports (JCR) database and that had appeared in the period 1998–2008. A web-based questionnaire was then designed to obtain peer judgments of these 101 researchers using the 5-point scale shown in Table 1. The levels of excellence listed here are based on those used in the most recent RAE in the UK, the outcomes of which were announced at the end of 2008 (at http://www.rae.ac.uk/news/2008/results.asp). After piloting of the questionnaire, 58 people were invited to provide judgments of the quality of the publications of these 101 researchers: 44 of these were drawn at random from the 101 researchers and the remaining 14 were the editors of LIS journals in the JCR database. In all, 42 people provided such judgments, an overall success rate of 72.4%. Each expert chose the researchers that they felt competent to evaluate (excluding themselves if they were on the list), and the median ranking was then computed for each researcher.

We have extended Norris and Oppenheim's basic data by categorising each of the 101 researchers in terms of their broad subject area and of their geographical location. Five very broad categories of LIS research were used, with each researcher being allocated to that single group that best reflected their principal areas of publication. These categories were BIB (Bibliometrics), IM (Information management), IRH (Information retrieval from the harder, computer-science end of the subject), IRS (Information retrieval from the softer end of the subject), and SOC (Social aspects of information science), with these five groups containing 24, 26, 17, 18 and 16 researchers, respectively. Three geographical categories were used. These were British (B), American (A) and Other (O), with these three groups containing 39, 31 and 31 researchers, respectively. The 42 experts comprising the peer review panel were categorised in the same way, with 17 B, 10 A and 15 O experts. Each of the 101 researchers hence had four peer review ratings, namely the median of the values assigned by the complete set of experts and by the experts from each of the three geographical groups. These ratings are listed in the left-hand part of Table 2 where it will be seen, e.g., that Judith Bar-Ilan's ratings were 2 (for the whole panel, denoted by W), 4, 2.5 and 2 (for the B, A and O experts).

Three citation indices were employed here: the *h*-index (Alonso et al., 2009; Hirsch, 2005), the *g*-index (Egghe, 2006; Schreiber, 2010) and the *H*-index (Randić, 2009). The *h*-index and the *g*-index are well known and were used in the previous study by Norris and Oppenheim (2010). The *h*-index was defined by Hirsch (2005) as follows: "*A scientist has index h if h of his/her $N_p$ papers have at least h citations each and other papers have no more than h citations each*. . ." while Egghe (2006) defined the g-index as: ". . . *an improvement of the h-index of Hirsch to measure the global citation performance of a set of articles. If [a set of articles] is ranked in decreasing order of the number of citations that they received, the g-index is the (unique) largest number such that the top g articles received (together) at least $g^2$ citations*." The *H*-index was introduced by Randić (2009) and seeks to describe the distribution of citation frequencies for the Hirsch core, i.e., those *h* publications that contribute to the *h*-index; in this way, the *H*-index provides a simple way of discriminating between individuals with the same *h*-index. Its calculation, which is quite complex, was described by Randić (2009) using a worked example; more recently Egghe (in press) provided a detailed description of the method, which involved ". . . *replacing the h-index of an author by a decreasing sequence of numbers* $A = (x_1, x_2, \ldots, x_N)$ *where* $x_1 = h$, *the h-index of this author and where the other values* $x_i$ ($i = 2, \ldots, N$) *are other ranks derived from the fact that papers in the h-core usually have a number of citations that is much higher than h*".

The WoS, Scopus and GS citation counts were obtained for each of the 101 researchers as follows. The WoS *AuthorFinder* facility was used to identify each of the publications in WoS, and the *h*-index, *g*-index and *H*-index for the resulting citations calculated via a Visual Basic Application in Excel (with analogous index calculations being carried out on the Scopus and GS citation data). In like vein, the *AuthorIdentifier* facility was used to obtain the corresponding citation data from Scopus, while the GS citations were searched using the popular *Publish or Perish* program (which is available for free download from http://www.harzing.com/pop.htm). The index values that were calculated are denoted in the following by *x*_*Y*, where *x* records the index type (*h*-index, *g*-index or *H*-index) and *X* denotes the data source (WoS for Web of Science, SCO for Scopus, and GS for Google Scholar); for example, *h*_SCO denotes an *h*-index value calculated using the Scopus database. Finally, some additional searches were carried out on WoS using a publication time limit of 1996–2008: the resulting index values are denoted by *x*_W96. In this way, 12 different index values were computed for each of the 101 researchers. The extent of

**Table 2**
Median rankings (W = Whole World, B = British, A = American, O = Other) of 101 LIS researchers and corresponding citation indices (see text for details of the indices).

| Author | W | B | A | O | *h*_GS | *g*_GS | *H*_GS | *h*_SCO | *g*_SCO | *H*_SCO | *h*_WoS | *g*_WoS | *H*_WoS | *h*_W96 | *g*_W96 | *H*_W96 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Bar-Ilan, J. | 2 | 4 | 2.5 | 2 | 22 | 35 | 59 | 14 | 22 | 37 | 12 | 18 | 30 | 11 | 16 | 26 |
| Bates, M.J. | 3 | 3 | 3 | 3 | 26 | 61 | 88 | 13 | 29 | 43 | 16 | 32 | 52 | 7 | 14 | 23 |
| Bawden, D. | 2 | 2 | 2 | 2 | 16 | 30 | 47 | 10 | 19 | 30 | 8 | 12 | 19 | 5 | 9 | 13 |
| Beheshti, J. | 2 | 2 | 2 | 2 | 16 | 27 | 46 | 9 | 16 | 27 | 10 | 17 | 30 | 7 | 13 | 23 |
| Belkin, N.J. | 4 | 4 | 3 | 4 | 33 | 75 | 116 | 19 | 33 | 57 | 17 | 39 | 63 | 4 | 9 | 14 |
| Bertot, J.C. | 3 | 0 | 2 | 4 | 15 | 22 | 38 | 9 | 13 | 20 | 7 | 11 | 18 | 6 | 8 | 14 |
| Borgman, C.L. | 3 | 3 | 3 | 3 | 25 | 53 | 82 | 12 | 25 | 41 | 18 | 34 | 58 | 7 | 17 | 26 |
| Brophy, P. | 2 | 2 | 0 | 2 | 12 | 19 | 31 | 5 | 7 | 11 | 4 | 6 | 10 | 4 | 4 | 8 |
| Buckland, M.K. | 3 | 4 | 3 | 2 | 20 | 44 | 62 | 5 | 8 | 12 | 8 | 17 | 24 | 4 | 8 | 12 |
| Budd, J.M. | 1 | 1 | 2 | 1 | 17 | 24 | 42 | 10 | 14 | 24 | 9 | 14 | 24 | 6 | 10 | 16 |
| Burrell, Q.L. | 2 | 2 | 3 | 2 | 16 | 23 | 40 | 10 | 13 | 23 | 13 | 19 | 33 | 6 | 8 | 12 |
| Case, D.O. | 2 | 2 | 2 | 2 | 13 | 29 | 43 | 7 | 13 | 21 | 10 | 17 | 30 | 4 | 8 | 13 |
| Cole, C. | 2 | 2.5 | 1.5 | 2 | 15 | 20 | 35 | 9 | 13 | 22 | 11 | 17 | 29 | 9 | 10 | 19 |
| Cronin, B. | 3 | 3 | 3 | 4 | 27 | 42 | 71 | 14 | 23 | 38 | 16 | 25 | 43 | 12 | 19 | 32 |
| Damodaran, L. | 1 | 1 | 0 | 0 | 10 | 22 | 33 | 4 | 11 | 16 | 4 | 10 | 15 | 4 | 6 | 14 |
| Davenport, E. | 2.5 | 3 | 2 | 2.5 | 15 | 24 | 40 | 7 | 10 | 16 | 5 | 9 | 13 | 5 | 9 | 13 |
| De Moya-Anegon, F. | 2 | 2.5 | 2 | 1.5 | 15 | 21 | 35 | 10 | 13 | 23 | 8 | 11 | 19 | 6 | 8 | 13 |
| Debackere, K. | 1 | 0 | 0 | 1.5 | 19 | 31 | 53 | 12 | 18 | 30 | 11 | 16 | 27 | 7 | 11 | 18 |
| Dilevko, J. | 1 | 0 | 0 | 2 | 7 | 10 | 17 | 4 | 7 | 12 | 4 | 6 | 10 | 4 | 6 | 9 |
| Dillon, A. | 2 | 2.5 | 2 | 3 | 26 | 51 | 86 | 11 | 24 | 38 | 9 | 19 | 30 | 8 | 16 | 26 |
| Egghe, L. | 3 | 3 | 3 | 3 | 22 | 39 | 63 | 14 | 23 | 37 | 17 | 25 | 44 | 10 | 16 | 27 |
| Ellis, D. | 2 | 2 | 3 | 2 | 22 | 42 | 67 | 13 | 20 | 36 | 16 | 29 | 48 | 9 | 15 | 26 |
| Enser, P.G.B. | 2 | 1 | 2 | 2 | 10 | 27 | 39 | 5 | 14 | 20 | 5 | 13 | 18 | 3 | 7 | 10 |
| Feather, J. | 2 | 2 | 0 | 2 | 11 | 24 | 33 | 2 | 3 | 4 | 3 | 4 | 7 | 2 | 2 | 4 |
| Fidel, R. | 2 | 0.5 | 2 | 2 | 22 | 40 | 69 | 10 | 20 | 31 | 15 | 26 | 47 | 4 | 6 | 17 |
| Foo, S. | 2 | 2.5 | 2 | 2 | 12 | 24 | 38 | 9 | 14 | 24 | 9 | 14 | 24 | 7 | 10 | 16 |
| Ford, N. | 2 | 3 | 2 | 2 | 23 | 36 | 64 | 16 | 23 | 40 | 18 | 23 | 43 | 13 | 18 | 33 |
| Garg, K.C. | 0 | 0 | 0 | 0 | 7 | 11 | 17 | 7 | 9 | 16 | 10 | 13 | 23 | 7 | 8 | 14 |
| Gibb, F. | 1.5 | 2 | 0 | 1 | 12 | 16 | 30 | 8 | 10 | 19 | 6 | 10 | 15 | 4 | 5 | 9 |
| Glanzel, W. | 3 | 2 | 3 | 4 | 30 | 45 | 81 | 22 | 32 | 56 | 25 | 35 | 62 | 17 | 23 | 39 |
| Goker, A. | 1 | 1 | 1 | 2 | 8 | 15 | 22 | 2 | 7 | 9 | 2 | 7 | 9 | 2 | 6 | 8 |
| Gunter, B. | 1 | 1 | 0 | 3 | 29 | 45 | 79 | 10 | 12 | 22 | 12 | 19 | 32 | 4 | 5 | 8 |
| Gupta, B.M. | 2 | 0 | 0 | 2 | 8 | 10 | 17 | 6 | 7 | 11 | 5 | 6 | 11 | 5 | 6 | 11 |
| Harnad, S. | 2 | 2 | 2 | 3 | 43 | 85 | 131 | 15 | 32 | 44 | 6 | 14 | 24 | 3 | 5 | 9 |
| Hernon, P. | 2 | 2 | 2 | 2 | 19 | 33 | 55 | 8 | 12 | 20 | 9 | 11 | 19 | 6 | 8 | 13 |
| Hjorland, B. | 2 | 2 | 2 | 2 | 19 | 37 | 59 | 14 | 21 | 37 | 13 | 20 | 35 | 11 | 17 | 30 |
| Huntington, P. | 2.5 | 2.5 | 0 | 0 | 17 | 24 | 43 | 13 | 17 | 31 | 11 | 15 | 26 | 10 | 14 | 24 |
| Ingwersen, P. | 3 | 3 | 3 | 3 | 26 | 59 | 89 | 15 | 32 | 51 | 15 | 33 | 51 | 10 | 26 | 40 |
| Jacso, P. | 2 | 3 | 1 | 2 | 10 | 21 | 31 | 7 | 14 | 22 | 7 | 11 | 17 | 5 | 10 | 15 |
| Jamali, H.R. | 0 | 0 | 0 | 0 | 7 | 10 | 17 | 7 | 9 | 14 | 6 | 8 | 12 | 5 | 6 | 10 |
| Jansen, B.J. | 1.5 | 1 | 1 | 2 | 27 | 64 | 94 | 18 | 39 | 60 | 12 | 27 | 37 | 9 | 23 | 31 |
| Jarvelin, K. | 3 | 3.5 | 2 | 4 | 23 | 48 | 74 | 15 | 28 | 46 | 12 | 22 | 33 | 9 | 13 | 22 |
| Jiang, J.J. | 0 | 0 | 0 | 0 | 20 | 31 | 54 | 14 | 21 | 35 | 10 | 16 | 27 | 9 | 13 | 21 |
| Jose, J.M. | 2 | 2 | 0 | 2 | 13 | 25 | 42 | 8 | 14 | 23 | 6 | 9 | 16 | 4 | 7 | 11 |
| Kantor, P.B. | 2 | 2 | 2 | 2 | 19 | 39 | 60 | 12 | 23 | 38 | 7 | 14 | 25 | 5 | 9 | 15 |
| Koenig, M.E.D. | 2 | 2 | 1.5 | 2 | 10 | 14 | 23 | 4 | 6 | 11 | 9 | 14 | 22 | 3 | 5 | 9 |
| Kretschmer, H. | 1 | 0.5 | 1.5 | 2 | 9 | 15 | 26 | 7 | 10 | 18 | 8 | 12 | 21 | 6 | 8 | 12 |
| Kuhlthau, C.C. | 3 | 3.5 | 2.5 | 3.5 | 22 | 53 | 73 | 7 | 14 | 22 | 8 | 24 | 32 | 3 | 7 | 11 |
| Lalmas, M. | 2 | 2 | 1.5 | 3 | 26 | 40 | 68 | 11 | 18 | 30 | 7 | 15 | 25 | 6 | 11 | 19 |
| Large, A. | 2 | 2.5 | 2.5 | 2 | 17 | 30 | 53 | 11 | 18 | 31 | 11 | 17 | 31 | 8 | 14 | 24 |
| Larson, R.R. | 2 | 0.5 | 2 | 3 | 14 | 31 | 46 | 6 | 11 | 17 | 7 | 14 | 22 | 4 | 6 | 9 |
| Lewison, G. | 2 | 1 | 0 | 2 | 14 | 18 | 35 | 12 | 16 | 28 | 10 | 14 | 24 | 9 | 12 | 22 |
| Leydesdorff, L. | 3 | 2.5 | 3 | 3 | 37 | 73 | 111 | 21 | 35 | 55 | 17 | 28 | 45 | 13 | 23 | 35 |
| Liddy, E.D. | 2.5 | 1 | 3 | 2.5 | 20 | 30 | 52 | 6 | 9 | 15 | 6 | 9 | 15 | 3 | 3 | 5 |
| Losee, R.M. | 2 | 0 | 2 | 2 | 17 | 29 | 49 | 10 | 14 | 24 | 11 | 16 | 28 | 6 | 9 | 14 |
| McCain, K. | 3 | 2 | 3 | 3 | 20 | 44 | 66 | 11 | 22 | 32 | 17 | 34 | 54 | 7 | 16 | 25 |
| McClure, C.R. | 2 | 3 | 2 | 3 | 22 | 38 | 64 | 10 | 13 | 23 | 10 | 12 | 21 | 6 | 8 | 13 |
| McKnight, C. | 2 | 1.5 | 3 | 3 | 18 | 35 | 58 | 8 | 11 | 18 | 7 | 10 | 17 | 5 | 7 | 12 |
| Marchionini, G. | 3 | 3 | 3 | 3 | 33 | 69 | 104 | 15 | 23 | 40 | 15 | 26 | 45 | 7 | 11 | 18 |
| Marcella, R. | 1.5 | 1.5 | 0 | 0 | 9 | 14 | 24 | 5 | 6 | 10 | 5 | 7 | 12 | 4 | 6 | 9 |
| Meyer, M. | 2.5 | 0 | 0 | 3 | 11 | 23 | 39 | 12 | 19 | 34 | 12 | 18 | 34 | 11 | 16 | 30 |
| Moed, H.F. | 3 | 2 | 2 | 3 | 29 | 47 | 82 | 19 | 30 | 52 | 21 | 34 | 59 | 13 | 20 | 33 |
| Morris, A. | 1.5 | 1 | 0 | 2 | 6 | 9 | 14 | 5 | 10 | 14 | 6 | 10 | 15 | 4 | 8 | 11 |
| Nicholas, D. | 2 | 3 | 1.5 | 2 | 19 | 27 | 47 | 13 | 18 | 31 | 12 | 15 | 28 | 11 | 15 | 25 |
| Oppenheim, C. | 3 | 3 | 2 | 3 | 22 | 35 | 60 | 14 | 21 | 37 | 10 | 14 | 22 | 9 | 13 | 20 |
| Ounis, I. | 3 | 1.5 | 0 | 3 | 16 | 29 | 49 | 7 | 9 | 15 | 4 | 6 | 10 | 3 | 4 | 6 |
| Raper, J.F. | 2 | 2 | 2 | 2 | 16 | 29 | 47 | 6 | 12 | 17 | 3 | 5 | 8 | 2 | 3 | 4 |
| Van Rijsbergen, C.J. | 4 | 3 | 4 | 3 | 35 | 65 | 111 | 15 | 25 | 46 | 6 | 13 | 20 | 6 | 12 | 19 |

Table 2 (*Continued* )

| Author | W | B | A | O | *h*_GS | *g*_GS | *H*_GS | *h*_SCO | *g*_SCO | *H*_SCO | *h*_WoS | *g*_WoS | *H*_WoS | *h*_W96 | *g*_W96 | *H*_W96 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Robertson, S. | 4 | 4 | 4 | 3 | 11 | 28 | 40 | 13 | 32 | 46 | 7 | 14 | 24 | 7 | 12 | 20 |
| Rousseau, R. | 3 | 2 | 3 | 3 | 23 | 43 | 66 | 17 | 28 | 46 | 18 | 25 | 43 | 14 | 19 | 33 |
| Rowland, F. | 1 | 1 | 0 | 1.5 | 10 | 20 | 33 | 5 | 8 | 13 | 4 | 6 | 10 | 3 | 5 | 9 |
| Rowlands, I. | 3 | 3 | 3 | 1 | 11 | 18 | 30 | 10 | 12 | 20 | 7 | 9 | 15 | 6 | 8 | 13 |
| Rowley, J. | 1 | 1 | 2 | 2.5 | 22 | 39 | 63 | 10 | 13 | 22 | 7 | 9 | 15 | 5 | 6 | 11 |
| Ruger, S. | 2 | 2 | 2 | 0 | 10 | 17 | 30 | 7 | 10 | 16 | 6 | 9 | 16 | 4 | 6 | 10 |
| Ruthven, I. | 2 | 2.5 | 2 | 2.5 | 19 | 32 | 55 | 10 | 16 | 28 | 7 | 11 | 19 | 6 | 10 | 16 |
| Sanderson, M. | 2 | 1 | 2 | 3 | 23 | 46 | 73 | 10 | 16 | 26 | 6 | 9 | 14 | 3 | 4 | 6 |
| Saracevic, T. | 3 | 4 | 3.5 | 3 | 32 | 72 | 111 | 17 | 42 | 63 | 18 | 39 | 60 | 11 | 16 | 37 |
| Savolainen, R. | 3 | 3 | 2.5 | 2 | 13 | 27 | 41 | 9 | 16 | 27 | 10 | 19 | 31 | 8 | 11 | 20 |
| Schamber, L. | 2 | 0 | 1.5 | 2 | 11 | 30 | 40 | 6 | 15 | 20 | 7 | 21 | 27 | 4 | 8 | 12 |
| Schubert, A. | 3 | 1 | 2 | 3 | 21 | 36 | 60 | 19 | 33 | 53 | 22 | 36 | 57 | 11 | 16 | 28 |
| Smith, L.C. | 2 | 2 | 2 | 2 | 11 | 25 | 35 | 5 | 10 | 13 | 6 | 15 | 19 | 3 | 3 | 6 |
| Soergel, D. | 2 | 1 | 2 | 2 | 19 | 38 | 62 | 8 | 15 | 25 | 7 | 15 | 22 | 4 | 8 | 12 |
| Spink, A. | 2 | 3 | 2 | 3 | 39 | 72 | 114 | 25 | 48 | 74 | 22 | 40 | 61 | 18 | 33 | 50 |
| Tait, J. | 1 | 0.5 | 0 | 2 | 12 | 22 | 35 | 5 | 7 | 12 | 3 | 4 | 6 | 2 | 2 | 4 |
| Tenopir, C. | 3 | 3.5 | 2 | 3 | 24 | 40 | 69 | 14 | 22 | 37 | 11 | 19 | 31 | 9 | 16 | 28 |
| Thelwall, M. | 3 | 3 | 3 | 3 | 30 | 49 | 87 | 23 | 35 | 62 | 20 | 30 | 51 | 17 | 26 | 46 |
| Vakkari, P. | 2.5 | 3 | 2 | 3 | 16 | 30 | 49 | 12 | 21 | 36 | 11 | 20 | 34 | 9 | 16 | 27 |
| Van House, N. | 2 | 2 | 2 | 2 | 16 | 32 | 53 | 5 | 9 | 15 | 2 | 3 | 8 | 2 | 3 | 8 |
| Van Leeuwen, T.N. | 2 | 0 | 1 | 3 | 18 | 32 | 55 | 14 | 24 | 42 | 12 | 18 | 33 | 12 | 17 | 31 |
| Van Raan, A.F.J. | 4 | 4 | 3.5 | 4 | 32 | 45 | 82 | 19 | 28 | 48 | 15 | 23 | 41 | 14 | 22 | 38 |
| Vaughan, L. | 2 | 0 | 1 | 2 | 16 | 29 | 48 | 12 | 22 | 36 | 11 | 19 | 34 | 10 | 16 | 28 |
| Warner, J. | 1 | 1 | 1 | 2 | 10 | 15 | 26 | 6 | 10 | 14 | 6 | 12 | 16 | 6 | 9 | 14 |
| White, H.D. | 3 | 1.5 | 3 | 3 | 19 | 44 | 62 | 14 | 28 | 45 | 16 | 34 | 51 | 9 | 19 | 29 |
| Whittaker, S.J. | 2 | 2 | 2 | 2 | 39 | 77 | 126 | 15 | 26 | 43 | 5 | 7 | 11 | 3 | 5 | 8 |
| Wildemuth, B.M. | 3 | 0 | 2 | 2 | 16 | 24 | 41 | 6 | 10 | 17 | 9 | 13 | 23 | 5 | 8 | 12 |
| Willett, P. | 4 | 4 | 3 | 4 | 50 | 92 | 144 | 45 | 80 | 121 | 46 | 82 | 125 | 30 | 59 | 84 |
| Williams, P. | 0 | 0 | 0 | 0 | 14 | 19 | 33 | 10 | 14 | 23 | 10 | 19 | 29 | 8 | 11 | 19 |
| Wilson, C.S. | 2 | 1 | 2.5 | 2.5 | 13 | 19 | 33 | 9 | 13 | 23 | 7 | 12 | 20 | 6 | 10 | 16 |
| Wilson, T.D. | 3 | 3 | 3 | 3 | 26 | 54 | 81 | 16 | 31 | 48 | 15 | 29 | 44 | 9 | 19 | 30 |
| Yang, C.C. | 3 | 2 | 0 | 3 | 13 | 23 | 38 | 13 | 30 | 45 | 7 | 12 | 19 | 7 | 11 | 18 |
| Zitt, M. | 2 | 1 | 1.5 | 3 | 15 | 21 | 36 | 9 | 13 | 24 | 8 | 12 | 20 | 8 | 11 | 18 |

the relationship between these index values and the median expert judgments was illustrated using scatter diagrams, and quantified using the Spearman rank correlation coefficient (Siegel & Castellan, 1988).

## 4. Results and discussion

The median rankings and the computed index values are listed in Table 2. It will be seen that the values derived from WoS and Scopus are broadly comparable, but noticeably less than those from GS. For example, the *h*_GS values of the 101 researchers are, on average, 1.7 times the *h*_SCO values and 1.8 times the *h*_WoS values. The ranges of *h*-, *g*- and *H*-indices for WoS are [2, 46], [3, 82] and [6, 125], respectively; those for Scopus are [2, 45], [3, 80] and [4, 121]; while those for GS are [6, 50], [9, 92] and [14, 144]. These figures demonstrate the much larger numbers of citations that can be expected for most researchers when the GS database is used.

When scatter diagrams are plotted of the median rankings against each of the median rankings, a marked, positive relationship between the median ranking and the chosen index is obtained in all 12 cases. This is exemplified by the two scatter diagrams in Fig. 1: these are based on the W rankings for *h*_W96 and for *g*_SCO. The correlation coefficients for the 12 scatter diagrams are listed in Table 3, where it will be seen that they range from 0.388 for *h*_W96 to 0.552 for *g*_SCO and *H*_SCO (i.e., the two diagrams in Fig. 1 are those with the lowest and highest correlation coefficients). Hardly surprisingly, all of the W96 values are lower than the corresponding values for the other three indices since they are based on less (and for some of the senior researchers, very much less) citation data. The magnitudes of the correlations in Table 3 are comparable to those for *h*_WoS and g_WoS listed by Norris and Oppenheim (whose figures relate only to citations for the period 1996–2008). It will be seen that for all three indices, the Scopus values are higher than the corresponding WoS and GS values, i.e., the index values computed from the Scopus database give higher correlations with the expert panel than do the values computed from the other two databases.

**Table 3**
Spearman correlation coefficients between median ranking and citation indices of the 101 researchers. All of the coefficients are significant at the 0.001 level (2-tailed test) of statistical significance.

| Index/database | GS | GS (cleaned) | SCO | WoS | W96 |
|---|---|---|---|---|---|
| *h* | 0.497 | 0.502 | 0.513 | 0.456 | 0.388 |
| *g* | 0.529 | 0.538 | 0.552 | 0.492 | 0.474 |
| *H* | 0.524 | 0.529 | 0.552 | 0.496 | 0.461 |

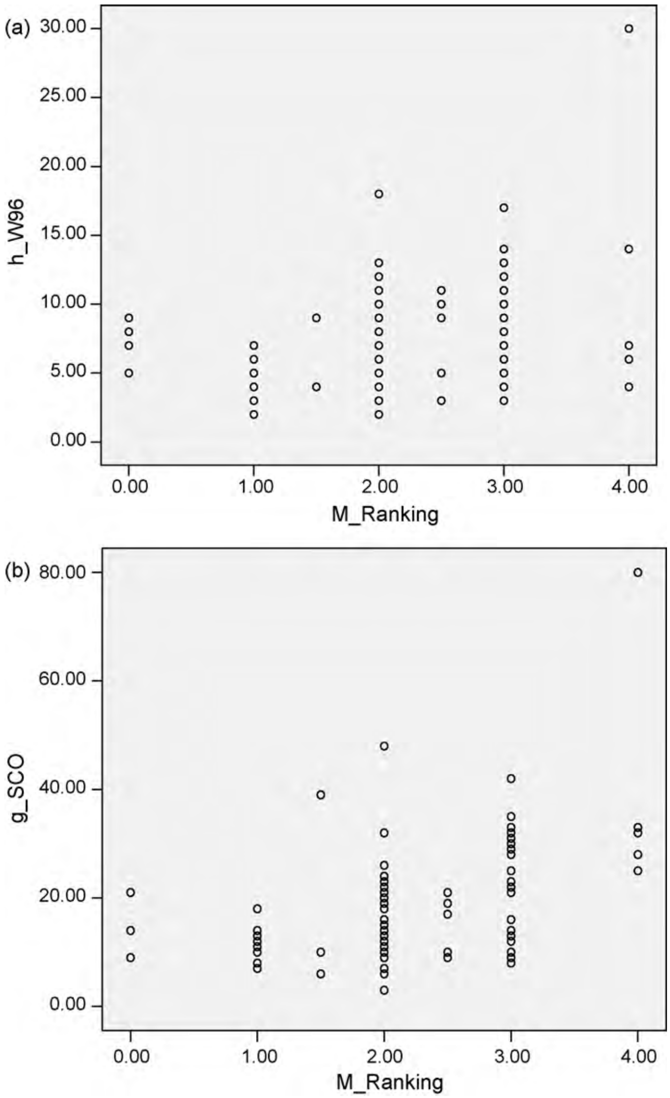

**Fig. 1.** Scatter diagram demonstrating the extent of the relationships between the median rankings (calculated over all of the experts) and (a) the *h*_W96 values, (b) the *g*_SCO values.

A criticism that has been leveled at the use of GS for bibliometric analyses is the large number of errors and duplications that are present in the data as a result of the largely automated procedures that are employed for the creation of the GS database. The raw *Publish or Perish* output was hence carefully filtered to remove all obviously erroneous and duplicate entries. Similar to the experience reported by (Meho & Yang, 2007) this increased the processing times of GS data by an order of magnitude. However, this substantial effort resulted in relatively small changes in results. Citation counts were about 5% greater and although correlations with the expert judgments were stronger, the Spearman values for the *h*-, *g*- and *H*-indices only increased a small amount: 0.502, 0.538 and 0.529, respectively against the values of 0.497, 0.529 and 0.524 in Table 3. The fact that cleaning of GS output has little effect on the ranking of researchers has also been noted by Baneyx (2008). His study was, however, very limited, involving the citations to five leading sociologists, and our studies hence provide a firmer basis than previously for the use of raw GS output as a cost-effective alternative to manual filtering.

There has been considerable discussion as to the extent to which WoS, Scopus and GS provide the same, or different, information (see, e.g., Bar-Ilan, 2008; Jacso, 2005; Norris & Openheim, 2007; Vieira & Gomes, 2009). Table 4 gives the correlations between each pair of index values, where it will be seen that they range from 0.344 (*h*_W96 and *g*_SCO) to 0.988 (*g*_SCO and *H*_SCO). It is noticeable that the correlations between the three indices for each data source are consistently large, with even the lowest value (for the correlation between *h*_WoS and *g*_WoS) being as high as 0.919. Hence, while the *g*-index and the *H*-index provide different ways of considering citation data, the values resulting from their application are very strongly correlated with those obtained from the basic *h*-index. Scopus covers citations from 1996 onwards, and we might hence expect that the W96 index values would correlate more strongly with the corresponding Scopus values than

**Table 4**
Correlation coefficients between the various citation indices. All of the coefficients are significant at the 0.001 level (2-tailed test) of statistical significance.

| | *h*_GS | *g*_GS | *H*_GS | *h*_SCO | *g*_SCO | *H*_SCO | *h*_WoS | *g*_WoS | *H*_WoS | *h*_W96 | *g*_W96 | *H*_W96 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *h*_GS | 1.000 | 0.937 | 0.972 | 0.797 | 0.755 | 0.777 | 0.628 | 0.642 | 0.651 | 0.459 | 0.532 | 0.533 |
| *g*_GS | 0.937 | 1.000 | 0.987 | 0.713 | 0.761 | 0.753 | 0.562 | 0.657 | 0.644 | 0.344 | 0.479 | 0.472 |
| *H*_GS | 0.972 | 0.987 | 1.000 | 0.762 | 0.779 | 0.784 | 0.602 | 0.662 | 0.664 | 0.402 | 0.516 | 0.515 |
| *h*_SCO | 0.797 | 0.713 | 0.762 | 1.000 | 0.932 | 0.962 | 0.783 | 0.731 | 0.769 | 0.787 | 0.807 | 0.811 |
| *g*_SCO | 0.755 | 0.761 | 0.779 | 0.932 | 1.000 | 0.988 | 0.731 | 0.778 | 0.791 | 0.703 | 0.803 | 0.803 |
| *H*_SCO | 0.777 | 0.753 | 0.784 | 0.962 | 0.988 | 1.000 | 0.763 | 0.772 | 0.799 | 0.744 | 0.820 | 0.823 |
| *h*_WoS | 0.628 | 0.562 | 0.602 | 0.783 | 0.731 | 0.763 | 1.000 | 0.919 | 0.952 | 0.820 | 0.824 | 0.840 |
| *g*_WoS | 0.642 | 0.657 | 0.662 | 0.731 | 0.778 | 0.772 | 0.919 | 1.000 | 0.985 | 0.703 | 0.790 | 0.798 |
| *H*_WoS | 0.651 | 0.644 | 0.664 | 0.769 | 0.791 | 0.799 | 0.952 | 0.985 | 1.000 | 0.747 | 0.814 | 0.830 |
| *h*_W96 | 0.459 | 0.344 | 0.402 | 0.787 | 0.703 | 0.744 | 0.820 | 0.703 | 0.747 | 1.000 | 0.937 | 0.949 |
| *g*_W96 | 0.532 | 0.479 | 0.516 | 0.807 | 0.803 | 0.820 | 0.824 | 0.790 | 0.814 | 0.937 | 1.000 | 0.980 |
| *H*_W96 | 0.533 | 0.472 | 0.515 | 0.811 | 0.803 | 0.823 | 0.840 | 0.798 | 0.830 | 0.949 | 0.980 | 1.000 |

they would with the corresponding WoS values (where no time limit has been applied). Table 4 shows that this is indeed the case; however, as noted previously, all of the W96 indices were less strongly correlated with the expert judgments than were the corresponding WoS indices.

An obvious question is whether the 12 citation measures in Table 2 are applying essentially similar criteria when ranking the 101 researchers. This was investigated using the Kendall *W* test (Siegel & Castellan, 1988), which measures the degree of concordance between multiple rankings of the same set of objects (i.e., the rankings of the 101 researchers resulting from use of the different citation measures). There are 12 rankings here (based on the three indices when applied to the four citation datasets) and the computed value for *W* is 0.771, demonstrating a statistically significant ($p \leq 0.01$) degree of correlation between the various rankings. Strong correlations are also observed between the *h*-, *g*- and *H*-indices when applied to a single dataset: the *W* values are 0.977, 0.974, 0.968 and 0.970 (all $p \leq 0.01$) for GS, Scopus, WoS and W96, respectively. Similar correlations are observed between the four datasets when they are analysed using a single index: the *W* values are 0.784, 0.783 and 0.802 (all $p \leq 0.01$) for the *h*-, *g*- and *H*-indices, respectively.

Table 2 lists the correlations obtained when the 101 researchers are considered as a whole. Very different trends are observed when the researchers were sub-divided into the five categories described previously, as shown in Table 5. The level of correlation for a dataset is strongly dependent on the range of values present, and significant correlation coefficients are typically larger for large datasets than for smaller ones (Bland & Altman, 1986). This is, however, not the case here since the 28 LIS researchers in the BIB category exhibit consistently stronger correlations with the citation indices than does the full dataset. Indeed, the magnitudes of the correlations here are sufficiently large to suggest that – for this subset at least – citation data might provide a cost-effective alternative to peer review for the evaluation of research excellence. Conversely, all of the coefficients for the IM and SOC subsets are lower than for the full dataset, with the IRH and IRS subsets falling between these two extrema.

Looking at the BIB researchers in more detail, the *h*-index values in Table 2 have been sorted to rank the researchers in descending order in Table 6 using each of *h*_GS, *h*_SCO and *h*_WoS: for example, Bar-Ilan is ranked 28th, 21st and 26th of the 101 researchers, respectively. Visual inspection of these data suggests that Scopus and WoS place researchers nearer the top of the overall ranking than does GS. BIB researchers would hence be disadvantaged, relative to the other subject groupings, were GS used to quantify research impact. This impression is confirmed by a Sign test (Siegel & Castellan, 1988), which shows that the *R*(*h*_GS) values in Table 6 are significantly greater than either the *R*(*h*_SCO) or the *R*(*h*_WoS) values ($p \leq 0.001$). As another example of this sort of differential behaviour, Sanderson has noted previously that LIS researchers

**Table 5**
Spearman correlation coefficients between median ranking and citation indices of the 101 academics when sub-divided into five categories. Correlations marked (**) are statistically significant at the 0.001 level of statistical significance, and those marked (*) at the 0.01 level; unmarked correlations are not significant at the 0.01 level.

| Index | Subject category | | | | |
|---|---|---|---|---|---|
| | BIB | IM | IRH | IRS | SOC |
| *h*_GS | 0.698 (**) | 0.276 | 0.357 | 0.575 (*) | 0.349 |
| *g*_GS | 0.727 (**) | 0.286 | 0.429 | 0.531 (*) | 0.466 |
| *H*_GS | 0.741 (**) | 0.271 | 0.404 | 0.574 (*) | 0.490 |
| *h*_SCO | 0.754 (**) | 0.224 | 0.646 (**) | 0.311 | 0.337 |
| *g*_SCO | 0.724 (**) | 0.196 | 0.671 (**) | 0.414 | 0.541 (*) |
| *H*_SCO | 0.782 (**) | 0.208 | 0.704 (**) | 0.423 | 0.465 |
| *h*_WoS | 0.756 (**) | 0.326 | 0.474 | 0.515 (*) | 0.358 |
| *g*_WoS | 0.762 (**) | 0.288 | 0.501 (*) | 0.633 (**) | 0.479 |
| *H*_WoS | 0.747 (**) | 0.306 | 0.508 (*) | 0.672 (**) | 0.492 |
| *h*_W96 | 0.652 (**) | 0.293 | 0.683 (**) | 0.028 | 0.355 |
| *g*_W96 | 0.780 (**) | 0.242 | 0.611 (**) | 0.264 | 0.485 |
| *H*_W96 | 0.727 (**) | 0.228 | 0.603 (*) | 0.281 | 0.495 |

**Table 6**
BIB academics' *h*-rankings in GS, Scopus and WoS.

| Author | $R(h\_GS)$ | $R(h\_SCO)$ | $R(h\_WoS)$ |
|---|---|---|---|
| Bar-Ilan, J. | 28 | 21 | 26 |
| Borgman, C.L. | 22 | 36 | 7 |
| Burrell, Q.L. | 55 | 47 | 24 |
| De Moya-Anegon, F. | 64 | 47 | 56 |
| Egghe, Leo | 28 | 21 | 11 |
| Garg, K.C. | 98 | 72 | 41 |
| Glanzel, W. | 11 | 4 | 2 |
| Gupta, B.M. | 96 | 81 | 87 |
| Harnad, S. | 2 | 15 | 76 |
| Jacso, P. | 87 | 72 | 62 |
| Lewison, G. | 69 | 36 | 41 |
| Leydesdorff, L. | 5 | 5 | 11 |
| Meyer, M. | 81 | 36 | 26 |
| Moed, H.F. | 13 | 6 | 5 |
| Oppenheim, C. | 28 | 21 | 41 |
| Rousseau, R. | 24 | 11 | 7 |
| Schubert, A. | 36 | 6 | 3 |
| Thelwall, M. | 11 | 3 | 6 |
| Van Leeuwen, T.N. | 49 | 21 | 26 |
| Van Raan, A.F.J. | 9 | 6 | 19 |
| Vaughan, L. | 55 | 36 | 33 |
| White, H.D. | 41 | 21 | 15 |
| Wilson, C.S. | 72 | 60 | 62 |
| Zitt, M. | 64 | 60 | 56 |

**Table 7**
IRH academics' *h*-rankings in GS, Scopus and WoS.

| Author | $R(h\_GS)$ | $R(h\_SCO)$ | $R(h\_WoS)$ |
|---|---|---|---|
| Goker, A. | 96 | 100 | 100 |
| Jarvelin, K. | 24 | 15 | 26 |
| Jose, J.M. | 72 | 67 | 76 |
| Kantor, P.B. | 41 | 36 | 62 |
| Lalmas, M. | 17 | 43 | 62 |
| Liddy, E.D. | 37 | 81 | 76 |
| Losee, R.M. | 51 | 47 | 33 |
| Ounis, I. | 55 | 72 | 92 |
| Van Rijsbergen, C.J. | 6 | 15 | 76 |
| Robertson, S. | 81 | 30 | 62 |
| Ruger, S. | 87 | 72 | 76 |
| Ruthven, I. | 41 | 47 | 62 |
| Sanderson, M. | 24 | 47 | 76 |
| Tait, J.I. | 77 | 88 | 97 |
| Whittaker, S.J. | 3 | 15 | 87 |
| Willett, P. | 1 | 1 | 1 |
| Yang, C.C. | 72 | 30 | 62 |

in strongly computational areas (such as the IRH subset here) may be disadvantaged when WoS is used in preference to GS for ranking purposes of LIS academics (Sanderson, 2008). This is also the case here, as shown in Table 7, where a Sign test shows that the IRH researchers have significantly greater $R(h\_WoS)$ values than either the $R(h\_GS)$ ($p \leq 0.05$) or the $R(h\_SCO)$ values ($p \leq 0.01$).

Finally, we have used the Sign test to analyse the data in Table 2 using the geographic categories (B, A and O). Specifically, we tested the hypothesis that there was no difference in the rankings for each pairing of expert category and researcher category, as shown in Table 8. Each element in this table gives the one-tailed level of significance. Only one cell shows a difference that is significant ($p \leq 0.01$), with Other experts ranking British researchers higher than did American experts;

**Table 8**
One-tailed Sign test results of peer review ratings.

| Researchers | Experts | | |
|---|---|---|---|
| | British vs. American | British vs. Other | American vs. Other |
| British (39) | British > American ($p = 0.015$) | $p = 0.271$ | Other > American ($p = 0.004$) |
| American (31) | $p = 0.412$ | $p = 0.133$ | Other > American ($p = 0.038$) |
| Other (31) | $p = 0.412$ | Other > British ($p = 0.026$) | Other > American ($p = 0.025$) |

there are also several other differences that are less significant ($p \leq 0.05$), e.g., British experts ranked British researchers higher than did American experts, and Other experts ranked Other researchers higher than did British experts. Thus, hardly surprisingly, there is some evidence of a geographic bias in the rankings listed in Table 2.

## 5. Conclusions

In this paper, we have considered the use of the *h*-, *g*- and *H*-indices for the ranking of 101 LIS researchers from around the world, based on citation data from the WoS, Scopus and GS databases.

Strongly significant correlations are obtained with the expert judgments of 42 LIS experts, demonstrating the strong relationship that exists between human and automated assessments of research impact. However, while the correlations are significant, their magnitudes are such that it would be premature in the extreme to suggest that citation-based indicators such as these could be used as a cost-effective alternative to expert judgments: the strongest correlations were obtained with the *g*- and *H*-indices computed using Scopus, but even this was only 0.552, i.e., the correlation explained only 30.5% of the variance in the data. The three citation indices and the four citation datasets applied essentially comparable rankings of the 101 researchers. It would appear that the apparently higher levels of noise in GS data compared to the other databases had minimal impact on its use in the types of bibliometrics analysis conducted in this study. Stronger correlations were obtained for subject-specific subsets of the 101 researchers, in particular for the more quantitative researchers in the BIB and IRH categories. GS disadvantaged BIB academics if used in preference to the other two citation databases and WoS disadvantaged IRH academics. Of the three databases, the strongest correlations overall were obtained using Scopus, despite the greater time-span of WoS. Geographic categorisation of the researchers and experts showed some degree of bias in that, e.g., British experts rated British researchers higher than did American experts. In possible future work, the degree to which geographic and web visibility might be affecting the performance measures and expert judgment should be further examined.

## Acknowledgments

We thank the Arts and Humanities Research Council and the China Scholarship Council for funding.

## References

Abramo, G., D'Angelo, C. A., & Caprasecca, A. (2009). Allocative efficiency in public research funding: Can bibliometrics help? *Research Policy*, *38*(1), 206–215.

Aksnes, D. W., & Taxt, R. E. (2004). Peer reviews and bibliometric indicators: A comparative study at a Norwegian university. *Research Evaluation*, *13*(1), 33–41.

Alonso, S., Cabrerizo, F. J., Herrera-Viedma, E., & Herrera, F. (2009). *h*-index: A review focused in its variants, computation and standardization for different scientific fields. *Journal of Informetrics*, *3*(4), 273–289.

Baneyx, A. (2008). "Publish or Perish" as citation metrics used to analyze scientific output in the humanities: International case studies in economics, geography, social sciences, phiolosophy, and history. *Archivium Immunologiae et Therapiae Experimentalis*, *56*(6), 363–371.

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*. Available from: http://www.dlib.org/dlib/september05/bauer/09bauer.html.

Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, *327*(8476), 307–310.

Bornmann, L., & Daniel, H.-D. (2006). Selecting scientific excellence through committee peer review—A citation analysis of publications previously published to approval or rejection of post-doctoral research fellowship applicants. *Scientometrics*, *68*(3), 427–440.

Cronin, B., & Meho, L. (2006). Using the *h*-index to rank influential information scientists. *Journal of the American Society for Information Science and Technology*, *57*(9), 1275–1278.

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

Egghe, L. (in press). Letter to the editor: On Randić's *H*-sequence. *Scientometrics*, doi:10.1007/s11192-009-0110-2.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(46), 16569–16572.

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.

Korevaar, J. C., & Moed, H. F. (1996). Validation of bibliometric indicators in the field of mathematics. *Scientometrics*, *37*(1), 117–130.

Maier, G. (2006). Impact factors and peer judgment: The case of regional science journals. *Scientometrics*, *69*(3), 651–667.

Meho, L., & 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.

Nicol, M. B., Henadeera, K., & Burtler, L. (2007). NHMRC grant applications: A comparison of "track record" scores allocated by grant assessors with bibliometric analysis of publications. *Medical Journal of Australia*, *187*(6), 348–352.

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

Norris, M., & Oppenheim, C. (2010). Peer review and the h-index: Two studies. *Journal of Informetrics*, *4*(3), 221–232.

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. (2007). Using the *h*-index to rank influential British researchers in information science and librarianship. *Journal of the American Society for Information Science and Technology*, *58*(5), 297–301.

Patterson, M. S., & Harris, S. (2009). The relationship between reviewers' quality-scores and number of citations for papers published in the journal *Physics in Medicine and Biology* from 2003–2005. *Scientometrics*, *80*(2), 343–349.

Randić, M. (2009). Citations versus limitations of citations: Beyond Hirsch index. *Scientometrics*, *80*(3), 809–818.

Reale, E., Barbara, A., & Costantini, A. (2007). Peer review for the evaluation of academic research: Lessons from the Italian experience. *Research Evaluation*, *16*(3), 216–228.

Rinia, E. J., van Leeuwen, T. N., van Vuren, H. G., & van Raan, A. F. J. (1998). Comparative analysis of a set of bibliometric indicators and central peer review criteria. Evaluation of condensed matter physics in the Netherlands. *Research Policy*, *27*(1), 95–107.

Sanderson, M. (2008). Revisiting *h* measured on UK LIS and IR academics. *Journal of the American Society for Information Science and Technology*, *59*(7), 1184–1190.

Schreiber, M. (2010). Revisiting the *g*-index: The average number of citations in the *g*-core. *Journal of the American Society for Information Science and Technology*, *61*(1), 169–174.

Seng, L. B., & Willett, P. (1995). The citedness of publications by United Kingdom library schools. *Journal of Information Science*, *21*(1), 68–71.

Siegel, S., & Castellan, N. J. (1988). Chapter 5, The case of one sample, two measures or paired replicates. In *Nonparametric statistics for the behavioural sciences* (2nd ed.). New York: McGraw-Hill.

van Raan, A. F. J. (2006). Comparison of Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups. *Scientometrics*, *67*(3), 491–502.

Vieira, E. S., & Gomes, J. A. N. F. (2009). A comparison of Scopus and Web of Science for a typical university. *Scientometrics*, *81*(2), 587–600.