Information Processing and Management 48 (2012) 234–241

Contents lists available at ScienceDirect

Information Processing and Management

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

Using the *h*-index to measure the quality of journals in the field of business and management

John Mingers *, Frederico Macri, Dan Petrovici

Centre for the Evaluation of Research Performance, Kent Business School, University of Kent, Canterbury CT7 2PE, UK

ARTICLE INFO

Article history:
Received 1 November 2010
Received in revised form 25 March 2011
Accepted 28 March 2011
Available online 19 April 2011

Keywords:
Citations
Google Scholar
h-Index
Impact factor
Journal quality
Research impact
Web of Science

ABSTRACT

This paper considers the use of the *h*-index as a measure of a journal's research quality and contribution. We study a sample of 455 journals in business and management all of which are included in the ISI Web of Science (WoS) and the Association of Business School's peer review journal ranking list. The *h*-index is compared with both the traditional impact factors, and with the peer review judgements. We also consider two sources of citation data – the WoS itself and Google Scholar. The conclusions are that the *h*-index is preferable to the impact factor for a variety of reasons, especially the selective coverage of the impact factor and the fact that it disadvantages journals that publish many papers. Google Scholar is also preferred to WoS as a data source. However, the paper notes that it is not sufficient to use any single metric to properly evaluate research achievements.

1. Introduction

There are many ways to evaluate the research quality of academic journals, the two main approaches being stated preference (i.e., peer review) or revealed preference (i.e., generally citations, or perhaps downloads) (Tahai & Meyer, 1999) although some studies use a combination of both (Mingers & Harzing, 2007). Both approaches have advantages and disadvantages, or perhaps supporters and critics, but in recent years there has been considerable research into revealed preference using citations.

Traditionally, the primary citation metric used has been the impact factor (IF), in either 2 year or 5 year form, originated by Garfield (1972). It is only available through the ISI Web of Science (WoS) and relies on the journals and citations that are collected in WoS. This measure takes the papers published in a journal over the last 2 years and calculates the mean citations per paper over that 2 year (or 5 year) period. There are a range of problems with this measure (Cameron, 2005; Seglen, 1997), in particular: (i) it uses a very short time window for many, non-science, disciplines. Mingers (2008) has shown that in the business and management field citations do not reach a peak until 10 years after publication and that there is little correlation between the number of citations in the first few years and the eventual total number of citations, (ii) being based on a mean, it can be quite sensitive to extreme values and it has been shown that citations tend to follow a negative binomial distribution (Mingers & Burrell, 2006) which is highly skewed and, (iii) having the number of papers in the denominator means that it penalises productivity, i.e., journals that publish a large number of papers.

Partly for these reasons, it has been suggested that the *h*-index (Hirsch, 2005) could be used instead of the IF to measure journal quality (Braun, Glänzel, & Schubert, 2006). The *h*-index is defined as: "a scientist has index *h* if *h* of his/her *N* papers have at least *h* citations each and the other (*N*–*h*) papers have no more than *h* citations each" (Hirsch, 2005, p. 16,569). It has

* Corresponding author. Tel.: +44 1227 824008; fax: +44 1227 76118.
E-mail addresses: j.mingers@kent.ac.uk (J. Mingers), fm83@kent.ac.uk (F. Macri), d.a.petrovici@kent.ac.uk (D. Petrovici).

0306-4573/$ - see front matter
doi:10.1016/j.ipm.2011.03.009

been used extensively in evaluating the work of individual researchers (see for example Bornmann, Mutz, and Daniel (2008), Egghe (2009), Egghe and Rousseau (2008)), and for research groups (van Raan, 2005). The main advantages of the *h*-index are: (i) it is very robust to poor data and to extreme values as it only counts the number of papers that have at least *h* citations, not how many citations they have, (ii) it is easy to calculate and to understand and, (iii) it values both the number of citations received, and the number of papers published by a journal which, we argue, is an advantage since it is the overall impact of a journal that we are trying to measure (Harzing & Van der Wal, 2009).

There are some disadvantages (Costas & Bordons, 2007; Glänzel, 2006): first, it only includes citations from the top *h* papers (known as the *h*-core) and ignores all lower papers; second, as it only counts the number of papers, and not their citations, it ignores the small number of papers that are extremely highly cited (Egghe, 2006); and third; as time goes on it continuously rises so its value is always a function of the time horizon (Burrell, 2007). In fact, these disadvantages are more prominent when it is used to assess individual researchers for whom total number of citations, very highly cited papers, and differing time spans are significant (Bornman & Daniel, 2005). In the case of journals, the time span is not a problem as we can choose an appropriate window that is suitable for all, and the very highly cited papers may, in any case, distort the overall comparison of contribution. Because of these limitations, many variants of the *h*-index (e.g., the *g*-index Egghe (2006)) have been produced but they in turn have their own biases and also increase the complexity and lack of transparency of the metric so in this paper we will concentrate on the fundamental *h*-index.

One of the first studies of the use of the *h*-index on journals was undertaken by Braun et al. (2006). The authors used Web of Science (WoS) to find the value of the *h*-index and they referred only to 1 year, 2001. They evaluated a top journal list according to the *h*-index with the rank based on the impact factor and concluded that the *h*-index looked "promising". Bornmann, Marx, and Schier (2009) analysed 20 organic chemistry journals found in the WoS database and analysed a number of indicators including the IF, and the *h*-index and several of its variants, over a 2 year period concluding that there was a high degree of correlation between the various measures. A recent article published in 2010 by Yin, Aris, and Chen (2010) analysed 20 top journals in the field of science and engineering using data from WoS. They combined the *h*-index with another recent measure – the Eigen factor – to provide graphical plots of the relative positions of journals as it was felt that the two measured slightly different characteristics.

Within the field of management, Moussa and Touzani (2010) analysed marketing journals using a variant of the *h*-index, the *hg*-index, which is the geometrical mean of the *h* and *g* indices. They used Google Scholar because marketing is poorly represented in Web of Science. Soutar and Murphy (2009) carried out a similar study of marketing journals with, as might be expected, very similar results. Finally, Harzing and van der Wal (2009) undertook a larger-scale investigation of over 800 business and management journals comparing the *h*-index generated from the citation data in Google Scholar (GS) with the impact factors generated by WoS. They concluded that the GS *h*-index gave a more accurate and comprehensive measure of journal impact, especially given that WoS does not have a good coverage across all subject areas (Walters, 2007).

These papers are all limited in various ways, especially concerning the number of journals covered, the timeframe, or the single source of data. In this paper we will be examining a large number of journals (460), over a 5 year period, utilising data from both Web of Science and Google Scholar, and also comparing the results with a peer reviewed journal ranking, that of the UK Association of Business Schools (ABS) (Association of Business Schools, 2010).

2. Methodology

The journals which compose the data set analysed come from the ABS 2010 journal ranking list. It is a hybrid list based in part upon peer review, in part upon citation measures and indicators and in part upon editorial judgement (Morris, Harvey, & Kelly, 2009). The list covers 821 journals across the broad spectrum of business and management and classifies the journals from 1 to 4 with a 4^* for a small number of top world journals. We have used the 460 journals which are included in Web of Science so that they all have an impact factor. This does not provide a representative sample in terms of subject area, of which there are 22 in the list, as some, e.g., finance, are under-represented in WoS. This does not mean that they are not high quality journals however (Mingers & Harzing, 2007).

The indicators on which the analysis of this work is based were found in two different databases: Google Scholar (GS) and Web of Science (WoS). In particular, Google Scholar was accessed through the Publish or Perish (PoP) software (Harzing, 2007) which allows easy and flexible access to GS. A previous study of these two data sources within business and management (Mingers & Lipitakis, 2010) found that while WoS was more accurate and rigorous, it was quite limited in many fields often covering less than 50% of papers published and ignoring totally non-journal outputs. On the other hand, GS has a wide coverage which does not vary much between different fields. It often includes nearly 90% of published outputs including books and reports. But, the citations it generates come from many different sources, often not research related such as masters theses or PPT presentations, and there are many mistakes necessitating much data cleaning. It is also limited to producing only 1000 results for any query. While this may be enough for an individual researcher, it is not always for a journal taken over a number of years. GS typically produces three or four times as many citations as WoS. By using both data sources we get a more comprehensive view of citation results and can compare the two (Bar-Ilan, 2008; Meho & Rogers, 2008).

The data was collected during June 2010 in two phases – the GS data through POP, and the WoS data. It covered publications during the period 2003–2008. One of the main problems with GS is the inaccuracy of the basic data and it requires considerable manual data cleaning. The main problems encountered were the journal name being incorrect with the search including other journals, and papers being duplicated in the results because of mistakes in spelling thus splitting the citations between them. On average, cleaning eliminated around 17% of the initial entries but the worst journals were "*Economic Policy*" (79.8%) and "*Transportation*" (76.7%) because their names are so common. The final results from POP gave not only the numbers of papers, numbers of citations and *h*-index but also a range of other bibliometric indicators. Note that one of the strengths of the *h*-index is its robustness to data problems as all that is required is that the top *h* papers be correct, all those below that are ignored.

The collection of citation data through WoS was less complicated as the accuracy is much higher even though the coverage is more limited. The impact factors came from the ISI Journal Citation Reports but these only apply to a particular year. This also produced other metrics such as the Eigen factor. There were also some problems over journal names between the different databases, mainly concerning spelling, which were generally resolved by reference to the ISSN.

The completed dataset was analysed for incorrect data and outliers. Some outliers were detected, mainly in the areas of Psychology and Finance, where the citations rates were very high in comparison the management as a whole and five journals were removed.

3. Results

3.1. Descriptive statistics of the metrics

Table 1 shows the main descriptive statistics for the various metrics.

In the main, these statistics present a comparison between GS and WoS. As is always found, GS generates significantly more citations than WoS, the mean number per journal 7377 as opposed to 2094, which is 3.5 times more in GS. This figure is comparable with previous studies in business and management (Mingers & Lipitakis, 2010). The resulting *h*-index and cites per paper (cpp) are similarly greater being 2.1 and 3.0 times higher respectively. The difference between these numbers is a result of the *h*-index not being overly affected by extremely high values which push the cpp upwards. The mean cpp of 5.5 in WoS is very much in line with previous studies as documented in Mingers and Lipitakis (2010). The statistical distribution of numbers of citations is highly skewed and has been shown previously to follow the negative binomial distribution (Mingers & Burrell, 2006). The distribution of the *h*-index is also skewed although less so than citations.

3.2. Relations between the h-index and other metrics

The first block of correlations in Table 2 compare the *h*-index with the IF's. First, we can see that the *h*-indices from the two sources are highly correlated (0.786), and the *h*-index from WoS is very highly correlated with the IF's, which are also based on WoS data (.807 and .767). The GS *h*-index is also well correlated with the IFs although at a lower level (.548, .597) as would be expected because the data sources are quite different and the two metrics measure different things. These figures can be compared with Harzing and van der Wal (2009) who obtained a correlation of 0.718 between the GS *h*-index and the WoS IF.

We should note that correlation analysis can be misleading especially, where the variables have underlying factors in common (West, Bergstrom, & Bergstrom, 2010), but we are not arguing that the metrics measure identical characteristics, merely that they do produce broadly similar rankings.

The next block considers the relations with number of papers and cpp. One of the main differences between the *h*-index and the cpp or IF (which is just a time-limited cpp) is that the *h*-index generally rewards productivity, in the sense that the more papers that are published the more opportunities to gain citations, while the mean citation metrics penalise productivity since the number of publications is the denominator. This makes it difficult for a journal that publishes a lot of papers to gain a high IF or cpp or, alternatively, these measures tend to reward journals that only publish a small number of papers. This can be seen clearly in the correlations, where the *h*-indices are strongly positively correlated to the number of papers, but the IF's and cpp's are weakly correlated, or not correlated at all.

Table 1
Descriptive statistics per journal (2003–2008), ($n = 455$).

	Mean	SD	Min	Median	Max
Papers (GS)	418.2	256.6	25	332	1000*
Papers (WoS)	390.2	304.9	12	298	1945
Citations (GS)	7376.8	8153.4	0	4401	51902
Citations (WoS)	2094.7	2492.7	0	1142	20983
Cites per paper (GS)	16.7	12.3	0	13.4	90.0
Cites per paper (WoS)	5.5	4.6	0	4.1	30.3
h-index (GS)	37.2	19.8	2	33	129
h-index (WoS)	17.5	9.7	1	15	68
IF (2 year)	1.455	1.060	0.125	1.188	9.028
IF (5 year)	2.125	1.566	0.248	1.743	11.582

* 1000 is the maximum GS will generate.

Table 2
Correlations between various indices ($n = 455$ except for 5 year IF which is 428).

	h-index (GS)	h-index (WoS)	IF	5 year IF	Papers (GS)	Papers (WoS)	Cpp (GS)	Cpp (WoS)
h-index (GS)	1	.786	.548	.597	.544	.328	.820	.650
h-index (WoS)		1	.807	.767	.454	.422	.637	.838
IF			1	.936	.162	.109	.575	.801
5 year IF				1	.139	.063*	.639	.870
Papers (GS)					1	–	.106*	–
Papers (WoS)						1	–	.009*
Cpp (GS)							1	–
Cpp (WoS)								1

* All correlations are significant at the 0.01 level except those marked.

3.3. Comparing the h-index with peer review

We have shown that the *h*-index is well correlated with other, citation-based, metrics such as the IF and the cpp, but how does it perform against a measure of journal quality based on peer review? In the UK, the journal list produced by the Association of Business Schools (ABS) has become a *de facto* standard, at least for UK academics. There are of course many other lists, some of which are available from the Harzing website (Harzing, 2009), but the process used in generating the ABS list is well documented (Association of Business Schools, 2010; Morris et al., 2009) and that is the one we have used.

The list classifies journals into five categories of quality, 1–4 and 4*, within 22 different fields. To do the comparison, we have classified our journals also into four groups (counting 4* as 4) based on their *h*-index value. The exact class boundaries are somewhat arbitrary, so we have chosen them so that roughly the same proportion of journals as in ABS fall into each category (see Table 3).

Using these boundaries, contingency tables were constructed for both versions of the *h*-index. The chi-square statistics were 133.7 for the GS one, and 200.8 for the WoS one with nine degrees of freedom so both were highly significant, with the WoS being more strongly related. This is not surprising since we have chosen specifically those journals from ABS that are included in WoS. Moreover, the IF is one of the metrics that the ABS Panel take into account when coming up with their judgements.

Of interest are the journals that are classified differently between the two indices and the ABS list. Table 4 shows those journals that rate highly in terms of *h*-index but are considered poor quality by ABS and Table 5 the opposite cases.

Table 4 shows journals that would be classified as four based on their *h*-index but are actually classified as 1 or 2 in ABS. Quite a few come from the information management/management science area, which may suggest that ABS under-grades this, but in fact four are *IEEE Transactions* which may just suggest that ABS does not recognise the quality of this series. One other commonality, not evident in the Table, is that these journals tend to have very high levels of citations (for example *J. Applied Social Psychology* has the largest number across all journals – 51, 902) and many have a large number of papers. Thus, the *h*-index recognises this overall contribution in a way that peer review does not.

In Table 5 we see journals that are judged to be of high quality under peer review but have very low citation rates. In the main, these come from the humanities/social sciences and this reflects the fact that citation rates are much lower in these disciplines. It is an important factor when interpreting any metrics based on citations – they should be normalised or at least contextualised to the appropriate field. A major question is, how specific do we need to be in the choice of field? As we shall see in a later section, there are major differences *within* business and management so it would not be satisfactory to normalise to this field as a whole. This also shows that citation measures should never be used by themselves – there is nothing intrinsically wrong with judging journals in a particular field to be of very high quality even though they have low citations.

3.4. The h-index or the impact factor?

In this section we will argue that the *h*-index is a better citation-based metric for measuring the research contribution of a journal than the impact factor.

The first problem with the IF is simply coverage. The IF is a proprietary product of Thompson and so can only be used on the basis of data from the Web of Science. As we will show in the next section, in many non-science fields the coverage of

Table 3
Boundaries for classifying journals based on the *h*-index.

	Grade 1	Grade 2	Grade 3	Grade 4
h-index (GS)	0–11	12–28	29–50	51–
No. of journals	17	160	189	89
h-index (WoS)	0–6	7–13	14–24	25–
No. of journals	23	159	185	88

Table 4
Journals with a high *h*-index but low ABS score.

Journal	Field	ABS 2010	*h*-index
From Google Scholar			
IEEE Trans. On Evolutionary Computing	Info Man	1	61
Health Economics	Econ	2	52
Computers and OR	OR & ManSci	2	52
Energy Policy	Sector	2	56
Applied Cognitive Psychology	Psych	2	52
J. Applied Social Psychology	Psych	2	99
World Economy	Econ	2	52
Political Studies	Soc Sci	2	58
From Web of Science			
IEEE Trans. On Systems, Man & Cybernetics C	Info Man	1	26
IEEE Trans. On Information Technology in Biomedicine	Info Man	1	26
IEEE Trans. On Evolutionary Computing	Info Man	1	44
Energy Economics	Econ	2	25
J. of Environmental Management	Sector	2	31
IEEE Trans. On Systems, Man & Cybernetics A	Info Man	2	25
Health Economics	Econ	2	29
Computers and OR	OR & ManSci	2	31

Table 5
Journals with a low *h*-index but high ABS score.

Journal	Field	ABS 2010	*h*-index
From Google Scholar			
J. of Financial and Quantitative Analysis	Finance	4	12
Business History	Bus Hist	4	12
Economic History Review	Soc Sci	4	17
Economic geography	Soc Sci	4	27
J. Experimental Psychology: Applied	Psych	4	28
American J. of Sociology	Soc Sci	4	12
Annual Review of Sociology	Soc Sci	4	19
From Web of Science			
Business History	Bus Hist	4	7
Economic History Review	Soc Sci	4	8
Human Resource Management (USA)	HRM and Emp	4	13

WoS is very patchy and often as many as 50% of the journals in an area are not included. Hence, it cannot be used for evaluating many journals. In contrast, the *h*-index via Google Scholar can be found for all journals.

Second, the IF is a very short term measure which only looks at citations over a 2 year period. There is now a 5 year IF which is better but still suffers from the other problems outlined here. In the social sciences and humanities citations take much longer to develop, typically not reaching a peak until 6 or 8 years after publication (Mingers, 2008). Moreover, there is little correlation between the number of citations in the early years and the eventual total. Two papers can begin with the same citation profile and then one will quickly become obsolete and no longer cited (a "shooting star") while the other will carry on for many years generating two or three times as many. Conversely, a paper may not be cited at all for several years and then suddenly be "discovered" and generate significant numbers of citations (a "sleeping beauty") (van Raan, 2004).

Third, all other things being equal, the IF penalises the number of papers a journal publishes while the *h*-index does not. This does not mean that as the number of papers published increases the IF *must* decrease but if there are more papers they must generate relatively more citations. We argue that it is the overall contribution of a journal in terms of impact, quantity and readership that is important. There are journals that publish relatively very few papers. By being so restrictive they can ensure that they publish only very high quality ones, receive many citations, and thereby gain a high IF. Whilst this may be good for the perceived quality of the journal, it does not make as significant contribution to the discipline as a journal that publishes a lot of very good papers as well as some outstanding ones.

We can show these effects with our data. In Table 2, the simple correlations between the WoS *h*-index and number of papers is 0.422; but between the IF and papers it is 0.109; and between 5 year IF and papers it is 0.63. If we regress IF against total citations (C) and number of papers (P) we get:

$$\begin{aligned} IF &= 1.49 + 0.00036C - 0.002P \quad R^2 = 42.0\%, \quad n = 459 \\ t &= \qquad 18.2 \quad - 10.9 \end{aligned} \tag{1}$$

We can get a better fit by regressing the IF against (C/P) which is really its definition:

$$IF = 0.34 + 0.21C/P \quad R^2 = 76.3\%, \quad n = 459 \qquad (2)$$
$$t = \qquad 38.4$$

Similarly for h:

$$h = 13.5 + 0.0033C - 0.008P \quad R^2 = 78.7\%, \quad n = 459 \qquad (3)$$
$$t = \qquad 35.8 \qquad -9.3$$

Here, it is interesting that when the citations are included the h-index also has a negative relationship with the number of papers despite its strong correlation with papers alone. The explanation for this is as follows. For a given total number of citations (C) there is a maximum possible h-index. Since the total number of citations accounted for by the h papers in the h-core is at least h^2, h will be a maximum when all the C citations are generated by the h-core papers, i.e., there are no excess papers. This occurs when $h^2 = C$, i.e., $h = \sqrt{C}$. So, for a journal with 2500 citations the maximum h would be 50 if exactly 50 papers each had 50 citations. In practice, this never happens and there are always considerably more papers, many of which have fewer citations. The more papers there are, the lower will be the h-index. So, whilst in general the more papers the more citations and the higher the h-index, once the number of citations is fixed the number of papers reduces the h-index. This can also be seen in Eq. (4), where we get an extremely good fit for h using $\sqrt{C}$ as a dependent:

$$h = 1.134 + 0.495\sqrt{C} - 0.0093P \quad R^2 = 95.4\%, \quad n = 459 \qquad (4)$$
$$t = \qquad 87.2 \qquad -23.9$$

This shows that, on average with this data set, the actual h-index is slightly less than half of the maximum that it could be for a given number of citations.

Rousseau (2006) takes the opposite approach by actually normalising the h-index to the number of papers published by a journal, which would seem to defeat the whole purposes and essentially reduce it to a form of IF.

3.5. Sub-fields within business and management

Table 6 shows the mean h-index, by source, for fifteen sub-fields of management together with the number of journals from each within our sample, and the mean impact factors. The first thing to notice is that there is significant variability between fields and that it is greater within the WoS data than the GS data, the difference between maximum and minimum is 236% for the former and 188% for the latter. ANOVA shows that the differences are statistically significant at the 1% level (for the GS figures, $F_{14,\,445} = 3.07$ and for WoS $F_{14,\,445} = 4.11$).

These differences do reflect different citation behaviour within the fields – the more "scientific" fields such as psychology and management science tend to cite more. But it also reflects the fact that some fields, again psychology and management science are examples, are better represented in WoS than others. Harzing and van der Wal (2009) found that 80% of the management science journals in their list were in WoS, but only 30% of finance and accounting ones were. This shows up clearly in Table 6, where finance and accounting have two of the highest h-indexes in Google Scholar but are near the bottom in WoS. In other words, WoS distorts the true citation rate for a journal in a way that GS does not.

A second consequence of these sub-field differences, mentioned earlier in the paper, is that citation rates have to be normalised or contextualised to the field. So, a GS h-index of 40 is only average in psychology or economics, but it would represent an extremely high figure in HRM. The impact factors show this even more clearly since the IF for psychology is twice

Table 6
Mean h-index by sub-field (Google Scholar and Web of Science).

	Mean h-index (WoS)	Mean h-index (GS)	No. of journals	Mean Impact Factor
Psychology	27.8	46.9	34	3.02
Operations research, man. science	21.8	39.5	27	1.40
Information management	20.6	37.3	39	1.87
Social science	18.6	35.7	46	1.46
Operations and technology	18.4	29.5	17	1.43
Marketing	17.8	35.4	22	1.72
International business	17.0	37.4	17	1.65
Sector (applied)	16.7	29.6	23	1.36
Organisation studies	16.5	30.5	16	1.64
Economics	16.5	43.2	108	1.21
Finance	16.5	45.3	25	1.32
Accounting	15.9	40.0	10	1.45
General management	15.7	34.0	35	1.49
Public sector	13.0	28.2	24	1.12
HRM, emp. relations	11.8	24.9	17	0.86

as much as most of the sub-fields showing that not only are there more citations in psychology, but they are generated much more quickly (and perhaps fade more quickly as well). The issue of normalisation is very important and has been studied intensively in terms of the Leiden "crown indicator", which is essentially the cites per paper normalised to the world average for the field (van Raan, 2003), but it has not developed much with regard to the *h*-index (Iglesias & Pecharromán, 2007). There is currently a significant debate about the proper basis for the crown indicator (Opthof & Leydesdorff, 2010; van Raan, van Leeuwen, Visser, van Eck, & Waltman, 2011; Waltman, van Eck, van Leeuwen, Visser, & van Raan, 2011a, 2011b).

4. Conclusions

The evidence in this paper suggests that the *h*-index would be a better citation-based metric for evaluating the quality and contribution of scholarly journals than other metrics such as the impact factor (IF) or the number of cites per paper (cpp). It also suggests that the data should be taken from Google Scholar rather than ISI Web of Science. The reasons for this are:

- The *h*-index is generally well correlated with more traditional measures such as those mentioned above, and also with peer review evaluations such as the ABS journal ranking list.
- In contrast to the IF or cpp, it recognises both the impact of a journal in terms of the number of citations, but also the overall contribution that the journal makes in terms of the number of papers published. The IF is actually inversely related to the number of papers and therefore particularly rewards journals that publish a small number of highly cited papers.
- In contrast to the IF, the *h*-index can measure citations over a long period which is much more appropriate for most non-science disciplines. The standard IF uses a 2 year period, whereas in many disciplines citations do not peak until around 10 years.
- The *h*-index is robust, and easily understood. It does not require that all the data about a journal is completely accurate, but only that concerning the top *h* papers. It is also insensitive to individual papers that are very highly cited, which can skew those metrics based on a mean figure. Future work could evaluate other variants of h, e.g., the *g*-index, to see if they offer any advantages.
- In terms of the source of data for the *h*-index, we would agree with Harzing and van der Wal (2009) that Google Scholar is better than ISI Web of Science, or indeed alternatives such as Scopus. WoS has a very poor coverage in many non-science disciplines and, as has been shown, even in one particular area such as business and management there are significant differences between sub-fields. Moreover, WoS only covers journal papers, and some conferences, thus missing out entirely on citations from other sources. We recognise that GS has limitations in terms of data quality and would urge Google to address this by making its sources more transparent and providing a greater degree of data cleaning, but the robustness of the *h*-index does not make this an insurmountable problem.

However, we would also caution that no one measure by itself, whether bibliometric or based on peer review, is adequate to properly measure the research contribution of researchers, journals or departments. They can be seen as complementary and more research should be carried out into way of appropriately combining them together to get a more well-rounded picture. We should also be aware that a concentration on purely scholarly measures of quality, such as academic citations, inevitably devalues the external impact of research on the economy and society (Adler & Harzing, 2009), and makes it less likely that researchers will address the really significant questions leading to research that "makes a visible difference" to our world.

References

Adler, N., & Harzing, A.-W. (2009). When knowledge wins: Transcending the sense and nonsense of academic rankings. *Academy of Management Learning and Education, 8*(1), 72–95.
Association of Business Schools. (2010, 1/3/2010). Academic journal quality guide, from <http://www.the-ABS.org.uk>.
Bar-Ilan, J. (2008). Which h-index? – A comparison of WoS, Scopus and Google Scholar. *Scientometrics, 74*(2), 257–271.
Bornman, L., & Daniel, H.-D. (2005). Does the h-index for ranking of scientists really work? *Scientometrics, 65*(3), 391–392.
Bornmann, L., Marx, W., & Schier, H. (2009). Hirsch-type index values for organic chemistry journals: A comparison of new metrics with the journal impact factor. *European Journal of Organic Chemistry, 2009*(10), 1471–1476.
Bornmann, L., Mutz, R., & Daniel, H. D. (2008). Are there better indices for evaluation purposes than the h index? A comparison of nine different variants of the h index using data from biomedicine. *Journal of the American Society for Information Science and Technology, 59*(5), 830–837.
Braun, T., Glänzel, W., & Schubert, A. (2006). A Hirsch-type index for journals. *Scientometrics, 69*(1), 169–173.
Burrell, Q. (2007). Hirsch index or Hirsch rate? Some thoughts arising from Liang's data. *Scientometrics, 73*(1), 19–28.
Cameron, B. (2005). Trends in the usage of ISI bibliometric data: Uses, abuses and implications. *Libraries and the Academy, 5*, 105–125.
Costas, R., & Bordons, M. (2007). The h-index: Advantages, limitations and its relation with other bibliometric indicators at the microlevel. *Journal of Informetrics, 1*, 193–203.
Egghe, L. (2006). Theory and practice of the g-index. *Scientometrics, 69*(1), 131–152.
Egghe, L. (2009). Mathematical study of h-index sequences. *Information Processing & Management, 45*(2), 288–297.
Egghe, L., & Rousseau, R. (2008). An h-index weighted by citation impact. *Information Processing & Management, 44*, 770–780.
Garfield, E. (1972). Citation analysis as a tool in journal evaluation. *Science, 178*, 471–479.
Glänzel, W. (2006). On the opportunities and limitations of the h-index. *Science Focus, 1*(1), 10–11.
Harzing, A.-W. (2007). Publish or Perish, from <http://www.harzing.com/pop.htm>.
Harzing, A.-W. (2009). Journal Quality List. Retrieved January, 2009, from <http://www.harzing.com/>.

Harzing, A.-W., & Van der Wal, R. (2009). A Google Scholar h-index for journals: An alternative metric to measure journal impact in economics and business? *Journal of the American Society for Information Science and Technology, 60*(1), 41–46.
Hirsch, J. (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.
Iglesias, J., & Pecharromán, C. (2007). Scaling the h-index for different scientific ISI fields. *Scientometrics, 73*(3), 303–320.
Meho, L., & Rogers, Y. (2008). Citation counting, citation ranking, and h-index of human-computer interaction researchers: A comparison of Scopus and Web of Science. *Journal of the American Society for Information Science and Technology, 59*(11), 1711–1726.
Mingers, J. (2008). Exploring the dynamics of journal citations: Modelling with S-curves. *Journal of the Operational Research Society, 59*(8), 1013–1025.
Mingers, J., & Burrell, Q. (2006). Modelling citation behavior in management science journals. *Information Processing and Management, 42*(6), 1451–1464.
Mingers, J., & Harzing, A.-W. (2007). Ranking journals in business and management: A statistical analysis of the Harzing dataset. *European Journal of Information Systems, 16*(4), 303–316.
Mingers, J., & Lipitakis, L. (2010). Counting the citations: A comparison of Web of Science and Google Scholar in the field of management. *Scientometrics, 85*(2), 613–625.
Morris, M., Harvey, C., & Kelly, A. (2009). Journal rankings and the ABS journal quality guide. *Management Decision, 47*(9), 1441–1451.
Moussa, S., & Touzani, M. (2010). Ranking marketing journals using the Google Scholar-based hg-index. *Journal of Informatrics, 4*, 107–117.
Opthof, T., & Leydesdorff, L. (2010). Caveats for the journal and field normalizations in the CWTS (''Leiden'') evaluations of research performance. *Journal of Informetrics, 4*(3), 423–430.
Rousseau, R. (2006). A case study: evolution of JASIS' Hirsch index. *Science Focus, 1*(1), 16–17.
Seglen, P. (1997). Why the impact factor of journals should not be used for evaluating research. *British Medical Journal, 314*(7079), 498–502.
Soutar, G., & Murphy, J. (2009). Journal quality: A Google Scholar analysis. *Australasian Marketing Journal, 17*, 150–153.
Tahai, A., & Meyer, M. (1999). A revealed preference study of management journals' direct influences. *Strategic Management Journal, 20*, 279–296.
van Raan, A. (2003). The use of bibliometric analysis in research performance assessment and monitoring of interdisciplinary scientific developments. *Technology Assessment – Theory and Practice, 1*(12), 20–29.
van Raan, A. (2005). Comparison of the hirsch-index with standard bibliometric indicators and with peer judgement for 147 chemistry research groups. *Scientometrics, 67*(3), 491–502.
van Raan, A., van Leeuwen, T., Visser, M., van Eck, N., & Waltman, L. (2011). Rivals for the crown: Reply to Opthof and Leydesdorff. *Journal of Informetrics, 4*(3), 431–435.
van Raan, A. J. (2004). Sleeping beauties in science. *Scientometrics, 59*, 467–472.
Walters, W. (2007). Google Scholar coverage of a multidisciplinary field. *Information Processing and Management, 43*, 1121–1132.
Waltman, L., van Eck, N., van Leeuwen, T., Visser, M., & van Raan, A. (2011a). Towards a new crown indicator: Some theoretical considerations. *Journal of Informetrics, 5*(1), 37–47.
Waltman, L., van Eck, N., van Leeuwen, T., Visser, M., & van Raan, A. (2011b). Towards a new crown indicator: An empirical analysis. *Scientometrics*. doi:10.1007/s11192-011-0354-5.
West, J., Bergstrom, T., & Bergstrom, C. (2010). Big Macs and Eigenfactor scores: Don't let correlation coefficients fool you. *Journal of the American Society for Information Science and Technology, 61*(9), 1800–1807.
Yin, C.-Y., Aris, M., & Chen, X. (2010). Combination of Eigenfactor and h-index to evaluate scientific journals. *Scientometrics, 84*(3), 639–648.