Journal of Documentation

The h-index: a broad review of a new bibliometric indicator

Michael Norris Charles Oppenheim

Article information:

To cite this document:
Michael Norris Charles Oppenheim, (2010),"The h-index: a broad review of a new bibliometric indicator", Journal of Documentation, Vol. 66 Iss 5 pp. 681 - 705
Permanent link to this document:
http://dx.doi.org/10.1108/00220411011066790

The *h*-index: a broad review of a new bibliometric indicator

Michael Norris and Charles Oppenheim
Department of Information Science, Loughborough University, Loughborough, UK

Received 19 June 2009
Revised 5 January 2010
Accepted 6 January 2010

Abstract

Purpose – This review aims to show, broadly, how the *h*-index has become a subject of widespread debate, how it has spawned many variants and diverse applications since first introduced in 2005 and some of the issues in its use.

Design/methodology/approach – The review drew on a range of material published in 1990 or so sources published since 2005. From these sources, a number of themes were identified and discussed ranging from the *h*-index's advantages to which citation database might be selected for its calculation.

Findings – The analysis shows how the *h*-index has quickly established itself as a major subject of interest in the field of bibliometrics. Study of the index ranges from its mathematical underpinning to a range of variants perceived to address the indexes' shortcomings. The review illustrates how widely the index has been applied but also how care must be taken in its application.

Originality/value – The use of bibliometric indicators to measure research performance continues, with the *h*-index as its latest addition. The use of the *h*-index, its variants and many applications to which it has been put are still at the exploratory stage. The review shows the breadth and diversity of this research and the need to verify the veracity of the *h*-index by more studies.

Keywords Information studies, Indexing, Information research

Paper type Literature review

1. Introduction

Scholars and departments frequently have the quality of their work assessed and ranked by their academic colleagues. The results of such assessments or rankings can lead to individual preferment, receipt of research funding or tokens of esteem.

Apart from peer review, the impact of a scholar's work could, in part be assessed by the use of bibliometric indicators. Traditionally, for example, the major foci of citation studies have been the evaluation of academic departments, or of a countries' entire research output or of journals. A few citation studies have, however, focused on the performance of individuals by measuring their citation counts (Baird and Oppenheim, 1994; Brittain, 2000; Kalyane *et al.*, 2004). Generally, the assessment of an individual's impact by the use of citation analysis has been very limited, given the lack of a citation metric that might reflect a scholar's overall achievement.

Recently, however, in 2005 Hirsch developed a new, simple and intuitively attractive measure of an individual's impact in his or her field – the *h*-index (Hirsch, 2005). Hirsch defined the index as follows:

> A scientist has index *h* if *h* of his/her N_p papers have at least *h* citations each and the other papers have no more than *h* citations each [...].

The authors wish to acknowledge the support of the AHRC in funding this research.

Journal of Documentation
Vol. 66 No. 5, 2010
pp. 681-705
Emerald Group Publishing Limited
0022-0418
DOI 10.1108/00220411011066790

Thus, a scholar who has a *h*-index of 20 has at least 20 papers with at least 20 citations each, but does not have 21 papers with at least 21 citations each. The new index has been viewed favourably in the academic community and may be readily calculated using automated features found in Web of Knowledge or Scopus citation databases and additionally for Google Scholar (GS). The index gives a broadly balanced view of an individual's impact rather than some of the alternative measures that might be used, such as simple citation counts or the number of publications.

Since 2005, the uses of the *h*-index and its properties have been the subject of much discussion and debate. A range of comment and scholarly articles on its uses, its methodological underpinning, how it might be improved and how it might be applied have appeared. Unlike other bibliometric measures, the *h*-index, it is argued, takes more account of the lifetime achievement of a scholar's work. Hirsch (2005) himself lists what he feels are the shortcomings of a number of existing single measures of academic success; these include the total number of papers, the total number of citations and the citations per paper. Using the *h*-index in place of these measures, it is argued, can give a fairer measure of an academic's overall impact. Such has been the interest in the index that it has been applied tentatively at least to groups of scholars, journals, subjects, countries, and Hirsch has suggested it can be used to measure the future performance of academics (Hirsch, 2007). The lure of a single measure of academic performance unencumbered by the trappings of peer review that is transparent, and open for all to quantify is considerable. In addition, it potentially rescues from obscurity those who have made a steady and consistent contribution without due recognition (Ball, 2005).

2. The *h*-index

Hirsch (2005) speculated in his original paper how the impact and relevance of an individual scholar could be fairly quantified, so that when limited resources are to be awarded, they can be equitably apportioned to those where it will be most beneficial. In deriving his index, Hirsch describes it as avoiding other single measures of impact that can be used to evaluate scholars and their output that are less discerning in their ranking of overall scholarly value.

Hirsch (2005) calculated the *h*-index for 21 physicists, some of whom were Nobel Prize winners. Highest among them was Edward Witten a physicist from Princeton Institute for Advanced Study, widely regarded by his peers as the most eminent living physicist (Ball, 2005). Witten had a *h*-index of 110, that is 110 publications with at least 110 citations each. Hirsch argues that a high *h*-index identifies those who consistently produce a stream of recognisably good work over a sustained period, rather than those who produce a handful of excellent papers which are heavily cited or those who produce much, but relatively poorly cited work.

The index does not take account of all the work that an author produces; uncited work is discounted at the outset. An author with a *h*-index of 4 will only need four articles that have been cited at least four times, regardless of how many other articles that have been written and cited beneath the threshold. Similarly, an author with just four very highly cited articles will still only have a *h*-index of 4 irrespective of their much higher level of citation. Thus, Hirsch (2005, p. 16569) suggests that the two academics with similar *h*-indexes "are comparable in terms of their overall scientific impact, even if their total number of papers or their total number of citations is very different". It follows then that the more productive an author is and the longer they have been active, the more likely they are to have a higher *h*-index than a comparative newcomer. In the field of physics at least, it is suggested by Hirsch (2005), that anyone with a index of 20 after 20 years of activity would be considered as successful, someone with an index of 40 active for the same period would be considered outstanding and anyone with a *h*-index of 60 or above after 20 years would be unique in their achievements. Clearly, the parameters for success will vary from field to field.

As noted earlier, the *h*-index has been well received; it is easy to calculate and gives a more balanced view of an individual's impact than some of the alternative measures such as total number of papers published or total number of citations received. If the academic continues to be active, the value of *h* will increase over time, reflecting the career progression as individual scholars continue to publish papers, some of which may be highly cited, and their earlier papers are likely to attract more citations. Once an author has acquired a particular *h*-index, it cannot go down even if the author then becomes inactive. Disciplines differ in the level of co-authorship and those authors who find themselves co-authoring with many authors may find that they are overly favoured; likewise, authors who are heavy self-citers may contribute disproportionately to their *h*-score, although articles with many self-citations may not reach the threshold count.

In an extension to his original work, Hirsch (2007) has suggested that the *h*-index could also be used to predict the future success of an individual when compared to other bibliometric indicators. Hirsch showed that the *h*-index was a better predictor of future success than the mean number of citations per paper, the number of papers published or the number of citations received. If further studies prove this to be correct then the *h*-index potentially at least, could help identify scholars who are likely to be more successful in the future than their peers as well as identifying those who are currently successful.

2.1 Advantages of the h*-index*

The literature on the *h*-index is already extensive. Bornmann *et al.* (2008) noted that 30 papers had been published just one year after Hirsch's (2005) original paper and recent counts suggest that this is now approaching 200. An entire issue of *Scientometrics* and a special issue of the *Journal of Informetrics* were devoted to the subject. Much of this literature considers alternatives to the *h*-index, discusses its theoretical aspects or applies it in different contexts to see how it performs. The numerous authors also consider the many advantages and disadvantages, as they perceive them, of the *h*-index. Hirsch (2005), Glänzel (2006a, b), Costas and Bordons (2007), Jin *et al.* (2007) and Liu and Rousseau (2009) amongst others systematically list a range of these advantages and disadvantages. Similarly, Alonso *et al.* (2009) provide a wide ranging review of the *h*-index including its advantages and disadvantages.

Foremost amongst these advantages is the relative simplicity of the *h*-index. It is easy to calculate and anyone who has access to Web of Science (WoS), Scopus or for that matter GS can calculate their own or others *h*-index. However, as Cronin and Meho (2006) note, obtaining accurate citation counts on which to reliably base *h*-index scores needs knowledge, of, for example, WoS. Academics, can however, get a quick estimate of their *h*-index using WoS by sorting articles they have written or co-written by the number of times they have been cited and then scrolling down the output until the rank of the paper is greater than the citation count; the rank of the preceding article indicates their *h*-index (Bornmann and Daniel, 2007). For those less inclined to do this, both the WoS and Scopus have automated this process, so anyone can then view (WoS/Scopus) *h*-index along with other statistics for any given author.

Robustness is an important issue and Braun *et al.* (2006, p. 170) find that the *h*-index is:

> [...] insensitive to an accidental excess of uncited papers and also to one or several outstandingly highly cited papers. Second, it combines the effect of "quantity" (number of publications) and "quality" (citation rate) in a rather specific, balanced way that should reduce the apparent overrating of some small review journals.

As Hirsch (2005) notes, the *h*-index corrects for "big hits" or "one hit wonders" where authors have written or co-written one or a handful of very highly cited papers. Thus, those authors who produce a steady stream of good quality work are more likely to have a higher *h*-index than authors who do not. Rewarding consistency in output is a feature recognised by Ball (2005), who finds one of the *h*-indexes' main attractions is that it can rescue consistent, but unrewarded scholars from obscurity. This is despite the fact that Van Raan (2006) considers that such an attraction has been available for sometime from the advanced bibliometric indicators that can be computed by his institute, albeit at some cost to those wanting the information.

2.2 Criticisms of the h*-index*

Those who have cited the advantages of the *h*-index also list its disadvantages. Bornmann and Daniel (2007) take up comments from Van Raan (2006) and Glänzel (2006b) about the dangers of trying to "force" the assessment of scholars into a single measure rather than using a broad range of indicators. The indicator can never decrease once an academic has achieved a particular *h*-index even if they cease to be productive, so scientists having achieved a good *h*-score could rest on their laurels (Jin *et al.*, 2007). The index is also potentially limited to the number of articles that are produced over a productive lifetime since, the index cannot exceed the number of publications produced, although realistically it is extremely unlikely that all articles would be included in the calculation for even a modest number of papers. Problems may occur when an author has only a small, but highly cited set of papers; despite their high citations counts, these sets of papers are at a disadvantage. As Glänzel (2006b) and Van Raan (2006) remark, "small is not beautiful" in this case. Such sets of papers that have high citation counts well beyond an individual's *h*-index cease to contribute to any improvement in their *h*-score even if the article attracts further citations (Costas and Bordons, 2007). In an attempt to overcome this insensitivity to highly cited articles, Egghe (2006a, b) has proposed the *g*-index, which gives more weight to highly cited article counts beyond the *h*-core. However, as Hirsch (2005) notes, one of the advantages of the *h*-index is that it is insensitive to highly cited articles, such as review articles.

Another criticism is that an *h*-index could be inflated by self-citation or by extensive co-authorship; all can distort *h*-scores. Self-citation is an issue of concern for any metric that counts citations as of a measure impact. The levels of self-citation Cronin and Meho (2006) found were low in their study of information science scholars and did not appreciably alter rankings once removed. This was not the case, however, for Schreiber (2007), who found in physics related fields and particularly for young scientists, marked differences in ranking when self-citations were removed both for the *h* and *g*-index.

The *h*-index is effective when ranking long-standing academics that have produced a steady stream of good quality work. On the other hand, early career researchers tend to be at a disadvantage inasmuch that their volume of work may be limited, as too, may be the number of citations they have accrued to these articles. To compare scientists, Hirsch (2005) proposed the *m* parameter, that is the division of the *h*-index by the number of years (their scientific age) that an academic has been research active. As Harzing (2008) notes:

> [. . .] that *m* generally does not stabilise until later in one's career and that for junior researchers (with low *h*-indices) small changes in the *h*-index can lead to large changes in *m*. [and] Moreover, *m* discriminates against academics that work part-time or have had career interruptions (generally women). However, in some cases *m* might be a useful additional metric to evaluate an academic's achievement.

Kelly and Jennions (2006) confirm these points. In their analysis, they chose to examine the *h*-index of editorial board members taken from seven evolution and ecology journals. In their conclusions, they noted that there was a perceptible gender effect where female scientists generally produced fewer papers, which affected their *h*-index and that "For younger researchers, a one-year difference in the onset of publishing and a tiny change in *h* can dramatically affect *m*".

3. Features of the *h*-index

3.1 Theoretical and mathematical considerations

As well as Hirsch's (2005) initial paper and his further work (2007) on the predictive power of the *h*-index, a substantial number of articles have examined the theoretical, statistical and axiomatic aspects of the *h*-index, providing models and discussion of its particular characteristics. A number of authors have also introduced alternatives to the *h*-index or variations of it and have shown their mathematical reasoning behind them. Egghe (2006a, b) is prominent amongst these, Egghe (2006c) extends Hirsch's (2005) original definition into a general framework of information production processes. This he follows by introducing his *g*-index (Egghe, 2006c), a variant of the *h*-index for which he not only provides a mathematical theory but illustrates by recourse to his own publication record and citations. Egghe writing alone and with fellow academics authors something in the order of eight papers on the *h*-index during 2008 and its variations. Egghe and Rousseau (2008) propose a version of the *h*-index which is weighted by citation impact and is drawn from the *h*-core. Egghe (2008b) also examines the successive *h*-index introduced by Prathap (2006) and Schubert (2006) and shows how this works collectively for Price Medallists[1] and when all their citations are merged. Extending this work, Egghe and Rao (2008) define three *h*-indices the first of which is successive and they also show, amongst their calculations, how closely correlated their *h*-indices are to author productivity and the number of citations they received. Further to this Egghe *et al.* (2009) show, using a power-law model how the *h*-index and the journal impact factor (JIF) can be unified into a single relation. The authors demonstrate this by recourse to real data by collecting data from the 304 journals which make up the eight physics categories in the Journal Citation Reports.

Taking a different approach Woeginger, in a series of three articles discusses the *h*-index (Woeginger, 2008c), the g-index (Woeginger, 2008b) and with some consolidation in the final paper (Woeginger, 2008a) the *h* and *w*-index, not to be confused with Wu's *w*-index, (Wu, 2008). In these articles, Woeginger (2008c, p. 231) takes an axiomatic approach to the characterisation of these indexes, with the *w*-index falling "out for free from our axiomatic investigations around the Hirsch-index".

The *w*-index is calculated in reverse to the *h*-index by working upward through the list of publications in increasing order, rather than downward[2]. Building on this work, Rousseau (2008) considers which of Woeginger's axiomatisations are satisfied by the *g*, *h*(2) and R^2 indices. In a not dissimilar approach, Van Eck and Waltman (2008) take to task those who have, in their view, accepted the arbitrary value that Hirsch has given in his definition of his *h*-index. Through a series of definitions and propositions, they derive a generalised version of the *h* and *g*-indices which they apply to Price Medallists to show their differing effects.

Glänzel (2008a, b) considers how two independent *h*-indexes can be concatenated when nothing is known except their *h*-indexes; he takes two test examples, one for the journals *Science* and *Nature* and the second for two bibliometricians who have never co-authored. Having derived the theorem, he applies this to the data and finds that he is readily able to determine their collective *h*-index without having to resort to the citation data.

3.2 Robustness, validity and convergence

Bornmann and Daniel (2005) asked whether the results of a peer review process that ranked and awarded young scientists would compare well when these scientists were ranked using their *h*-index. To investigate this, they compared, with the applicant's *h*-index the results of a peer review process over a five-year period where post-doctoral research fellowships were awarded for research in biomedicine. The 414 applicants had published 1,586 papers prior to their application, and these received 60,882 citations when counted using the WoS's Science Citation Index. The results showed that the 64 successful applicants had on average a higher *h*-index than those who were unsuccessful. Without exception for the successful applicants, in every year the *h*-indexes and for that matter, the citations per paper were higher than those who were rejected.

In a similar study, Bornmann *et al.* (2008) considered the applicants for long-term fellowships and young investigator programmes to the European Molecular Biology Organization. They attempted to determine if there was a relationship between their respective *h*-index, the number of publications, total citation counts, average JIFs and peer assessments. In the first stage, the bibliometric indicators were compared to the *h*-index and in the second, the results of peer review. The authors found, as in the earlier work of Bornmann and Daniel (2005), a strong correlation between number of publications, total citation counts and the applicant's *h*-scores. Correlation between average impact factor (IF) and *h*-score was poor.

Usefully, Bornmann *et al.* (2008) show their results in comparison with earlier work where other authors show the degree of correlation between their results and *h*-index scores, (Bornmann and Daniel, 2005; Costas and Bordons, 2007; Cronin and Meho, 2006; Kelly and Jennions, 2006; Saad, 2006; Symonds *et al.*, 2006; Van Raan, 2006). In the second stage of the work, Bornmann *et al.* (2008) show how the *h*-index scores for those successful applicants for funding are on average higher than those who were not. However, the authors found that the level of correlation between the scores obtained by peer review and the *h*-scores could only be classed as "medium or typical in the applied social sciences" (2008, p. 155). This they reasoned was because the publication record was only part of the assessment criteria that the committee would use in its assessment.

The citation records of 55 scholars who had published five or more papers in the *Journal of Consumer Research* between 1989 and 2005 were compared by Saad (2006). Using the WoS database, he compared the number of citations that each author had received for their papers and their *h*-index. The correlation between the two was 0.87 ($p < 0.001$). He also noted the correlation between the GS *h*-index and the WoS citation counts was 0.83 ($p < 0.001$). The correlation between the two *h*-indices was 0.82 ($p < 0.001$). Finally, using a block sampling design a random sample of journals was taken from 508 journals of interest to business schools, he showed that the journal's *h*-index and IF were reasonably correlated (0.64, $p < 0.001$). Saad shows that there is convergence between the different bibliometric measures; this is not surprising since they are derived from the same pool of citation counts. Of interest, however, is the difference in correlation, a point noted by Hirsch (2007); he found that of the other measures he considered (number of papers and number of citations per paper) that the cumulative author citation counts after 12 years and *h*-index scores were the most closely correlated.

In another approach, Van Raan (2006) measured the correlation between the *h*-index, standard bibliometric measures and peer judgement. Posing the crucial question of whether there is a correlation between these indicators, he addressed this question to research groups rather than the individual scientist. Van Raan took for his sample 147 chemistry university research groups working in The Netherlands and their 18,000 or so publications which appeared between the years 1991 and 1998 and whose journals appeared in the WoS citation indexes. A range of bibliometric indicators and the *h*-index were calculated for these publications, based on citation counts limited to the first three years after their publication. The results showed a strong correlation between total citation counts and the *h*-index but, as others above found, there was a weaker correlation with total publications. Van Raan then introduces the use of a subject specific measure; the crown indicator[3]. Using the crown indicator score obtained for the chemistry group and *h*-index, the scores were then compared to the peer ranking of these research groups. Both measures discriminated well research groups that were rated excellent or good with the less good, but less well when trying to discriminate between the excellent or good categories themselves. Both indicators were able to relate in a comparable way with peer review but the *h*-index was less successful at discriminating those departments whose peer review had placed them at lower levels.

Vanclay (2007) considers the robustness of the *h*-index in terms of its relative stability when encountering errors in citation records and the need to verify, if necessary, the citation counts around "threshold" records that may affect a *h*-index score. He also points to its relative simplicity, the ease with which it may be calculated and its single integer form. Vanclay calculated his own *h*-index using citation records from WoS and GS, correcting for errors and showing the result both with and without self-citations. He demonstrates that the original uncorrected score changes little when all corrections are made. Similarly, Vanclay calculated the *h*-index of the journal *Forest Ecology and Management,* finding that each of the scores remained the same for WoS and GS after corrections.

Rousseau (2007) notes that it is now possible to count and collect citations from a range of sources, including, for example, GS, as well as Google and from WoS. Highly cited items may appear in sources other than WoS. This then raises the possibility that uncounted sources may affect a scientist's *h*-index. Rousseau (2007) examines the effect of adding five highly cited articles to former Price Medallists and finds that this results in an increase in their *h*-index of between 0 and 4, with a mean of 2.21, less than half of the missing number of publications. Rousseau (2007, p. 6) concludes that a small "number of missing highly cited publications has only a small influence on the value of the *h*-index". This influence may be minor on Price Medallists, who have relatively high *h*-indexes; such omissions might be quite important when *h*-scores are quite low. Sanderson (2008) noted similar, but larger variations in *h*-scores when different databases were used to count citations; most noticeably, scores varied considerably when GS counts were included.

3.3 Self-citation and the h*-index*

Academics cite themselves to a greater or lesser extent; the frequency of self-citation often varies depending on the field in which the author works (Snyder and Bonzi, 1998). Self-citation will clearly inflate an author's overall citation count and potentially at least could also inflate their *h*-index. A number of authors have noted the affect that self-citations can have on the *h*-index. Vinkler (2007) examined 21 eminent scientists working in different fields of chemistry, using a battery of measures, the *h*-index among them, he concluded calculating a *h*-index with self-citations may inflate *h*-scores and this may be particularly unsatisfactory when individual scores are compared.

Engqvist and Frommen (2008) examined the effect of self-citations on the *h*-index. From a slightly different perspective, they argued that it would be quite difficult for an author to target the appropriate article and selectively self-cite in order to inflate their *h*-index, given that this would always need to be done at the *h*-index borderline and would have to increase at the same rate as the *h*-index itself. Taking 40 authors from evolutionary biology, they identified the citation causing the most recent increase in the author's *h*-index and the first citation appearing thereafter which would have the same effect. The difference between the dates of these two citations gives the length of time that any one citation would affect the *h*-index. Within one month, more than 20 per cent of all *h*-increasing citations were redundant; this rose to 50 per cent in two months and no more than 10 per cent were crucial to an authors *h*-index after nine months. Dealing with a second issue, that of random self-citing inflating a *h*-index score the authors conjecture that within a discipline where self-citing norms are likely to be very similar that any inflation by self-citation is likely to be roughly uniform. Asking whether "there is a systematic variation in the *h*-index that is linked to the authors self-citation rate", Engqvist and Frommen (2008, p. 251), taking 40 authors, calculated the decrease in their *h*-score after excluding all self-citations. They found that only five of the authors had a decrease in their *h*-index by more than two, but that there was a correlation between consistent self-citation and an elevated *h*-index. However, the authors noted that in this case, that on average, each author had 2.02 ± 1.13 self-citations per paper and it would need an author on average to have three self-citations per paper to increase their *h*-index by one.

Schreiber (2007), on the other hand, proposes that it is necessary to sharpen the *h*-index by taking into account the effect of self-citation. Taking his own publication record as a physicist, he found that when all self-citations were removed that his *h*-index dropped by 18 per cent and those of two of his colleagues were variously, but noticeably affected. Overall, Schreiber (2007) found that from a methodological point of view it was sufficient to check for self-citation counts around the "borders" of the *h*-index but that Hirsch's (2005) view that only a few papers need to be dropped from the *h*-count if self-citations are taken into account is erroneous. This is particularly, the case for the middle ranking or junior scientist if an accurate *h*-index is to be calculated; these are perhaps, the group for whom the *h*-index is likely to be important, rather than the very well established senior academic. Extending his examination of the *h*-index and the effect of self-citation, Schreiber (2008b) turned his attention to Egghe's (2006c) *g*-index. Using data sets similar to those found in Schreiber (2007), Schreiber examined the influence of self-citations on this variant of the *h*-index and found that there was a consistent reduction in the value of the *g*-index (described below) across the academic subjects being investigated. The reduction in score varied between one and 12 with an average of 5 or 21.7 per cent.

In a larger study, Cronin and Meho (2006) examined the *h*-index of 31 mid-to late-career information science authors from the USA. After removing self-citations, they found a somewhat smaller drop in *h*-index values of the order of 8 per cent, a notably smaller drop than noted in the work reported above, but with little change in the relative ranking of the authors. However, Kosmulski (2006), in contrast, in his examination of 19 professors affiliated to a department of chemistry found that after removing self-citations that this reduced, on average, their *h*-index scores by 26 per cent.

3.4 Co-author frequency and the h*-index*

Apportioning credit amongst the co-authors of a cited item is problematic. A simple division of the citations by the number of authors might seem to be fair, but many articles, particularly in the sciences may have hundreds of authors. It is likely that being included in a large number of articles as one of many co-authors that these articles may rapidly accrue citations when the contribution of each author to the work is not so easily discerned (Wan *et al.*, n.d.). Similarly, any *h*-index that is based on many co-authored articles may disadvantage those scholars who work alone in the same field. Several authors have suggested solutions to this problem. Hirsch (2005, p. 16571) notes the problem in his initial article and suggests that it may be useful "[...] in comparing different individuals to normalise *h* by a factor that reflects the average of number of co-authors". Schreiber (2008c) feels that this approach penalises authors with a number of papers with many co-authors and has an excessively reductive effect on the contribution of single author publications to a *h*-score. Egghe (2008a) gives two methods for fractionalising the contribution of an article that has more than one author. In the first method, authors are ranked by their fractionalised citation count; for an article with ten citations and two authors, this paper will have a value of five. In the second method authors are ranked by the fractionalised author count; a paper with four authors has a value of one quarter. Egghe (2008a) then shows how his own *h* and *g*-indexes are affected by fractionalisation. His initial *h*-score at 14 is only marginally reduced to 12 for the fractionalised citation count and 12.83 for the author count. The initial *g*-index of 20 ranges similarly, from 18 to 22. Schreiber (2008c) takes the fractionalised author count approach to assessing the ranking of eight famous physicists and finds that their initial *h* ranking is changed, in his opinion substantially. However, Schreiber does note that the Spearman Rank Order Correlation between the physicist's original *h* and the fractionalised *h* is high at 0.90. Nonetheless, Schreiber feels that while the fractionalised *h* may be regarded only as a modification, it is an important and necessary one.

3.5 Variants and alternatives to the h*-index*

The perceived shortcomings of the *h*-index have been explored by a number of authors who have suggested a range of variants or alternatives that attempt to improve its usefulness as an indicator of research performance. Hirsch (2005), for example, was aware that his index was dependent on how long a scholar had been active and suggested that to compare scientists with different lengths of service, the *h*-index could be divided by their research active years. This *m*-quotient goes some way to take account of the disparity in *h*-index scores where scientists have widely differing productive years.

Egghe (2006a, b) has proposed the *g*-index, which takes into account those highly cited articles which are included in the *h*-core[4] but whose high citation counts beyond the *h*-core make no further contribution to an individual's *h*-index. Egghe (2006a, p. 8) defines the *g*-index "[. . .] as the highest number *g* of papers that together received g^2 or more citations". The range of *g*-index scores for a given group of authors will be either the same or more usually higher than their *h*-index scores so those with a number of highly cited articles will find almost certainly find themselves more highly visible than if only ranked by their *h*-index. Taking this a step further, Egghe (2006c) has extended the work of Glänzel and Persson (2005) who calculated the *h*-index of Price Medallists by calculating their *g*-index as well. The results show a difference between the two indexes with a noticeably higher *g*-index for all of the medallists; with the result better reflecting their lifetime achievement. Egghe (2006c, p. 143) also introduces the measure "*g/h*, i.e. the relative increase of *g* with respect to *h*" which gives some indication of the scale of difference between the indices for individual authors. Aware of the limitations of the *h*-index and the advantages of the *g*-index, Anderson *et al.* (2008) introduce a "tapered *h*-index" (h_T) which also scores all citations but in a way that allows for marginal and smooth increases from year to year. The authors take theoretical examples from Vinkler (2007) and show how in their view their index h_T is fairer than *h*. The authors then applied their h_T index in a practical application where they ranked six eminent scientists using their index and the *h*-index, the two were strongly correlated but the former showing a smoother transition between rankings from year to year and a higher index score.

Similarly Jin (2007), and Jin *et al.* (2007) introduce the *A*, *R* and *AR* indexes as variants of the *h*-index in an attempt, like the *g*-index, to take account of all citations that appear in the Hirsch core. Whilst the *A* index takes into account all citations, it is prone to distortion as indicated by Jin *et al.* (2007). To compensate for this, they propose that the square root of all the citations, which appear in the Hirsch core, be taken; this they call the *R* index. Further, they suggest the *AR* index, which is an adaptation of the *R* index taking into account both the citation intensity in the Hirsch core and the age of the publications in that core. Jin (2007, p. 6) defines the index as "the square root of the sum of the average number of citations per year of articles included in the *h*-core". Rousseau and Jin (2008) elaborate on this a stage further where they consider the AR^2-index introduced by Jin which takes the actual number of citations into account as well as their age and can decrease, unlike the *h*-index, in value. After some theoretical discussion of the properties of the AR^2-index the authors give a practical example of its application where the *h*-index, the *R*-index and the AR^2-index are shown together, most notably to illustrate how the AR^2-index declines while the other two indexes stagnate when modelled against Rousseau's own citation record.

Jarvelin and Persson (2008) also tackle the problem addressed by Jin (2007) with her *AR*-index, which, whilst compensating for the age of articles, has, they feel, too high an impact on relatively young articles, whose citation count, for example, would be halved if it were just two years old. To compensate for this and other issues, the authors introduce their Discounted Cumulative Impact index. Still using the *h*-core, or in fact any other publication data set, the authors propose that by devaluing older citations and rewarding authors for new citations to their work, in a smooth and parameterizable way and weighting citations by their "quality", however, measured, that a fairer assessment of a scholars work can be achieved.

Again giving more weight to highly cited articles, Kosmulski's (2006, p. 5) *h*(2) index is described as "A scientist's *h*(2) is defined as the highest natural number such that his *h*(2) most cited papers received each at least $[h(2)]^2$". Therefore, someone with a *h*(2) of ten would have ten papers each of which would have been cited at least 100 times. Kosmulski argues that this significantly reduces the number of articles in the core that is needed to make the calculation while still giving a relevant result. However, in their analysis of alternatives indexes Jin *et al.* (2007) regard the Kosmulski index as having poor sensitivity and only consider it briefly.

Rousseau and Ye (2008) introduce another variant of the *h*-index by proposing that it is more useful for hiring purposes than Hirsch's original version. Describing and defining their new index as time dependent, the two authors explain that it has three elements; "the size and contents of the *h*-core, the number of citations received, and the *h*-velocity" (Rousseau and Ye, 2008, p. 1854). The *h*-velocity reflects the recent increase in *h*. For two scientists with the same *h*-score and citation count within the *h*-core, Rousseau and Ye suggest they may be differentiated by their version of the index if one of the scientist's *h*-index was on the rise and the other had remained the same for some time. The scientist whose *h*-index is on the rise is, they suggest, probably the better choice for employment.

Wan *et al.* (n.d.), introduce the notion of a "pure" *h*-index which takes into account the number of co-authors and potentially their position in the byline as a way of more fairly apportioning credit between authors. Taking a more generalised approach, Sidiropoulos *et al.* (2007) introduce the possibility of five new variants of Hirsch's *h*-index. Two of these variants, the "contemporary" and "trend" *h*-indexes are attempts to reveal as they describe them as "brilliant young scientists and trendsetters".

3.6 Successive indexes

Prathap (2006) and Schubert (2006) saw the possibility of calculating a series of *h*-indexes at successively higher levels of aggregation. Egghe (2008b) examines this concept by presenting a theoretical model of successive *h*-indices and then showing in practice how Price Medallists can be ranked by their *h*-indexes at an individual level. As Egghe and Rao (2008, p. 1276) explain, taking for their example authors in an institution:

> If we rank these authors in decreasing order of their h_1 index, we can then apply the definition of the *h*-index to this ranked list and obtain the *h*-index h_2 of this group of authors (e.g., an institute, hereby giving a visibility of this institute).

Therefore, a department with an *h*-index of ten would have ten of its authors with at least an individual *h*-index of ten each. Egghe and Rao (2008) continue this line of reasoning by building successive layers of indexes for 167 authors publishing on the topic of "optical flow estimation". They ranked the number of publications and the citations to these publications that each of the authors in their sample had acquired. Not surprisingly, Egghe and Rao (2008) found that there was a strong correlation between the different rankings with the same authors appearing in the same ranks.

Taking this a step further, Schubert (2009) suggests that a Hirsch type index can be used to assess the impact of highly cited items by calculating the *h*-index of the citing papers. Using two journals, two samples of published items were taken from each, one set of the same age with different citation counts and the other set with the same citation count but of different ages. Schubert (2009) suggests that such an indicator has some striking features which are worthy of further analysis.

3.7 Comparisons of the different indices

Helpfully, using data from a sample of authors from biomedicine, Bornmann *et al.* (2008) have compared the various alternatives to the *h*-index; they examined the *h*-index and those of the different indices identified above, which they regard as the most important variants identified so far. They add to this the *m*-index proposed by themselves; this is the median number of citations received by papers in the Hirsch core.

Using data reported earlier (Bornmann and Daniel, 2005), which compared the *h*-index scores of applicants for post-doctoral research fellowships in biomedicine and their success or otherwise with the awarding trustees' assessments of their worth, Bornmann *et al.* (2008) using the various indices calculated the respective scores of the applicants and using factor analysis identified how these indices cluster. Importantly, the results of this and further statistical analysis indicate that there are two types of indices, which the authors describe as (2008, p. 836):

> The one type of indices (*h* index, *m* quotient, *g* index, and *h(2)* index) describe the most productive core of the output of a scientist and tell us the number of papers in the core.
>
> The other indices (*a* index, *m* index, *r* index, *ar* index, and the *hw* index) depict the impact of the papers in the core.

The authors found that the "impact of the productive core" was the better predictor of the trustees' decision on whether to award fellowships.

Costas and Bordons (2008) ask if the *g*-index is better than the than *h*-index at discriminating between scientists of different standing. They note that because the *h*-index is size dependent, those authors who publish fewer, but highly cited papers can be less favourably treated by the index than those who are more productive. Building a profile of their scientists, Costas and Bordons gathered a range of bibliometric indicators ranging from total number of documents, citation counts, median IF of the journals in which these scientists published. In conclusion, the authors find that the *g*-index is more sensitive than the *h*-index for those scientists who employ a more selective publishing strategy, but do also recognise the impact that a "big hit" can have on the index quite apart from them cautioning on using any single measure to evaluate performance.

Schreiber (2008a) compared the *g*-index with the *h*, *A* and *R* indexes for a group of 26 physicists. The results showed that the *g*-index allows for a better discrimination between academics, with some rearrangement of the rankings. This rearrangement was traced to different citation patterns with "one hit wonders" advancing in the *g*-listed rankings and some greater discrimination for those not so prominent academics. Whilst cautioning at over interpretation, Schreiber thinks the *g*-index more appropriate to rank the overall impact of a scientist's publications.

3.8 Comparisons across disciplines

It has always been difficult to assess, make comparisons or rank fairly across disciplines using bibliometric indicators; this includes using the *h*-index. Difficulties arise because of the different cultures in citation, self-citation, co-authorship and how the discipline disseminates its research, for example, journals, monographs or conferences proceedings. In an effort to deal with the problem of comparing scholars across disciplines and some of the perceived shortcomings of the *h*-index, Antonakis and Lalive (2008) propose the *IQp* index. Calculating the *IQp* index takes two steps, the first, for estimating "quality" requires calculating "the expected total number of citations [a] scholars work would receive if it were of average quality in the field" and "productivity" "is measured in 'adjusted papers' reflecting the number of papers published and the impact of the average paper in the author's top three subject areas" (Antonakis and Lalive, 2008, p. 959) A ratio is then derived from these two components and an index figure found. Taking Nobel laureates from economics, physics, chemistry, medicine and a group of psychologists the authors regarded as "great", the index was applied across three varied populations covering a sample of 80 scholars and was able to predict group membership.

4. Citation databases and the *h*-index

Calculation of the *h*-index for scholars cannot be separated from the citation databases that hold their citation history; as such, the coverage of the literature appropriate to the field of study by these databases can affect the accuracy of *h*-index scores. Deciding which database to use to calculate *h*-index scores would not have been an issue until recently, but with the advent of Scopus and GS, and in some disciplines other databases, there is now a choice (Neuhaus and Daniel, 2008; Jacsó, 2008b).

In an introductory paper, Jacsó (2008b) picking up Hirsch's (2005) point about the importance of database coverage, seeks to explain:

> [. . .] the essential practical requirements for creating *h*-index lists, and the considerations in interpreting the results, as they may be highly distorted due to shortcomings in the breadth (retrospectivity, depth) and consistency of coverage in the databases [. . .].

To show the variation that can occur when different databases are used, Jacsó (2008b) calculates Hirsch's own *h*-index using eight different databases and recorded a range of *h*-scores between 0 and 51. As Jacsó points out, even though EBSCO MegaFILE has 33 million master records, similar to that of Scopus, it only has 2 per cent of these records as a cited reference subset, thus severely limiting its use for the calculation of citation metrics. Whilst it may not be the primary concern of database suppliers to facilitate the calculation of *h*-indexes, those who are interested in calculating their own or others' *h*-index need to be aware of the limitations of the databases they use. Further, Jacsó (2008b) cautions on relying on *h*-indexes automatically generated from master records that are unlikely to deal effectively with those references that cannot be adequately matched and in addition, those references to books, book chapters, conference papers and journals not included in any particular database.

Following in this series, Jacsó (2008c) examines the use of GS to compute the *h*-index for individuals and journals. His overall and somewhat damning analysis suggests that while GS can be used for calculating *h*-indexes, it can only be done so when the data have been thoroughly checked and corroborated by extensive verification processes. This process is extremely time consuming and tedious and, as Jacsó notes, other authors, for example, Meho and Yang (2007) spent a "gruelling" 3,000 hours cleaning the data derived from the database. In a further two articles, Jacsó (2008d, e) examines the pros and cons of using Scopus and the WoS to compute the *h*-index. Both databases have notable characteristics, although Scopus (Jacsó, 2008d) is limited to having cited references available only from 1996, thus limiting its use for calculating the *h*-index of academics active before this period. Whilst this limitation is evident, what is less obvious is that Scopus disregards citations received after 1996 for work published before this date, thus limiting the database's utility for calculating any *h*-index (Jacsó, 2008d). However, Scopus does have good coverage for the period that it does cover with a broader range of journal and conference proceedings currently than WoS. WoS, as Jacsó (2008e) points out, does, however, have the greater historical coverage overall but this is dependent on the edition of the database that is used. Both Scopus and WoS offer quick and easy *h*-index calculators for the selected academic, although these do not take account of "orphan and stray" references (Jacsó, 2008d, e).

4.1 Database comparisons – choices and techniques

Harzing and van der Wal (2008) consider the value of GS as a source for citation analysis and draw a somewhat different conclusion to that of Jacsó (2008c). Describing the advantages and disadvantages of GS and WoS, the authors point to GS greater citation counts and the range of items from which it draws them, quite apart from the errors that they believe Jacsó made in his research. Harzing and van der Wal (2008) go on to compute the *h* and *g*-indexes for 20 top management journals using GS and compare this to the IF found for the same journals through WoS; the authors found a good correlation between IF and the two indexes. Extending this work and still using GS, Harzing and van der Wal (2009) derive *h*-scores for 838 journals in economics and business and compare this to their respective IF. The authors show that there is a strong correlation between IF and the GS *h*-index and by implication that those journals which do not have an official IF, can be just as readily ranked using the GS *h*-index.

In a review of the literature on the coverage of Scopus, GS and WoS, Bar-Ilan (2008a, b) found variable coverage and overlap between them depending on the topic being assessed. Calculating the *h*-indices of a number of highly cited Israeli scholars based on citation counts taken from these databases, Bar Ilan (2008a) found different results. The 37 scholars and three recent Israeli Nobel prize-winners examined had similar *h*-index scores when citation counts derived from WoS and Scopus were used in the calculation, but there was a marked difference between them and GS. The GS results for computer science and mathematics scholars had overall, higher *h*-scores, reflecting for computer science at least a better coverage of conference proceedings. Similarly, there were widely differing citation counts for the articles included in the *h*-core for each of the databases. Although *h* rankings were similar for Scopus and WoS, those generated from GS citations were less consistent, and the accuracy of some of the GS citations was uncertain.

Cronin and Meho (2006) calculated the *h*-index of a number of influential information scientists from faculty based in the USA; they accessed Institute for Scientific Information (ISI) citation indexes on dialog showing their results with and without self-citations. Oppenheim (2007) on the other hand used WoS to calculate the *h*-index for influential library and information scientist (LIS) in the UK; he counted citations to all authors but did not exclude self-citations. Both used cited reference searches, a more laborious process but more accurate than the general search facility which only includes citation counts from core journals indexed by ISI that do not have ambiguous citation records or citations outside of the core items indexed. Sanderson (2008), taking 26 of Oppenheim's LISs, compared these to those calculated using GS, and extending the work; used the "Author Finder" feature from WoS, and the "Author Identifier" feature from Scopus to calculate *h*-scores. Whilst Scopus and WoS *h*-scores were relatively close, those from GS were overall higher, indicating that its conference and book coverage was greater.

Although ranking by citation counts alone, Meho and Yang (2007) compared the citation coverage of Scopus, WoS and GS for publications from 15 US LIS faculty members. There were noticeable differences in the rate of citation, with Scopus finding 13.7 per cent more citations than WoS, and when these were combined, the total unique citation count rose by 35.1 per cent. Applying the revised citation count resulted to a change in the relative ranking for eight of the 15 scholars previously ranked by their total citation count using WoS. Most noticeably, those scholars in the middle rank had their relative rank altered the most by the combined count. Coverage by Scopus of conference proceedings was found to be superior to that of WoS. Meho and Yang (2007) then considered the citation counts that the 15 LIS scholars had using GS they found combining the collective unique citation counts increased the total citation count by 93.4 per cent. The authors found that most of the additional citations came from conference proceedings and lower impact journals, but even though considerable, these additional citations did not alter significantly the relative rankings of the scholars when added to either Scopus or WoS separately.

Meho and Rogers (2008) examined the ranking of 22 top human-computer interaction researchers from the UK. Identifying the 7,439 documents that had cited the work of these researchers, they found that Scopus had a greater coverage of them at 93 per cent compared to WoS at 54 per cent, the greater coverage being attributed to the more extensive indexing of conference proceedings by Scopus. The citation rankings derived from either database for the 22 researchers were significantly correlated. Using these two databases, the authors calculated the *h*-scores of the 22 researchers using system generated scores and manual counting methods. Both databases produced, not surprisingly, higher *h*-scores using the manual system, with Scopus, scores were noticeably greater, reflecting its broader coverage of conference proceedings. The authors also showed that combining the citation counts from Scopus and WoS did not significantly alter the *h* ranking of the 22 researchers. The implication is that it would be unnecessary to use both databases and laboriously combine the two citation counts.

The authors (Meho and Rogers, 2008) also compared on the grounds of free availability, a broader-based coverage and the generation of "manual" type counts, GS generated *h*-index scores with the union results from Scopus and WoS. The authors argue if there is consistent evidence of a high correlation between manually calculated:

"[. . .] scores in Scopus and/or Web of Science, one could potentially use Google Scholar as a possible alternative" (Meho and Rogers, 2008, p. 1724). Their results showed:

> [. . .] a very significant correlation between the *h*-index ranking in Google Scholar with that of the union of Scopus and Web of Science-Spearman Rank Order Correlation coefficient for the two rankings being 0.960 [. . .]

The *h*-scores from GS are higher than the union scores. This they attribute to the wider range of material, including books, book chapters, theses and reports, without any language or geographical limitations that GS covers. However, as the authors point out, interpretation of these higher *h*-scores needs to be done with caution as the source of these citations is not always of the same quality as those derived from commercial databases.

Although somewhat thin on research methodology, Saad (2006) also used WoS and GS to calculate the *h*-index for those who had published five or more papers in the *Journal of Consumer Research* during the period 1989-2005. He found as others, (Hirsch, 2005; Cronin and Meho, 2006) a good correlation between *h*-index and total citation counts for both of the databases, and also recorded a high correlation of 0.82 ($p < 0.001$) between the two *h*-indices from the databases. Results from the comparison showed that GS yielded 39 *h*-scores higher than WoS, with eight being equal and the remaining eight being higher for WoS. How the overall rankings were affected by these differences was unclear.

Van Leeuwen (2008) takes a range of Dutch scientists from a number of disciplines and calculates several bibliometric indicators and compares these to their *h*-scores. He found that there were noticeable differences between the disciplines in terms of their productivity and their *h*-index as well as noting that age did not seem to be correlated with *h*-index scores. In all, Van Leeuwen is cautious about the use of the *h*-index in research assessment at an individual level.

Jacsó (2008a) examines the lifetime achievement of a single scholar using the occasion to examine how effective GS, Scopus and WoS are in measuring that achievement, by manual calculation, rather than by using the automatic features of these databases to establish F.W. Lancaster's lifetime *h*-index. Predictably perhaps GS comes out rather badly from Jacso's analysis given the general inadequacies he has identified in earlier work, some of which is noted above. Scopus on the other hand is more successful but is severely hampered by it only indexing cited references from 1996 and therefore is unable to give a reliable lifetime *h*-index for Lancaster, given his publishing career started well before this date. Limitations by Scopus in its coverage of library and information journals also hinders an accurate coverage of the Lancaster's work. Overall, WoS delivers the most comprehensive citation record on which to base an adequate *h*-index for Lancaster. But this was only established after painstaking analysis of the cited references which doubled Lancaster's initial *h*-index of 13-26 the former obtained through the automated feature provided by WoS for the calculation of an individual's *h*-index.

5. Library and information management case studies

Cronin and Meho (2006) took for their sample mainly mid-to-late career information science scholars from the USA. Using Dialog to access the Thomson citation indexes, they counted the total citations of these scholars and calculated their *h*-index both with and without self-citations. It appears that only first author self-citations were removed. The result from a Spearman Rank Order Correlation found a strong correlation between citation ranks and *h*-index (without self-citations) of 0.9, significant at the 0.01 level. The overall result showed total citations to be a reliable indicator of impact but Cronin and Meho (2006) further suggest that the *h*-index gave additional discriminatory power. For example, they noted that some authors, whilst having greater citation counts than their colleagues, had their ranks reversed when ranked by their *h*-index; similarly, some authors who had shorter in-field times had a very similar *h*-index when compared to some of their longer serving fellows. The removal of self-citations did little to alter the overall rankings but did reduce 18 out of the 31 *h*-index scores by between one and three.

In a similar study, Oppenheim (2007) examined the *h*-index rankings of senior UK library and information science academics. The majority of them were still active, but the sample of 84 did include 13 retired scholars and for comparison, Eugene Garfield (the founder of modern citation studies). Using citation counts from the WoS, including self-citations he initially ranked active senior academics who had a *h*-index of 5 or above. A high scoring individual was identified with a *h*-index of 31, well outside the next highest score of 18 and higher than the highest score found by Cronin and Meho (2006) above. Oppenheim (2007) attributes this high score to the specialist nature of the subject of this particular academic (cheminformatics).

In a follow up to the work of Oppenheim (2007), Sanderson (2008) used GS to count citations. Taking the 26 active academics who Oppenheim (2007) had identified as having a *h*-index above five, Sanderson added a further two authors and computed a revised *h*-index. Differences were recorded, mostly an increase in *h*-scores being on average 3.4 points higher. The citations counted, were not limited to the 1992-2005 time frame that Oppenheim (2007) had used; rather all relevant citations were counted from the full range of years that GS covered. Sanderson (2008) found, however, from a sample of his authors that less than 10 per cent of the pre 1992 publications contributed to the *h*-index score, adding on average 1.1 to the *h*-score.

Sanderson, like Oppenheim (2007), did not remove self-citations arguing that these made little difference to the overall rankings. Noticeably, however, *h*-scores were higher when using citations from GS, in part from the longer time frame and from the greater range of materials covered. An important difference was noted by Sanderson for five scholars who published more in computer science. Their *h*-scores were higher when citations found in GS to conference proceeding were added. These publications he reasoned were not adequately covered by WoS. Finally, Sanderson recalculated the *h*-index scores using the WoS "author finder", and the Scopus "author identifier" services comparing these to the original GS scores and those from Oppenheim (2007) finding, not surprisingly a range of results, dependent on database coverage. Sanderson suggests that the largest score achieved for each individual from any one of the methods used might be the most appropriate measure of their *h*-index.

Examining the *h*-index of Price Medallists, Glänzel and Persson (2005) speculated on what might be a reasonable *h*-score for those still active in this particular field of study. Selecting only work related to the field, they found a range of *h*-scores of nine to 17 using the WoS for publications from 1986 onward. They thought anyone with a *h*-index of 10 or above might be considered a suitable candidate for the award, given that latterly medallists have tended to have lower *h*-scores than their predecessors. Following up this work, Bar-Ilan (2006) calculated the *h*-index of these same medallists but used GS instead of WoS. Noting the difficulty of using GS, she found very similar *h*-scores for both databases, although searches did go back as far as 1986 which may be a little too early for the records that GS indexes. This result is at odds with the foregoing studies, which recorded overall higher *h*-scores when compared to Scopus and WoS.

6. Broader applications

The *h*-index cannot only be applied to individuals. It has been applied by Banks (2006) to differentiate between "hot" and "older" topics or compounds by carrying out a topic or compound search using WoS. The resulting articles can then be ranked by their citation count and then a *h*-score found as for individuals, dividing this result by the number years since the compound or topic was first mentioned gives an indication of its popularity. Other authors have suggested an *h*-type index for journals, Braun *et al.* (2006) examined, with the exception of two review journals, those journals with the highest IFs from 2001 and compared their rankings to their *h*-index and showed that there is a strong correlation between JIFs and the *h*-index. Bornmann *et al.* (2009) took 20 organic chemistry journals and calculated their JIF and compared these to a battery of *h* variants derived from the same citation and article data. The authors found a range of results with a strong correlation between the rankings suggesting that there is some redundancy amongst them as they appear to measuring similar aspects of scientific performance. Vanclay (2008) compared the results of an expert ranking of forestry journals with JIFs and with *h*-indices calculated by using WoS and Harzing's Publish or Perish (PoP) (2008)[5] which harvests its data from GS. Amongst his findings, Vanclay found a strong correlation between the IF of journals and their *h*-index, both when calculated using WoS ($r >= 0.88$) and when using PoP ($r >= 0.84$). In addition to this, it was evident that using the *h*-index to rank journals in conjunction with PoP could help identify candidate journals for inclusion in the WoS. In another exercise, Bador and Lafouge (2009) examined the correlation between the JIF for two samples of pharmacology and pharmacy and psychiatry journals drawn from the WoS Journal Citation Reports with their *h*-index. Of the two journal sets there was a higher correlation for the psychiatry journals with their *h*-index than for the pharmacology and pharmacy journals with albeit a much coarser *h*-index scale than that offered by JIF's.

Evaluation of post graduate research programmes in Brazil has been traditionally by peer review. However, using an institutional *h*-index, and traditional bibliometric indicators Pires da Luz *et al.* (2008) ranked six Brazilian postgraduate psychiatry programmes and found that they were similarly ranked by both and evident was a close correlation between the institutional *h*-index and most of the bibliometric indicators they used.

Taking Turkey, Croatia, European Union countries and six other countries including the USA and China, Csajbók *et al.* (2007) take a novel approach to ranking them by their *h*-index in differing scientific fields. Rather than counting all the citation counts directly, the authors approximated the *h*-index of these countries by subject area using the WoS's Essential Science Indicators database. This allowed the authors to access the highly cited papers feature of the database and by sequential examination was able to rank each country by their *h*-index within 19 different subject areas. Guan and Gao (2009) applied the *h*-index to patents. Collecting semiconductor patent records between the years 1996 and 2005 and the citations made to them by other patents, the authors were able to identify and rank the top 20 companies by their *h*-index just as if they were academics, i.e. a *h*-index of 20 requires 20 patents granted to a company to receive at least 20 patent citations each. As in other research, there was a strong correlation between citation counts and the *h*-index rankings and characteristically, the *h*-index balances quantity, the number of patents against quality, the number of citations.

In an intriguing application, Liu and Rousseau (2009) show how they used the *h*, *g* and *R* indexes to assess the rankings of library classification categories by book loans from main and sub-branches of a university library in China. This was achieved by using the library classification categories as "authors" and the book loans as "citations" and then counting, for comparison, the issues from reading rooms and the library. The ranked results were compared between the indexes and a high correlation was found between them suggesting in this case that the *h*-index alone was adequate for this application and suggested that newer books have a higher *h*-index than older ones.

A recent report from the Research Information Network (2009) has some surprising insights into the information seeking behaviour of scholars. The authors of the report, by deep log analysis, show how in the life sciences that there is a "strong negative correlation between the research rating of life scientists in each institution, as measured by the Hirsh index, and the length of their sessions" (Research Information Network, 2009, p. 33) in ScienceDirect, a database of journal articles, books and online tutorials. In effect, the higher the collective *h*-index of the life scientists at each institution the shorter their session length in the database. Although it is not quite clear how the authors identified the scholars in question or how the *h*-score was calculated, it is assumed the *h*-score was based on the collective citation counts for the life scientists at each institution. At a broader scale still, the authors rank ten institutions on a research rating scale using the *h*-index and show that the more highly ranked the institution the more likely they will have accessed ScienceDirect, through a gateway such as Google or GS.

7. Conclusion and implications for research

In a very short time, Hirsch's *h*-index has become established feature for those interested in the evaluation of science and has become a research subject in its own right. It is already being used as an indicator of research impact by individuals who wish to use it, amongst other indicators, to show their standing in their discipline. Like any indicator of performance, the *h*-index has its advantages and disadvantages and these have led to a range of variants and applications being developed. In particular, to compensate those individual scholars whose high citation counts are discounted by the *h*-index, a number of variants and refinements have been suggested and these tend toward compensating, perhaps more fairly, those individuals who have many highly cited items. It has to noted, however, that whilst these numerous alternative indexes attempt, and sometimes succeed, in addressing these shortcomings in the *h*-index, they are all more complex (sometimes significantly so) to calculate than the *h*-index, whose simplicity is likely to ensure it remains the metric of choice for the bibliometric assessment of individuals.

Nonetheless, the *h*-index has proved, in its short lifetime, to be a versatile indicator. It can, and has been used to rank journals, hot topics, individuals, departments and companies by citations, in this case, to their patents and has the potential to be applied imaginatively to other areas where rankings may be useful. The index can also be applied to almost any level of aggregation, from single to groups of scholars, to departments, to faculties, to universities or indeed to countries. In so doing, it lends itself to far more applications than other bibliometric indicators and is perhaps more readily accessible and understood by a wider audience outside of academia. Added to this the index is robust.

In use, care needs to be exercised in the way that the *h*-index and its variants are applied, including careful selection of the database from which citations are counted to ensure at whatever level of aggregation, the individual, subject or topic that it is adequately covered by that database. Similarly, it is apparent that particular care needs to exercised at the individual level to ensure as far as possible that citations are identified for all types of publications that the author may have produced. This is particularly important in subjects where monographs and edited contributions to them are important to the subject. At aggregated levels of use, the *h*-index is little affected by self-citation rates and overall rankings are generally little affected. It is apparent, though, that there are some very high self-citers, sometimes in particular subjects, and if comparative exercises are undertaken between scholars, where preferment may be an issue, proper note should be taken of self-citation rates, even though these may be perfectly legitimate.

A number of studies have shown the utility of the *h*-index and some of these have used elements of peer review to show that the index does correlate well with assessment by peers. There is still, however, a need to undertake more studies at various levels of aggregation to show that the *h*-index or its variants are reliable indicators of performance, even if this highlights their limitations as well as their strengths.

Notes

1. The Price Medal is periodically awarded by the journal *Scientometrics* for outstanding contributions in the field of quantitative studies of science.
2. The *w*-index is determined as follows. Work through the list of publications in INCreasing order of citations (as $xn \leqslant xn - 1 \leqslant \cdots \leqslant x1$), and keep a counter ctr that is initialized at 0. Every time you move to a new xk, compare it against the current value of the counter. If ctr $<$ xk, then increase ctr, otherwise do nothing. Once you reach your top publication x1, the value of the counter contains your *w*-index (Woeginger, 2008a, b, c).
3. Crown indicator compares the mean number of citations per publication for these research groups divided by the mean of the citations per publication on a worldwide basis for the journals in the same subject area, using the same time window and document type, both without self-citations. A score greater than one for this measure indicates a higher than average citation score for publications in that field, and hence greater impact.
4. Rousseau (n.d.) defines the "*Hirsch core*" as the first *h* papers, although more generally understood to be those items whose citations are included in an author's *h*-index.
5. PoP is a software program that retrieves and analyses academic citations. It uses GS to obtain the raw citations, then analyses these and presents a range of bibliometric indicators.

References

Alonso, S., Cabrerizo, F.J., Herrera-Viedma, E. and Herrera, F. (2009), "*h*-Index: a review focused in its variants, computational and standardization for different scientific fields", *Journal of Informetrics*, Vol. 3 No. 4, pp. 273-89.

Anderson, T.R., Robin, K.S. and Killworth, P.D. (2008), "Beyond the Durfee square: enhancing the *h*-index to score total publication output", *Scientometrics*, Vol. 76 No. 3, pp. 577-88.

Antonakis, J. and Lalive, R. (2008), "Quantifying scholarly impact: IQp versus the Hirsch h", *Journal of the American Society for Information Science and Technology*, Vol. 59 No. 6, pp. 956-69.

Bador, P. and Lafouge, T. (2009), "Comparative analysis between impact factor and *h*-index for pharmacology and psychiatry journals", *Scientometrics*, Vol. 84 No. 1, pp. 65-79.

Baird, L.M. and Oppenheim, C. (1994), "Do citations matter?", *Journal of Information Science*, Vol. 20 No. 1, pp. 2-15.

Ball, P. (2005), "Index aims for fair ranking of scientists", *Nature*, Vol. 436 No. 900.

Banks, M.G. (2006), "An extension of the Hirsch index: indexing scientific topics and compounds", *Scientometrics*, Vol. 69 No. 1, pp. 161-8.

Bar-Ilan, J. (2006), "H-index for price medalists revisited", *ISSI Newsletter*, Vol. 2 No. 1, pp. 3-5.

Bar-Ilan, J. (2008a), "The *h*-index and of other informetric topics", *Scientometrics*, Vol. 75 No. 3, pp. 591-605.

Bar-Ilan, J. (2008b), "Which *h*-index? – a comparison of WoS, Scopus and Google scholar", *Scientometrics*, Vol. 74 No. 2, pp. 257-81.

Bornmann, L. and Daniel, H.-D. (2005), "Does the *h*-index for ranking scientists really work?", *Scientometrics*, Vol. 65 No. 3, pp. 391-2.

Bornmann, L. and Daniel, H.-D. (2007), "What do we know about the *h* index?", *Journal of the American Society for Information Science and Technology*, Vol. 58 No. 9, pp. 1381-5.

Bornmann, L., Mutz, R. and 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*, Vol. 59 No. 5, pp. 830-7.

Bornmann, L., Werner, M. and Schier, H. (2009), "Hirsh-type index values for organic chemistry journals: a comparison of new metrics with the journal impact factor", *European Journal of Organic Chemistry*, Vol. 10, pp. 1471-6.

Braun, T., Glänzel, W. and Schubert, A. (2006), "A Hirsch-type index for journals", *Scientometrics*, Vol. 69 No. 1, pp. 169-73.

Brittain, J.M. (2000), "A highly visible scientist: Jack Meadows", *Journal of Information Science*, Vol. 26 No. 4, pp. 267-72.

Costas, R. and Bordons, M. (2007), "The *h*-index: advantages, limitations and its relation with other bibliometric indicators at the micro level", *Journal of Informetrics*, Vol. 1 No. 3, pp. 193-203.

Costas, R. and Bordons, M. (2008), "Is the *g*-index better than the *h*-index? An exploratory study at the individual level", *Scientometrics*, Vol. 77 No. 2, pp. 267-88.

Cronin, B. and Meho, L. (2006), "Using the *h*-index to rank influential information scientists", *Journal of the American Society for Information Science and Technology*, Vol. 57 No. 9, pp. 1275-8.

Csajbók, E., Berhidi, A., Vasaa, L. and Schubert, A. (2007), "Hirsch-index for countries based on essential science indicators data", *Scientometrics*, Vol. 73 No. 1, pp. 91-117.

Egghe, L. (2006a), "An improvement of the *h*-index: the g-index", *ISSI Newsletter*, Vol. 2 No. 1, pp. 8-9.

Egghe, L. (2006b), "An informetric model for the Hirsch-index", *Scientometrics*, Vol. 69 No. 1, pp. 121-9.

Egghe, L. (2006c), "Theory and practise of the *g*-index", *Scientometrics*, Vol. 69 No. 1, pp. 131-52.

Egghe, L. (2008a), "Mathematical theory of the *h*- and *g*-index in case of fractional counting of authorship", *Journal of the American Society for Information Science and Technology*, Vol. 59 No. 10, pp. 1608-16.

Egghe, L. (2008b), "Modelling successive *h*-indexes", *Scientometrics*, Vol. 77 No. 3, pp. 377-87.

Egghe, L. and Rao, I.K. (2008), "Study of different h-indices for groups of authors", *Journal of the American Society for Information Science and Technology*, Vol. 59 No. 8, pp. 1276-81.

Egghe, L. and Rousseau, R. (2008), "An *h*-index weighted by citation impact", *Information Processing & Management*, Vol. 44 No. 2, pp. 770-80.

Egghe, L., Liang, L. and Rousseau, R. (2009), "A relation between *h*-index and impact factor in the power-law model", *Journal of the American Society for Information Science and Technology*, Vol. 60 No. 11, pp. 2362-5.

Engqvist, L. and Frommen, J.G. (2008), "The *h*-index and self-citations", *Trends in Ecology and Evolution*, Vol. 23 No. 5, pp. 250-2.

Glänzel, W. (2006a), "On the *h*-index – a mathematical approach to a new measure of publication activity and citation impact", *Scientometrics*, Vol. 67 No. 2, pp. 315-21.

Glänzel, W. (2006b), "On the opportunities and limitations of the *H*-index", available at: http://eprints.rclis.org/archive/00009535/01/H_Index_opprtunities.pdf (accessed 26 June 2008).

Glänzel, W. (2008a), "*H*-index concatenation", *Scientometrics*, Vol. 77 No. 2, pp. 369-72.

Glänzel, W. (2008b), "On some new bibliometric applications of statistics related to the *h*-index", *Scientometrics*, Vol. 77 No. 1, pp. 187-96.

Glänzel, W. and Persson, O. (2005), "*H*-index for Price medallists", *ISSI Newsletter*, Vol. 1 No. 4, pp. 15-18.

Guan, J.G. and Gao, X. (2009), "Exploring the *h*-index at patent level", *Journal of the American Society for Information Science and Technology*, Vol. 60 No. 1, pp. 35-40.

Harzing, A. (2008), "Reflections on the *h*-index", available at: www.harzing.com/resources.htm#/pop_hindex.htm (accessed 30 June 2008).

Harzing, A. and van der Wal, R. (2008), "Google Scholar as a new source for citation analysis", *Ethics in Science and Environmental Politics*, Vol. 8 No. 1, pp. 61-73.

Harzing, A. and van der Wal, R. (2009), "Google Scholar *h*-index for journals: an alternative metric to measure journal impact factors in economics and business", *Journal of the American Society for Information Science and Technology*, Vol. 60 No. 1, pp. 41-6.

Hirsch, J.E. (2005), "An index to quantify an individual's scientific research output", *Proceedings of the National Academy of Sciences*, Vol. 102 No. 46, pp. 16569-72.

Hirsch, J.E. (2007), "Does the *h* index have predictive power?", *Proceedings of the National Academy of Science*, Vol. 104 No. 49, pp. 19193-8.

Jacsó, P. (2008a), "Testing the calculation of a realistic *h*-index in Google Scholar, Scopus and Web of Science for F.W. Lancaster", *Library Trends*, Vol. 56 No. 4, pp. 784-815.

Jacsó, P. (2008b), "The plausibility of computing the *h*-index of scholarly productivity and impact using reference-enhanced databases", *Online Information Review*, Vol. 32 No. 2, pp. 266-83.

Jacsó, P. (2008c), "The pros and cons of computing the *h*-index using Google Scholar", *Online Information Review*, Vol. 32 No. 3, pp. 437-52.

Jacsó, P. (2008d), "The pros and cons of computing the *h*-index using Scopus", *Online Information Review*, Vol. 32 No. 4, pp. 524-35.

Jacsó, P. (2008e), "The pros and cons of computing the *h*-index using Web of Science", *Online Information Review*, Vol. 32 No. 5, pp. 673-88.

Jarvelin, K. and Persson, O. (2008), "The DCI index: discounted cumulated impact-based research evaluation", *Journal of the American Society for Information Science and Technology*, Vol. 59 No. 9, pp. 1433-40.

Jin, B. (2007), "The AR-index complementing the *h*-index", *ISSI Newsletter*, Vol. 3 No. 1, p. 6.

Jin, B., Liang, L., Rousseau, R. and Egghe, L. (2007), "The *R*- and *AR*-indices: complementing the *h*-index", *Chinese Science Bulletin*, Vol. 52 No. 6, pp. 855-63.

Kalyane, V.L., Kumar, V., Prakasam, E.R. and Swama, T. (2004), "Eponymous citations to Homi Jhangar Bhaba", *Malaysian Journal of Library and Information Sciences*, Vol. 9 No. 1, pp. 81-99.

Kelly, C.D. and Jennions, M.D. (2006), "The *h* index and career assessment by numbers", *Trends in Ecology and Evolution*, Vol. 21 No. 4, pp. 167-70.

Kosmulski, M. (2006), "A new Hirsch-type index saves time and works equally well as the original *h*-index", *ISSI Newsletter*, Vol. 2 No. 3, pp. 4-6.

Liu, Y. and Rousseau, R. (2009), "Properties of Hirsch-type indices: the case of library classification categories", *Scientometrics*, Vol. 79 No. 2, pp. 235-48.

Meho, L.I. and Rogers, Y. (2008), "Citation counting, citation ranking, and *h*-index of human-computer interaction researchers: a comparison between Scopus and Web of Science", *Journal of the American Society for Information Science and Technology*, Vol. 59 No. 11, pp. 1711-26.

Meho, L.I. and Yang, K. (2007), "Impact of data sources on citation counts and ranking of LIS faculty: Web of Science versus Scopus and Google Scholar", *Journal of the American Society for Information Science and Technology*, Vol. 58 No. 13, pp. 2105-25.

Neuhaus, C. and Daniel, H.-D. (2008), "Data sources for performing citation analysis: an overview", *Journal of Documentation*, Vol. 64 No. 2, pp. 193-210.

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*, Vol. 58 No. 2, pp. 297-301.

Pires da Luz, M., Marques-Portella, C., Gleiser, S., Silva Freire Coutinho, E. and Figueira, I. (2008), "Institutional *h*-index: the performance of a new metric in the evaluation of Brazilian psychiatric post-graduation programs", *Scientometrics*, Vol. 77 No. 2, pp. 361-8.

Prathap, G. (2006), "Hirsch-type indices for ranking institutions' scientific research output", *Current Science*, Vol. 91 No. 11, p. 1439.

Publish or Perish (2008), available at: www.harzing.com/resources.htm (accessed 30 June 2008).

Research Information Network (2009), *E-Journals: Their Use, Value and Impact*, Research Information Network, London.

Rousseau, R. (2007), "The influence of missing publications on the Hirsch index", *Journal of Informetrics*, Vol. 1 No. 1, pp. 2-7.

Rousseau, R. (2008), "Woeginger's axiomatisation of the *h*-index and its relation to the *g*-index, the $h^{(2)}$-index and the R^2-index", *Journal of Informetrics*, Vol. 2 No. 4, pp. 335-40.

Rousseau, R. (n.d.), "New developments related to the Hirsch index", available at: http://sci2s.ugr.es/impact/Hirsch_new_developments.pdf (accessed 25 June 2008).

Rousseau, R. and Jin, B. (2008), "The age-dependent *h*-type AR^2-index: basic properties and a case study", *Journal of the American Society for Information Science and Technology*, Vol. 59 No. 14, pp. 2305-11.

Rousseau, R. and Ye, F.R. (2008), "A proposal for a dynamic *h*-type index", *Journal of the American Society for Information Science and Technology*, Vol. 59 No. 11, pp. 1853-5.

Saad, G. (2006), "Exploring the *h* index at the author and journal levels using bibliometric data of productive consumer scholars and business-related journals respectively", *Scientometrics*, Vol. 69 No. 1, pp. 117-20.

Sanderson, M. (2008), "Revisiting *h* measured on UK LIS academics", *Journal of the American Society for Information Science*, Vol. 59 No. 7, pp. 1184-90.

Schreiber, M. (2007), "Self-citation corrections for the Hirsch index", *European Physics Letters*, Vol. 78 Art. No. 30002, pp. 1-6.

Schreiber, M. (2008a), "An empirical investigation of the *g*-index for 26 physicists in comparison with the *h*-index, the A-index, and the R-index", *Journal of the American Society for Information Science and Technology*, Vol. 59 No. 9, pp. 1513-22.

Schreiber, M. (2008b), "The influence of self-citation corrections on Egghe's *g*-index", *Scientometrics*, Vol. 76 No. 1, pp. 187-200.

Schreiber, M. (2008c), "To share the fame in a fair way, h_m modifies h for multi-authored manuscripts", *New Journal of Physics*, Vol. 10, available at: www.njp.org (accessed 24 June 2008).

Schubert, A. (2006), "Successive *h*-indexes", *Scientometrics*, Vol. 70 No. 1, pp. 201-5.

Schubert, A. (2009), "Using the *h*-index for assessing single publications", *Scientometrics*, Vol. 78 No. 3, pp. 559-65.

Sidiropoulos, A., Katsaros, D. and Manolopoulos, Y. (2007), "Generalised Hirsch *h*-index for disclosing latent facts in citation networks", *Scientometrics*, Vol. 72 No. 2, pp. 253-80.

Snyder, H. and Bonzi, S. (1998), "Patterns of self-citation across disciplines (1980-1989)", *Journal of Information Science*, Vol. 24 No. 6, pp. 431-5.

Symonds, M.R.E., Gemmell, N.J., Braisher, T.L., Gorringe, K.L. and Elgar, M.A. (2006), "Gender differences in publication output: towards an unbiased metric of research performance", *PloS ONE*, Vol. 1, available at: www.plosone.org/article/info:doi/10.1371/journal.pone.0000127 (accessed 12 May 2009).

Vanclay, J.K. (2007), "On the robustness of the *h*-index", *Journal of the American Society for Information Science and Technology*, Vol. 58 No. 10, pp. 1547-50.

Vanclay, J.K. (2008), "Ranking forestry journals using the *h*-index", *Journal of Informetrics*, Vol. 2 No. 4, pp. 326-34.

Van Eck, N. and Waltman, L. (2008), "Generalising the *h*- and *g*-indexes", *Journal of Informetrics*, Vol. 2 No. 4, pp. 263-71.

Van Leeuwen, T. (2008), "Testing the validity of the Hirsch-index for research assessment purposes", *Research Evaluation*, Vol. 17 No. 2, pp. 157-60.

Van Raan, A.F.J. (2006), "Comparison of the Hirsch index with standard bibliometric indicators and with peer judgement for 147 chemistry research groups", *Scientometrics*, Vol. 67 No. 3, pp. 391-502.

Vinkler, P. (2007), "Eminence of scientists in the light of the *h*-index and other scientometric indicators", *Journal of Information Science*, Vol. 33 No. 4, pp. 481-91.

Wan, J., Hua, P. and Rousseau, R. (n.d.), "The pure *h*-index: calculating an author's *h*-index by taking co-authors into account", available at: http://eprints.rclis.org/archive/00011401/01/pure_h.pdf (accessed 16 July 2008).

Woeginger, G.J. (2008a), "A symmetry axiom for scientific impact indices", *Journal of Informetrics*, Vol. 2 No. 4, pp. 298-303.

Woeginger, G.J. (2008b), "An axiomatic analysis of Egghe's *g*-index", *Journal of Informetrics*, Vol. 2 No. 4, pp. 364-8.

Woeginger, G.J. (2008c), "An axiomatic characterization of the Hirsch-index", *Mathematical Social Sciences*, Vol. 56 No. 2, pp. 224-32.

Wu, Q. (2008), "The *w*-index: a significant improvement of the *h*-index", available at: http://arxiv.org/ftp/arxiv/papers/0805/0805.4650.pdf (accessed 12 December 2008).

Further reading

Bornmann, L. and Daniel, H.-D. (2009), "The state of *h* index research", *EMBO*, Vol. 10 No. 1, pp. 2-6.

Bornmann, L., Wallon, G. and Ledin, A. (2008), "Is the *h* index related to (standard) bibliometric measures and to the assessments by peers? An investigation of the *h* index by using molecular life sciences data", *Research Evaluation*, Vol. 17 No. 2, pp. 149-56.

Egghe, L. (2007), "Dynamic *h*-index: the Hirsch index in function in time", *Journal of the American Society for Information Science and Technology*, Vol. 58 No. 3, pp. 452-4.

Egghe, L. and Rousseau, R. (2006), "An informetric model for the *h* index", *Scientometrics*, Vol. 69 No. 1, pp. 121-9.

Iglesias, J.E. and Pecharroman, C. (2007), "Scaling the *h*-index for different scientific ISI fields", *Scientometrics*, Vol. 73 No. 3, pp. 303-20.

Moed, H. (2005), *Citation Analysis in Research Evaluation*, Springer, London.

Norris, M. and Oppenheim, C. (2003), "Citation counts and the Research Assessment Exercise V: Archaeology and the 2001 RAE", *Journal of Documentation*, Vol. 59 No. 6, pp. 709-30.

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*, Vol. 51 No. 1, pp. 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*, Vol. 53 No. 5, pp. 477-87.

Corresponding author

Charles Oppenheim can be contacted at: c.oppenheim@lboro.ac.uk

This article has been cited by:

1. Muzammil Tahira, Rose Alinda Alias, Arayti Bakri, Ani Shabri. 2014. h-index, h-type Indices, and the Role of Corrected Quality Ratio. *Journal of Information Science Theory and Practice* **2**, 20-30. [CrossRef]
2. Ana Paula dos Santos Rubem, Ariane Lima de Moura, João Carlos Correia Baptista Soares de Mello. 2014. Comparative analysis of some individual bibliometric indices when applied to groups of researchers. *Scientometrics* . [CrossRef]
3. Cameron Barnes. 2014. The emperor's new clothes: the h -index as a guide to resource allocation in higher education. *Journal of Higher Education Policy and Management* **36**, 456-470. [CrossRef]
4. Denis Bouyssou, Thierry Marchant. 2014. An axiomatic approach to bibliometric rankings and indices. *Journal of Informetrics* **8**, 449-477. [CrossRef]
5. Gerhard J. Woeginger. 2014. Investigations on the step-based research indices of Chambers and Miller. *Journal of Informetrics* **8**, 659-666. [CrossRef]
6. Star X. Zhao, Paul L. Zhang, Jiang Li, Alice M. Tan, Fred Y. Ye. 2014. Abstracting the core subnet of weighted networks based on link strengths. *Journal of the Association for Information Science and Technology* **65**:10.1002/asi.2014.65.issue-5, 984-994. [CrossRef]
7. Lutz Bornmann, Werner Marx. 2014. How to evaluate individual researchers working in the natural and life sciences meaningfully? A proposal of methods based on percentiles of citations. *Scientometrics* **98**, 487-509. [CrossRef]
8. Christopher Walton, Anne Morris. 2013. A bibliometric study of taxonomic botany. *Journal of Documentation* **69**:3, 435-451. [Abstract] [Full Text] [PDF]
9. Mauricio Romero-Torres, Luis Alberto Acosta-Moreno, María Alejandra Tejada-Gómez. 2013. Ranking de revistas científicas en Latinoamérica mediante el índice <i>h</i>: estudio de caso Colombia. *Revista española de Documentación Científica* **36**:10.3989/redc.2013.v36.i1, e003. [CrossRef]
10. Mu-Hsuan Huang. 2012. Exploring the h-index at the institutional level. *Online Information Review* **36**:4, 534-547. [Abstract] [Full Text] [PDF]
11. Jerome K. Vanclay, Lutz Bornmann. 2012. Metrics to evaluate research performance in academic institutions: a critique of ERA 2010 as applied in forestry and the indirect H2 index as a possible alternative. *Scientometrics* **91**, 751-771. [CrossRef]
12. Panagiota Altanopoulou, Maria Dontsidou, Nikolaos Tselios. 2012. Evaluation of ninety-three major Greek university departments using Google Scholar. *Quality in Higher Education* **18**, 111-137. [CrossRef]
13. Lutz Bornmann, Werner Marx. 2012. HistCite analysis of papers constituting the h index research front. *Journal of Informetrics* **6**, 285-288. [CrossRef]
14. K. Jonkers, G.E. Derrick. 2012. The bibliometric bandwagon: Characteristics of bibliometric articles outside the field literature. *Journal of the American Society for Information Science and Technology* **63**, 829-836. [CrossRef]
15. John S. Liu, Louis Y.Y. Lu. 2012. An integrated approach for main path analysis: Development of the Hirsch index as an example. *Journal of the American Society for Information Science and Technology* **63**:10.1002/asi.v63.3, 528-542. [CrossRef]
16. Ludo Waltman, Nees Jan van Eck. 2012. The inconsistency of the h-index. *Journal of the American Society for Information Science and Technology* **63**:10.1002/asi.v63.2, 406-415. [CrossRef]

17. Khaled El Emam, Luk Arbuckle, Elizabeth Jonker, Kevin Anderson. 2012. Two h-Index Benchmarks for Evaluating the Publication Performance of Medical Informatics Researchers. *Journal of Medical Internet Research* **14**, e144. [CrossRef]
18. Lutz Bornmann, Rüdiger Mutz, Sven E. Hug, Hans-Dieter Daniel. 2011. A multilevel meta-analysis of studies reporting correlations between the h index and 37 different h index variants. *Journal of Informetrics* **5**, 346-359. [CrossRef]