Pertanika J. Sci. & Technol. 20 (2): 197 - 203 (2012)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Short Communications

The *h*-index in Elsevier's *Scopus* as an Indicator of Research Achievement for Young Malaysian Scientists

Yap, C. K.

Department of Biology, Faculty of Science, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

HOW TO MEASURE A SCIENTIST'S RESEARCH PERFORMANCE OUTPUT?

If we were given a questionnaire of "How do we measure a researcher as a true scientist? with optional answers like (a) Having a good number of publications, (b) having attending numerous conferences, (c) with a high popularity as always appeared in mass media, and (d) good international networking and good public relations. Options (c) and (d) always come later after option (a) has been achieved, while option (b) can be simply achieved or abstract be accepted for presentation in any conference. Hitherto, publishing in any peer-reviewed journals carry a certain quality since they are highly subjected to peer review evaluation before the paper can be accepted for publication in a journal. Needless to say, those constructive comments given by the reviewers are very crucial in shaping our scientific understanding in our subject area rather than rejection experience (Yap, 2009). Having said so, option (a) will definitely be the best answer. The fact is that option (a) should not be argued whatsoever as the best answer [since publications speaks louder than anything else] and options (a), (b) and (c) are supplementary criteria to option (a) but they are not as vital as option (a).

When we are asked 'What is your scientific research performance or research output?', the answer could always be 'Having a good number of publications.' Then, the next question forwarded is that 'What is the quality and impact of your published papers to the scientific community?' Of course, good and high impact factor journals always accept papers with high novelty in the subject area. Therefore, papers published in good journals are always highly cited and subsequently resulting in high impact (or citations) of the research done to the scientific community. However, the last question is sometimes very subjective and difficult to answer until *h*-index is introduced and discussed among the researchers. This paper aimed to discuss the *h*-index based on Elsevier's *Scopus* database as an indicator of research achievement for young Malaysian scientists.

Article history:
Received: 17 January 2011
Accepted: 4 October 2011

Email address:
yapckong@hotmail.com (Yap, C. K.)
*Corresponding Author

ISSN: 0128-7680

When Hirsch (2005) published his popular and highly cited paper entitled, 'An index to quantify an individual's scientific research output in Proceedings of the National Academy of Sciences of the United States of America' in 2005, the paper was cited 681 times in *Scopus* and 1211 times in Google Scholar, as searched on 8 Nov 2010. This shows the tremendous impacts and concerns of many researchers around the world who really care about what is the true or the best measurement of their research performance output. This is due to the fact that recognition of a researcher varies from university to university, both locally and internationally.

What is h-index?

The Hirsch's index (*h*-index) by Jorge Hirsch (Hirsch 2005) was introduced as an indicator of lifetime achievement (since the number will not decrease or change once the number is produced) as measured by the number of received citations. Hirsch (2007) proposed the *h*-index as a better alternative to other bibliometric indicators, such as the number of publications, average number of citations, and sum of all citations. The *h*-index is based on a scientist's lifetime citedness (Seglen 1992), which incorporates productivity as well as citation impacts. According to Hirsch (2005), "A scientist has index *h* if *h* of his or her number of papers published over *n* years (*Np*) have at least *h* citations each and the other (*Np* −*h*) papers have ≤*h* citations each." Therefore, all the research papers published by a scientist having at least *h* citations are called the 'Hirsch core' (Rousseau, 2006). Those papers in the 'Hirsch core' are the publications within a scientist's publication list that have the greatest visibility or greatest impact, according to Burrell (2007). Meanwhile, Egghe and Rousseay (2006) stated that the *h*-index is an original and simple new measure that incorporates both quantity and visibility of publications.

The introduction of the *h*-index by the physicist Hirsch (2005) as an indicator for quantifying the research output of scientists has since been discussed and studied theoretically and empirically in a number of disciplines (Bornmann & Daniel 2009). Han *et al*. (2010) compared between the journal impact factors and *h*-indices in the journals of reproduction biology computed from the ISI WoS, and found that the *h*-index (2001-2008) exhibited a positive correlation with a five years' Journal Impact Factor (2004-2008) (r= 0.64, p= 0.001). This clearly shows the relevance of the *h*-index as an indicator of the scientific performance output of researchers.

Thus, the *h*-index is advantageous since the necessary data for calculation are easy to access in the database without the need for any offline data processing (Batista *et al*., 2006). Perhaps, the most comprehensive review on the advantages and limitations of the *h*-index was reported by Costas and Bordons (2007), as presented in Table 1 below.

The *h*-index has already been used by major citation databases to evaluate the academic performance of individual scientists. Although effective and simple, the *h*-index suffers from some drawbacks that limit its use in accurately and fairly comparing the scientific output of different researchers. These drawbacks include information loss and low resolution; the former refers to the fact that in addition to *h*2 citations for papers in the *h*-core, excess citations are completely ignored, whereas the latter means that it is common for a group of researchers to have an identical *h*-index. Zhang (2009) suggested that *e*-index is a necessary *h*-index

Table 1: Advantages and limitations of the *h*-index as reviewed by Costas and Bordons (2007)

No.	Advantages	Limitations
1.	It combines a measure of quantity (publications) and impact (citations) in a single indicator	There are inter-field differences in typical *h* values due to the differences among the fields in productivity and citation practices, so the *h*-index should not be used to compare scientists from different disciplines.
2.	It allows us to characterize the scientific output of a researcher with objectivity, and therefore, may play an important role when making decisions about promotions, fund allocations and awarding prizes.	The *h*-index depends upon the duration of each scientist's career because the pool of publications and citations increases over time, in order to compare scientists at different stages of their career.
3.	It performs better than other single-number criteria that are commonly used to evaluate the scientific output of a researcher (impact factor, total number of documents, total number of citations, citation per paper rate and number of highly cited papers).	Highly cited papers are important for the determination of the *h*-index, but once they are selected to belong to the top *h* papers, the number of citations they receive is not important anymore.
4.	The *h*-index can be easily obtained by anyone with access to the citation databases, such as *Scopus;* in addition, it is also easy to understand.	Since the *h*-index is easily obtained, the authors run the risk of indiscriminate use, such as relying only on it for the assessment of scientists. Research performance is a complex multifaceted endeavour that cannot be assessed adequately by means of a single indicator.
5.	-	The use of the *h*-index could provoke changes in the publishing behaviour of scientists, such an artificial increase in the number of self-citations distributed among the documents on the edge of the *h*-index.
6.	-	There are also technical limitations, such as the difficulty in obtaining the complete output of scientists with very common names, or whether self-citations should be removed or not. Self-citations can increase a scientist's *h*, but their effect on *h* is much smaller than on the total citation count since only self-citations with a number of citations just >*h* are relevant.

complement, especially for evaluating highly cited scientists or for precisely comparing the scientific output of a group of scientists having an identical *h*-index. The *e*-index as a complementary to *h*-index has been recently supported by Dodson (2009).

Regardless of the limitations listed in Table 1, and with the many different modifications to h-index suggested in the literature, the h-index is still recommended to be used for our young Malaysian scientists. Although the *h*-index may sometimes not reflect the real publication citation for a researcher, the *h*-index is still a useful supplementary indicator, enrichment for the bibliometric toolset but not a substitution to the advanced indicator and long recognized standard such as Journal Impact Factor (Han *et al*., 2010).

Why h-index in Elsevier's Scopus database?

Elsevier's *Scopus* is the largest searchable abstract and citation database of research literature and selected web sources (Rew, 2010) and for this reason, the *Scopus* database was preferred in this paper. Moreover, *Scopus* databases are only based on cited papers while Google Scholar includes not only the cited papers but also the non-cited ones as well (Bar-Ilan *et al*. 2007). Therefore, *Scopus* is theoretically better in terms of quality of papers inclusion in its database although it could still be subjected to arguments and revisions. Bar-Ilan (2008) compared the *h*-indices of a list of highly-cited Israeli researchers based on citation counts retrieved from the Web of Science (WoS), *Scopus* and Google Scholar. In several cases, the results obtained through Google Scholar are considerably different from the results based on the WoS and *Scopus*. Meanwhile, Bar-Ilan *et al*. (2007) found that *Scopus* and the WoS are comparable in terms of the rankings induced. Some of the differences between the different databases are caused by the differing indexing strategies of the databases. Google Scholar does not have a clear policy, but unlike WoS, it indexes books and proceedings as well, and thus, resulting in citing more than journal papers.

Why h-index for young researchers in Malaysia?

The 'number of published papers' and the 'values of *h*-index' can show an evolution of how a researcher is recognized as being a true scientist. Perhaps, the saying 'published or perish' should now be upgraded to 'publish in high impact factor journals with potentially getting good h-index in future'. Here, the *h*-index is suggested to be used as an indicator of a researcher's achievement, particularly for young Malaysian scientists, due to four important points discussed below:

1. Most of the young research-based PhD degree holders in Malaysia (graduated in 2000 and after), both trained locally and overseas, are familiar and encouraged to disseminate their research findings to cited international journals as parts of the prerequisite to earn their PhD degrees. As a matter of fact, those researchers who are trained in Malaysian local universities are equally competitive with those trained overseas and are also able to write good scientific papers that are published in international cited journals (personal communication with Dr. Ahmad Zaharin Aris, the youngest PhD degree holder who graduated from Universiti Malaysia Sabah at the age <30).

2. Since *h*-index is highly dependent on the seniority of a researcher, a young researcher should have accumulated citations in order to get a high *h*-index. Of course, this is also continually being affected by several factors such as popularity of the research area, name of the journals and year or time when the paper was first published. Some papers were highly cited when it was newly published but less cited after a period of time. This could be due to the different subject area or field of study. For example, the *h*-index in biological

sciences tends to be higher than that in physics (Hirsch, 2005). Therefore, a range of the *h*-index value should be quantified for a specific field of study when a researcher is qualified to be promoted as an associate professor and a full professorship. Hirsch (2005) suggested that for faculty at major research universities, the *h*-index 12 might be a typical value for advancement to tenure (associate professorship) and that *h*-index 18 might be a typical value for advancement to full professorship.

3. There are only citations received since the year 1996 in *Scopus*. Prior to that, it was rather unfair for the old timer researchers to compare their number of citations or *h*-index with those researchers (usually those of the younger generation) who published their papers after 1996. To exemplify this, a young researcher from Japan named Assoc. Prof. Dr. Takaomi Arai (age < 40 years), who has published 83 papers in *Scopus* with an *h*-index of 16 as compared to a truly recognized academician in Malaysia, an old timer researcher (age > 60 years) named Prof. Dr. Tan Soon Guan (TSG) who has had a total of 119 papers in *Scopus* with an *h*-index of 17 (the above search was done on 26 November 2010 in *Scopus*). The first papers by Arai and TSG in *Scopus* were found to have been published in 1997 and 1976, respectively. Another good comparative example is between an excellent and highly profiled researcher in UPM, Prof. Dr. Yaakob Che Man (age > 50 years) and a young researcher (age < 40 years), Assoc. Prof. Dr. Tan Chin Ping (UPM), who had won the prestigious Prosper.Net-*Scopus* Young Scientist Award in Sustainable Development recently in Shanghai, China, on 5th July 2010 in the category of Agriculture and Food Security. Prof. Yaakob has a total number of 203 papers published in *Scopus* with an *h*-index of 19 while Dr. TCP has 105 publications with an *h*-index of 16. The first papers by Prof. Yaakob and TCP in *Scopus* were found published in 1991 and 1999, respectively (the search in *Scopus* for the last two researchers was done on 3rd December 2010). Therefore, I believe that Prof TSG's and Prof. Yaakob's *h*-indices should have simply reached at least more than 25 if the citations before 1996 were to be fully obtained in *Scopus*.
4. Since the number of publications sometimes does not reflect the impacts of the published works, the *h*-index should be employed to check the impact of the papers published apart from their quantity. Further details of the advantages of the *h*-index can be found in Table 1.

CONCLUSION

As a final note on this short note paper, I personally foresee the importance of the *h*-index for our Malaysian scientists as a good indicator to quantify an individual's scientific research output (Hirsch, 2005; Bornmann *et al*., 2010). Having reviewed all the above literature, in conclusion, I highly recommend the use of the *h*-index as a better and transparent indicator to complement the number of citations and the number of papers published. Regardless of the criticisms on the accuracy of *h*-index, including not taking into account the citation counts of papers with fewer than *h* citations as reviewed by Bornmann *et al*. (2010), this *h*-index-based bibliometric evaluation should still be ideally and routinely implemented as a must for the promotion committee in evaluating the academic performance of individual scientist, especially the young ones at all universities in Malaysia (be it governmental or non-governmental universities) if

Malaysia was to be recognized as one of the top research-based centre in this region. Certainly, a single indicator can never give more than a rough approximation to an individual's multifaceted profile, and therefore, many other factors should be taken into consideration in combination to evaluating an individual's performance output. This communication paper may potentially shed light to more literature-based studies on the proposed ranges of *h*-index for different academic achievements according to the field of study or subject area.

ACKNOWLEDGMENTS

I would like to thank Assoc. Prof. Dr. Takaomi Arai and Dr. Ahmad Zaharin Aris for stimulating and useful discussion on this topic.

REFERENCES

Bar-Ilan, J. (2008). Which *h*-index? – A comparison of WoS, Scopus and Google Scholar. *Scientometrics, 74*(2), 257–271.

Bar-Ilan, J., Levene, M., & Lin, A. (2007). Some measures for comparing citation databases. *Journal of Informetrics, 1*, 26–34.

Batista, P. D., Campiteli, M. G., Kinouchi, O., & Martinez, A. S. (2006). It is possible to compare researchers with different scientific interests? *Scientometrics, 68*, 179-189.

Bornmann, L., & Daniel, H. D. (2009). The state of *h* index research. Is the *h* index the ideal way to measure research performance? *EMBO Reports, 10*(1), 2–6.

Bornmann, L., Mutz, R., & Daniel, H. (2010). The *h* index research output measurement: Two approaches to enhance its accuracy. *Journal of Informetrics, 4*, 407–414.

Burrell, Q. L. (2007). On the *h*-index, the size of the Hirsch core and Jin's *A*-index. *Journal of Informetrics, 1*(2), 170–177.

Costas, R., & Bordons, M. (2007). The h-index: Advantages, limitations and its relation with other bibliometric indicators at the micro level. *Journal of Informetrics, 1*, 193–203.

Dodson, M. V. (2009). Citation analysis: Maintenance of *h*-index and use of *e*-index. *Biochemical and Biophysical Research Communications, 387*, 625–626.

Egghe, L., & Rousseay, R. (2006). An informatric model for the Hirsch-index. *Scientometrics 69,* 121-129.

Han, W., Yu, Q., & Wang, Y. (2010). Comparative analysis between impact factor and *h*-index for reproduction biology journals. *Journal of Animal and Veterinary Advances, 9*(11), 1552-1555.

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

Hirsch, J. E. (2007). Does the h index have predictive power? *Proceedings of the National Academy of Sciences of the United States of America, 104*(49), 19193–19198.

Rew, D. (2010). Editorial: SCOPUS: Another step towards seamless integration of the world's medical literature. *EJSO, 36*, 2-3.

Rousseau, R. (2006). New developments related to the Hirsch index. *Science Focus, 1*(4), 23–25.

Seglen, P. O. (1992). The skewness of science. *Journal of the American Society for Information Science, 43*(9), 628–638.

Yap, C. K. (2009). Acceptance and rejection of peer-reviewed articles in environmental sciences: My personal publication experience. *Pertanika Journal of Tropical Agricultural Science, 32*(1), 25-27.

Zhang, C. -T. (2009). The *e*-Index, complementing the *h*-Index for excess citations. *PLoS ONE 4*(5): e5429.