

## POINT/COUNTERPOINT

Suggestions for topics suitable for these Point/Counterpoint debates should be addressed to Colin G. Orton, Professor Emeritus, Wayne State University, Detroit: ortonc@comcast.net. Persons participating in Point/Counterpoint discussions are selected for their knowledge and communicative skill. Their positions for or against a proposition may or may not reflect their personal opinions or the positions of their employers.

### The $h$ index is the best measure of a scientist's research productivity

Clive Baldock, Ph.D.

Institute of Medical Physics, School of Physics, University of Sydney, Sydney, New South Wales 2006, Australia  
(Tel: +61 2 9351 8731, E-mail: c.baldock@physics.usyd.edu.au)

Ruimin Ma, M.S.

Research Center for Chinese Science Evaluation, Wuhan University, Wuhan 430072, People's Republic of China  
(Tel: 86 13871473654, E-mail: ruimin.ma@yahoo.com.cn)

Colin G. Orton, Ph.D., Moderator

(Received 4 February 2009; accepted for publication 4 February 2009; published 5 March 2009)

[DOI: [10.1118/1.3089421](https://doi.org/10.1118/1.3089421)]

#### OVERVIEW

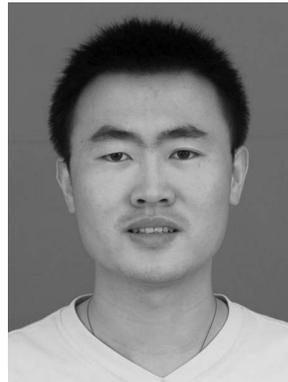
It is sometimes necessary to evaluate the research output of a scientist when, for example, making decisions on recruitment or advancement of faculty, or considering a candidate for an award. The quantity of a scientist's publications is easy to evaluate, but not the quality. Hirsch<sup>1</sup> proposed a single number, the  $h$  index, for the characterization of both the quantity and significance of a scientist's research publications. Some, however, have questioned the "accuracy" of this index for a variety of reasons. The premise that the  $h$  index is the best measure of the research productivity of a scientist is the topic of this month's Point/Counterpoint.



Arguing for the Proposition is Clive Baldock, Ph.D. Dr. Baldock obtained his M.S. in Radiation Physics from St. Bartholomew's Medical College, University of London and his Ph.D. from Kings College, University of London. His Ph.D. research was on the development of MRI of radiation sensitive gels for three-dimensional radiotherapy dosimetry. In 1997 he moved to

Australia where he initially worked at the Centre for Medical, Health and Environmental Physics, School of Physical Sciences, Queensland University of Technology in Brisbane. In 2003, he was appointed to his current position in the School of Physics at the University of Sydney where he is Professor of Medical Physics. At the University of Sydney he is also the Director of the Institute of Medical Physics and Associate Dean and Director of Postgraduate Coursework in

the Faculty of Science. His research interests include most aspects of radiation physics, radiotherapy physics and dosimetry, and medical imaging.



Arguing against the Proposition is Ruimin Ma, M.S. Mr. Ma obtained his B.S. from Shanxi University in 2000 and his M.S. from Wuhan University, Wuhan, People's Republic of China, in 2004. He is currently completing research for his Ph.D. in Informatics at Wuhan University, where he is an assistant researcher in the Research Centre for Chinese Scientific Evaluation. His major

research interest is scientometrics, especially domain analysis, visualization, and evaluation of research competitiveness. He has participated as one of the research leaders in several evaluation projects for Chinese and other universities, and Chinese academic journals, and has prepared consultant reports commissioned by the Ministry of Education of the People's Republic of China and several universities. He has published over 20 papers on research evaluation in both international and domestic journals.

#### FOR THE PROPOSITION: Clive Baldock, Ph.D.

##### Opening Statement

The  $h$  index has been devised by Hirsch<sup>1</sup> as a research performance indicator intended to be an improved measure of the impact and quality of the work of an individual researcher. The  $h$  index is that value where  $N_p$  of their papers

has at least  $h$  citations each and the other  $(N_p - h)$  papers have  $\leq h$  citations each. Researchers with an  $h$  index of 30 have, when their papers are ordered by the number of citations received from highest to lowest, their 30th paper having been cited at least 30 times and papers 1–30 having greater than or equal to 30 citations. The  $h$  index reflects both the number of publications and the number of citations per publication. It is designed to improve upon simpler measures such as the total number of citations or publications. The  $h$  index works properly only for comparing scientists from the same discipline because citation conventions differ for different disciplines. It has rapidly become an alternative to the more traditional metric of journal impact factor in the evaluation of the impact of the work of a particular researcher.<sup>2</sup> Numerous papers published in the literature explore the role of the  $h$  index further,<sup>3–5</sup> including in the discipline of medical physics.<sup>6</sup>

As only the most highly cited articles contribute to the  $h$  index, its determination is a relatively simple process. The  $h$  index can be determined using databases such as Web of Science®, Scopus®, or Google Scholar®. Different databases used to calculate the  $h$  index for the same researcher, however, often produce a slightly different result, with Google Scholar® having more citations than Scopus® and Web of Science®, with the latter citation collections tending to be more accurate than the former.<sup>6</sup>

Hirsch calculated the highest  $h$  value among physicists to be that of Witten from the Princeton Institute for Advanced Study, for whom it is 110, with Hawking having an  $h$  index of 62.<sup>1</sup> Hirsch subsequently demonstrated that the  $h$  index is highly predictive as to whether a scientist will be elected to a fellowship of a national academy or even awarded a Nobel Prize.<sup>7</sup> For physicists, Hirsch suggested that a value for  $h$  of about 10–12 might be a useful guideline for making a decision regarding tenure at major research universities. A value of 18 might be a useful guideline for a full professorship, 15–20 for fellowship in the American Physical Society, and 45 or higher for membership in the National Academy of Sciences. As a numerical value it can be used to quantify both the *productivity* and *impact* of a scientific researcher and is becoming the metric of choice in academic circles for assessment when considering grant allocation, and making offers of employment, tenure, promotion, and fellowship in learned societies.<sup>8</sup>

## AGAINST THE PROPOSITION: Ruimin Ma, M.S.

### Opening Statement

The  $h$  index, which measures the quality and quantity of an author's research papers,<sup>1</sup> has become an accepted indicator for evaluation of the research productivity of scientists. However, because of several drawbacks, it is not the best measurement, as I will demonstrate.

First, ingenious though the  $h$  index is, it neglects the dynamicity of citations. The  $h$  index is significantly influenced by highly cited papers which, once selected, have no influence whatsoever on the computation of the author's  $h$  index in subsequent years, no matter how their citations increase.<sup>9</sup>

The  $g$ ,  $R$ , and  $AR$  indices<sup>9,10</sup> have been developed to correct this deficiency. In addition, the  $h$  index has been shown to be easily influenced by subterfuges such as excessive self-citation.

Second, the relevance of the  $h$  index varies considerably between specialties. It is more logical to apply it to the measurement of research productivity in the natural sciences than humanities and social sciences, for which researchers publish significantly more in monographs rather than journals: These are not included in  $h$ -index computations. However, even *within* the natural sciences, citing patterns vary considerably between subjects. For example, the average  $h$  index for the top ten scientists is 147.1 for the life sciences but only 63.7 in computer science.<sup>11</sup> Several methods have had to be developed to address this problem of noncomparability.<sup>12–14</sup>

Third, the  $h$  index is highly dependent on the length of time scientists have devoted to research work. The  $h$  index is biased toward older scientists because younger scientists have had less time to generate sufficient numbers of papers and subsequent citations. It is possible for a scholar who has hardly published any papers for several years to have a high  $h$  index. Therefore, the  $h$  index puts newcomers at a disadvantage, potentially hampering the recognition of young but excellent scientists.<sup>13</sup>

Fourth, the  $h$  index overvalues the quantity of papers published and undervalues the quality. A researcher with a large number of poorly cited papers will have a higher  $h$  index than one with half as many very highly cited papers,<sup>15</sup> which means it is impossible for a scientist with a limited number of excellent papers to obtain a high  $h$  index.<sup>16</sup>

Lastly, the  $h$  index alone cannot substitute for indicators from traditional bibliometrics. After so many years of application, the combined use of bibliometric indicators, such as number of papers, citations, average citations, and impact factor, is recommended rather than a single  $h$  index.<sup>15</sup> The  $h$  index alone relies too heavily on the number of papers and total citations.<sup>15</sup> It is significant that the standard adopted by the Institute for Scientific Information (ISI) and the Essential Science Indicators (ESI) databases for selection of the most highly cited and “hot” papers considers both the time period of publication and the subject field of a paper and reduces the impact of the total number of citations, in contrast to the  $h$  index.

In summary, it is a complicated process to evaluate a scientist's research productivity. The  $h$  index should be considered as a supplement to traditional bibliometric indicators but definitely not an omnipotent indicator. The  $h$  index is neither perfect nor the best.

### Rebuttal: Clive Baldock, Ph.D.

The introduction by Hirsch of the  $h$  index generated much interest in the academic community and discussion in the literature as to whether the  $h$  index is the “best” metric to measure how “good” a scientist is.<sup>17</sup> In this context, and in order to investigate whether there are better comparative indices than the  $h$  index, a study was undertaken of nine different Hirsch-type variants on the  $h$  index, which had been

suggested in the literature.<sup>18</sup> These variants were the  $m$  quotient,  $g$  index,  $h(2)$  index,  $a$  index,  $r$  index,  $ar$  index, and  $h_w$  index. The authors found from factor analysis that there are two types of indices: Some describe the most productive “core” of a scientist’s output and give the number of papers in that core, and the others describe the impact of the papers in the core. The use of pairs of indices was suggested as a meaningful indicator for comparing scientists, where one index relates to the number of papers in the researcher’s productive core (either the  $h$  or  $g$  index) and the other relates to the impact of the papers in the researcher’s productive core (either the  $a$  or  $m$  index). It is noted however that a number of the proposed Hirsch-type indices are actually derived variants of the  $h$  index, thereby perhaps reinforcing the premise that the fundamental metric, the  $h$  index, or a variant of, is the best “measure” of a scientist’s research productivity.

It will perhaps be only after more studies have been published in this area, particularly by those working in the information sciences and researching into evaluative bibliometrics, that the true value of the  $h$  index will become evident. However, since the  $h$  index is automatically calculated in the “citation report” function of Web of Science®, it is likely to be used by researchers in the foreseeable future as it does indicate the broad impact of a scientist’s cumulative research contributions.

#### Rebuttal: Ruimin Ma, M.S.

Indeed, the  $h$  index is quite remarkable for its ingenuity and ease of use, and it does measure the quantity and quality of the papers of an author. However, these advantages do not guarantee that it is the best measurement.

Dr. Baldock, in citing Ball’s paper,<sup>2</sup> stated that the  $h$  index has become an alternative to the impact factor, which is ambiguous. The impact factor evaluates the quality of journals and is not capable of measuring the impact of the work of any specific researcher directly. It is possible that the  $h$  index might eventually gain stature equal to that of the impact factor, but they have completely different functions.

Also, Dr. Baldock, by citing Hirsh,<sup>1</sup> demonstrated that the  $h$  index may be a useful guideline to classify researchers and make decisions on promotion, fellowship, and so on. However I have to indicate that some drawbacks exist in the  $h$  index’s algorithm. For example, author A has published six papers which, respectively, have 20, 15, 9, 7, 6, and 4 citations. In contrast, author B has also published six papers but with citations of 100, 25, 8, 7, 6, and 4, respectively. Obviously, both authors have the same  $h$  value (viz., 5). The total citations of author B are, however, much higher than those of

author A, which means that the  $h$  index has ignored some significant aspects of the outputs of these researchers. It is not sufficient to apply the  $h$  index, therefore, without combining it with traditional metric indicators.

In addition, Dr. Baldock considered that the  $h$  index should only be used to compare the outputs of researchers within specific disciplines, which is a major drawback. Some suggestions have been made to address this deficiency.<sup>12–14</sup>

In summary, whether the  $h$  index should be widely used to measure the outputs of researchers has still to be adequately tested. It is certainly an interesting and attractive index, but definitely not the best indicator of the impact of a researcher’s work.

- <sup>1</sup>J. E. Hirsch, “An index to quantify an individual’s scientific research output,” *Proc. Natl. Acad. Sci. U.S.A.* **102**, 16569–16572 (2005).
- <sup>2</sup>P. Ball, “Index aims for fair ranking of scientists,” *Nature (London)* **436**, 900 (2005).
- <sup>3</sup>S. Lehmann, A. D. Jackson, and B. E. Lautrup, “Measures for measures,” *Nature (London)* **444**, 1003–1004 (2006).
- <sup>4</sup>J. K. Vanclay, “On the robustness of the  $h$ -index,” *J. Am. Soc. Inf. Sci. Technol.* **58**, 1547–1550 (2007).
- <sup>5</sup>P. Vinkler, “Eminence of scientists in the light of the  $h$ -index and other scientometric indicators,” *J. Inf. Sci.* **33**, 481–491 (2007).
- <sup>6</sup>C. Baldock, “The  $h$ -index and Medical Physics,” *Australas. Phys. Eng. Sci. Med.* **31**, 10–12 (2008).
- <sup>7</sup>J. E. Hirsch, “Does the  $h$  index have predictive power?,” *Proc. Natl. Acad. Sci. U.S.A.* **104**, 19193–19198 (2007).
- <sup>8</sup>C. D. Kelly and M. D. Jennions, “The  $h$  index and career assessment by numbers,” *Trends Ecol. Evol.* **21**, 167–170 (2006).
- <sup>9</sup>L. Egghe, “Theory and practise of the  $g$ -index,” *Scientometrics* **69**, 131–152 (2006).
- <sup>10</sup>B. H. Jin, L. M. Liang, R. Rousseau, and L. Egghe, “The  $R$ - and  $AR$ -indices: Complementing the  $h$ -index,” *Chin. Sci. Bull.* **52**, 855–863 (2007).
- <sup>11</sup>J. M. Zhao, J. P. Qiu, K. Huang, and B. H. Liu, “A new scientometric indicator—Review on  $h$ -index and its applications,” *Bulletin of the National Natural Science Foundation* **1**, 23–32 (2008).
- <sup>12</sup>J. E. Iglesias and C. Pecharrmán, “Scaling the  $h$ -index for different scientific ISI fields,” *Scientometrics* **73**, 303–320 (2007).
- <sup>13</sup>J. P. Qiu, R. M. Ma, and N. Cheng, “New exploratory work of evaluating a researcher’s output,” *Scientometrics* **77**, 335–344 (2008).
- <sup>14</sup>P. D. Batista, M. G. Campiteli, O. Kinouchi, and A. S. Martinez, “Is it possible to compare researchers with different scientific interests,” *Scientometrics* **68**, 179–189 (2006).
- <sup>15</sup>R. Costas and M. Bordons, “The  $h$ -index: Advantages, limitations and its relation with other bibliometric indicators at the micro level,” *Journal of Informetrics* **1**, 193–203 (2007).
- <sup>16</sup>W. Glänzel, “On the opportunities and limitations of the  $h$ -index,” *Science Focus* **1**, 10–11 (2006).
- <sup>17</sup>L. Bornmann and H. Daniel, “What do we know about the  $h$  index?,” *J. Am. Soc. Inf. Sci. Technol.* **58**, 1381–1385 (2007).
- <sup>18</sup>L. Bornmann, R. Mutz, and H. Daniel, “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,” *J. Am. Soc. Inf. Sci. Technol.* **59**, 830–837 (2008).