viewpoint

The state of *h* index research

Is the *h* index the ideal way to measure research performance?

Lutz Bornmann & Hans-Dieter Daniel

How does one measure the quality of science? The question is not rhetorical; it is extremely relevant to promotion committees, funding agencies, national academies and politicians, all of whom need a means by which to recognize and reward good research and good researchers. Identifying high-quality science is necessary for science to progress, but measuring quality becomes even more important in a time when individual scientists and entire research fields increasingly compete for limited amounts of money. The most obvious measure available is the bibliographic record of a scientist or research institute—that is, the number and impact of their publications.

Identifying high-quality science is necessary for science to progress…

Currently, the tool most widely used to determine the quality of scientific publications is the journal impact factor (IF), which is calculated by the scientific division of Thomson Reuters (New York, NY, USA) and is published annually in the Journal Citation Reports (JCR). The IF itself was developed in the 1960s by Eugene Garfield and Irving H. Sher, who were concerned that simply counting the number of articles a journal published in any given year would miss out small but influential journals in their Science Citation Index (Garfield, 2006). The IF is the average number of times articles from the journal published in the past two years have been cited in the JCR year and is calculated by dividing the number of citations in the JCR year—for example, 2007—by the total number of articles published in the two previous years—2005 and 2006.

Owing to the availability and utility of the IF, promotion committees, funding agencies and scientists have taken to using it as a shorthand assessment of the quality of scientists or institutions, rather than only journals. As Garfield has noted, this use of the IF is often necessary, owing to time constraints, but not ideal (Garfield, 2006). *Nature* has, for example, shown "how a high journal impact factor can be the skewed result of many citations of a few papers rather than the average level of the majority, reducing its value as an objective measure of an individual paper" (Campbell, 2008).

Needless to say, the widespread use of the IF and the way in which it is calculated have attracted much criticism and even ridicule (Petsko, 2008). Many commentators have argued that it wrongly equates the importance of a paper with the IF of the journal in which it was published (Notkins, 2008), and that some scientists are now more concerned about publishing in high-IF journals than they are about their research, which negatively affects the peer-review and scientific publication process. "Scientists … submit their papers to journals at the top of the impact factor ladder, circulating progressively through journals further down the rungs when they are rejected" (Simons, 2008); this is a waste of time for both editors and reviewers. Moreover, Thomson Reuters itself has been criticized for a lack of transparency and for the fact that journals can negotiate what to include in the denominator and are thus able to influence their IF (Anon, 2006).

In August 2005, Jorge Hirsch—a physicist at the University of California, San Diego, USA—introduced a new indicator for quantifying the research output of scientists (Bornmann & Daniel, 2007a; Hirsch, 2005).

Hirsch's so-called *h* index was proposed as an alternative to other bibliometric indicators—such as the number of publications, the average number of citations and the sum of all citations (Hirsch, 2007)—and is defined as follows: "A scientist has index *h* if *h* of his or her N_p papers have at least *h* citations each and the other $(N_p - h)$ papers have ≤ *h* citations each" (Hirsch, 2005). All papers by a scientist that have at least *h* citations are called the "Hirsch core" (Rousseau, 2006). An *h* index of 5 means that a scientist has published five papers that each have at least five citations. An *h* index of 0 does not inevitably indicate that a scientist has been completely inactive: he or she might have already published a number of papers, but if none of the papers was cited at least once, the *h* index is 0.

…the widespread use of the IF and the way in which it is calculated have attracted much criticism and even ridicule…

Shortly after Hirsch submitted his 2005 paper on the *h* index to the electronic archive arXiv.org as a preprint, both *Nature* (Ball, 2005) and *Science* (Anon, 2005) reported on it. At the initiative of Manuel Cardona, Emeritus Professor at the Max Planck Institute for Solid State Research in Stuttgart, Germany, and a member of the US National Academy of Sciences (Washington, DC, USA), the preprint was published a few weeks later in a revised form in the *Proceedings of the National Academy of Sciences*. Until late 2008, the paper has been cited about 200 times, which shows that Hirsch's proposal to represent the research achievements of a scientist as a single number is fascinating
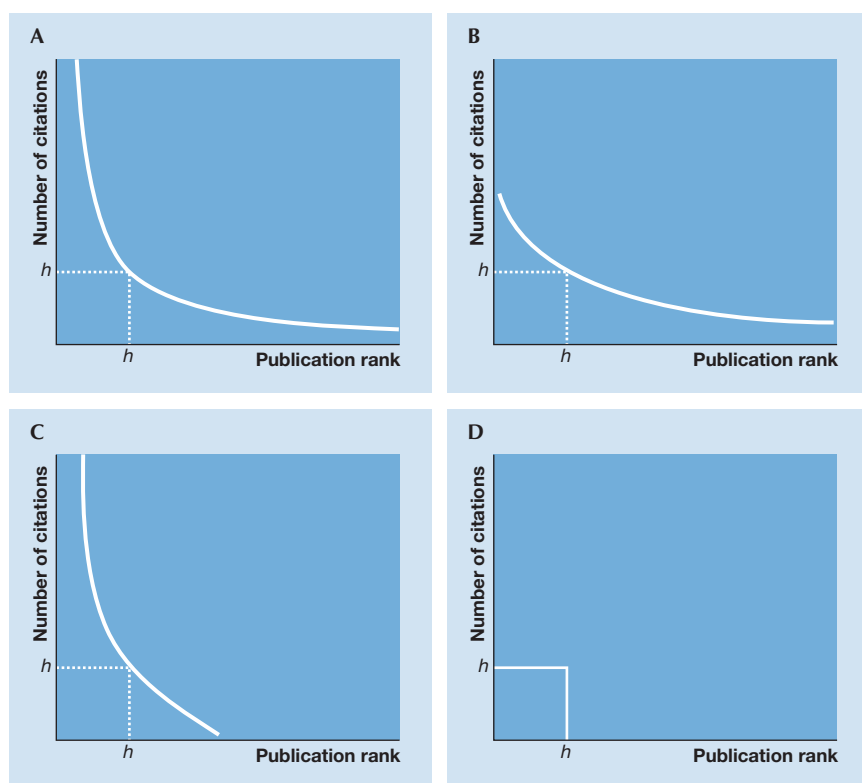
Fig 1 | The *h* index for distributions of citation frequencies (the publications are sorted in graphs **A–D** by number of citations).

Table 1 | Example of *h* indices for different publication records

Paper	Citations	
	Scientist A	**Scientist B**
1	51	6
2	34	5
3	29	4
4	22	4
5	3	3
6	1	3
7	0	2
8	–	2
9	–	1
10	–	0
11	–	0
h	4	4

Two scientists with the same *h* index: scientist A with few, highly cited papers and scientist B with many rarely cited papers.

to many people, not only to science news editors. However, the combination of publications and citation frequencies into one value has been criticized by some scientists as making little sense: "The problem is that Hirsch assumes an equality between incommensurable quantities. An author's papers are listed in order of decreasing citations with paper *i* having *C*(*i*) citations. Hirsch's index is determined by the equality, $h = C(h)$, which posits an equality between two quantities with no evident logical connection" (Lehmann *et al*, 2008).

The *h* index can now be calculated automatically for any publication set in the Web of Science (WoS; provided by Thomson Reuters) and is already regarded as the counterpart to the IF (Gracza & Somoskovi, 2007). WoS is not the only literature database that allows a user to calculate the *h* index; any database that includes the references cited in the publication will do, such as Chemical Abstracts provided by Chemical Abstracts Services (Columbus, OH, USA), Google Scholar, or Scopus provided by Elsevier (Amsterdam, The Netherlands; Jacso, 2008). Depending

on what publications a database covers and analyses, however, the calculation of the *h* index will produce different results. Using Google Scholar, for example, can lead to different results from using databases that require subscription fees, such as Chemical Abstracts, Scopus and WoS (Bornmann *et al*, 2008a).

In any case, the main debate surrounding the *h* index concerns how meaningful it is as a measure of a scientist's performance. Fig 1 shows a range of distributions of citation frequencies for a given publication set and the corresponding *h* index. As a rule, the distribution of citation frequencies for a larger number of publications is skewed to the right according to a power law (Joint Committee on Quantitative Assessment of Research, 2008). A scientist's publication record usually contains a few highly cited papers and many rarely cited papers. As Fig 1A shows, the *h* index captures only a part of the publication and citation data if the distribution is right-skewed, as it fails to represent highly and rarely cited or non-cited papers. Scientists with very different

citation frequencies can therefore have the same *h* index (Table 1; Fig 1B,C): "Think of two scientists, each with 10 papers with 10 citations, but one with an additional 90 papers with 9 citations each; or suppose one has exactly 10 papers of 10 citations and the other exactly 10 papers of 100 each. Would anyone think them equivalent"? (Joint Committee on Quantitative Assessment of Research, 2008).

An *h* index that completely captures the distribution of citation frequencies is shown in Fig 1D: a scientist has published *h* papers, of which each has received *h* citations; however, such 'constant performers' are very rare. The *h* index therefore provides an incomplete picture for most scientists, whose publication and citation data have a right-skewed distribution. Evidence Ltd, a company in Leeds, UK, that analyses research performance, considers the *h* index to be an indicator with low information content that is not applicable to the general body of researchers (Evidence Ltd, 2007). Nonetheless, *h*-index values are already being used in various disciplines—for example, in chemistry, information sciences, medicine, physics and economics—to produce ranking lists of scientists.

The *h* index is no longer being used only as a measure of scientific achievement for

Sidebar A | Definition of the *h* index and its variants.

Index	Definition
h index	"A scientist has index *h* if *h* of his or her N_p papers have at least *h* citations each and the other $(N_p - h)$ papers have fewer than ≤ *h* citations each" (Hirsch, 2005, p 16569)
m quotient	$\dfrac{h}{y}$ where *h* is *h* index, *y* is number of years since publishing the first paper
g index	"The *g*-index *g* is the largest rank (where papers are arranged in decreasing order of the number of citations they received) such that the first *g* papers have (together) at least g^2 citations" (Egghe, 2006, p 144)
a index	$\dfrac{1}{h}\sum_{j=1}^{h}\mathrm{cit}_j$ where *h* is *h* index, cit is citation counts
ar index	$\sqrt{\sum_{j=1}^{h}\dfrac{\mathrm{cit}_j}{a_j}}$ where *h* is *h* index, cit is citation counts, *a* is number of years since publishing

individual researchers, but also to measure the scientific output of research groups (van Raan, 2006), scientific facilities and even countries. The ranking list for the research of individual nations in biology and biochemistry for the period 1996–2006, for example, shows that the USA, the UK and Germany have *h* indices of 400, 219 and 206 respectively (Csajbók *et al*, 2007). The indices at such aggregated levels—group, research facility or country—are calculated analogously to those of individual researchers. In addition, it is also possible to calculate successive *h* indices at higher aggregate levels (Prathap, 2006): "The institute has an index h_2 if h_2 of its *N* researchers have an h_1-index of at least h_2 each, and the other $(N - h_2)$ researchers have h_1-indices lower than h_2 each. The succession can then be continued, for example, for networks of institutions or countries or other higher levels of aggregation" (Schubert, 2007).

> **The *h* index [...] is already regarded as the counterpart to the IF...**

Braun *et al* (2005) also recommend using the *h* index as an alternative to the IF to qualify journals: "Retrieving all source items of a given journal from a given year and sorting them by the number of times cited, it is easy to find the highest rank number which is still lower than the corresponding times cited value. This is exactly the *h*-index of the journal for the given year." As the *h* index for a journal cannot be higher than the number of papers that are published in a certain period, journals that publish only a few highly cited papers should therefore not be included in a ranking list that is based on the *h* index—this concerns mainly journals that predominantly publish reviews. For example, the *Annual Review of Biochemistry* published only 28 papers in 2005, which were cited on average about 100 times up until mid-2008. Other authors have proposed an *hb* index to assess research in selected areas (Banks, 2006; Egghe & Ravichandra, 2008).

Any new bibliometric indicator to measure scientific performance should be carefully checked for its validity and its ability to represent scientific quality correctly. In a series of investigations (Bornmann & Daniel, 2007b; Bornmann *et al*, 2008b), we have shown that, for individual scientists, the *h* index correlates well with both the number of their publications and the number of citations these publications have attracted—which is hardly surprising given that the *h* index was proposed to do exactly that. More importantly, however, are studies that have examined the relationship between the *h* index and peer judgements of research performance (Moed, 2005), of which there are only four available so far. Bornmann & Daniel (2005) and Bornmann *et al* (2008b) have shown that the average *h*-index values of accepted and rejected applicants for biomedicine research fellowships differ statistically significantly, while van Raan (2006) found that the *h* index "relates in a quite comparable way with peer judgments" for 147 Dutch research groups in chemistry. Finally, Lovegrove & Johnson (2008) have reported similar findings for the relationship between the *h* index and peer judgements in the context of grant applicants to the National Research Foundation of South Africa (Pretoria, South Africa). Although these studies provide confirmation of the *h* index's validity, it will require more time and research before it can be used in practice to assess scientific work.

Several disadvantages to the *h* index have also been pointed out (Bornmann & Daniel, 2007a; Jin *et al*, 2007). For example, "like most pure citation measures it is field-dependent, and may be influenced by self-citations…; [t]he number of co-authors may influence the number of citations received…; [t]he *h*-index, in its original setting, puts newcomers at a disadvantage since both publication output and observed citation rates will be relatively low…; [t]he *h*-index lacks sensitivity to performance changes: it can never decrease and is only weakly sensitive to the number of citations received" (Rousseau, 2008).

> **To measure the quality of scientific output, it would therefore be sufficient to use just two indices: one that measures productivity and one that measures impact...**

This has led to the development of numerous variants of the *h* index (Sidebar A). The *m* quotient, for example, is computed by dividing the *h* index by the number of years that the scientist has been active since the first published paper (Hirsch, 2005). Unlike the *h* index, the *m* quotient avoids a bias towards more senior scientists with longer careers and more publications. The h_I index is "a complementary index $h_I = h^2/N_a^{(T)}$, with $N_a^{(T)}$ being the total number of authors in the considered *h* papers" (Batista *et al*, 2006). It is meant to reduce the bias towards scientists that publish frequently as co-authors. The h_c index excludes self-citations to avoid a bias towards scientists who disproportionately often cite their own work (Schreiber, 2007).

The *a* index indicates the average number of citations of publications in the Hirsch core (Jin, 2006), whereas the *g* index is defined as follows: "[a] set of papers has a *g*-index *g* if *g* is the highest rank such that the top *g* papers have, together, at least g^2 citations" (Egghe, 2006). In contrast to the *h* index, which corresponds to the number of citations for the publication with the fewest

citations in the Hirsch core, the *a* index and the *g* index are meant to give more weight to highly cited papers. The *ar* index is defined as the square root of the sum of the average number of citations per year of articles included in the *h*-core (Jin, 2007) and is meant to avoid favouring scientists who have stopped publishing because the *h* index can never decrease over time; even if a scientist is no longer active, his or her *h* index remains constant in the worst case. Of the various indices that have been proposed in recent years, the *g* index by Egghe (2006) has received most attention, whereas many other derivatives of the *h* index have had little response.

All empirical studies that have tested the various indices for scientists or journals have reported high correlation coefficients. Apparently, this indicates a redundancy among the various indices to measure achievement. The results of two studies by Bornmann *et al* (2008c,d) state more precisely that the *h* index and its variants are, in effect, two types of index. "The one type of indices […] describe the most productive core of the output of a scientist and tell us the number of papers in the core. The other indices […] depict the impact of the papers in the core" (Bornmann *et al*, 2008c). To measure the quality of scientific output, it would therefore be sufficient to use just two indices: one that measures productivity and one that measures impact—for example, the *h* index and the *a* index.

As described above, only four studies have examined the validity of the *h* index by testing the relationship between a scientist's *h*-index value and peer assessments of his or her achievements. Although the results of the studies mentioned are positive, we need further studies that use extensive data sets to examine the *h* index and possibly to select variants for use in various fields of application. Future research on the *h* index should no longer be aimed at developing new variants, but should instead test the validity of the existing ones. Only once such studies have confirmed the fundamental validity of the *h* index and certain variants should it be used to assess scientific work.

As a basic principle, it is always prudent to use several indicators to measure research performance (Glänzel, 2006a; van Raan, 2006). The publication set of a scientist, journal, research group or scientific facility should always be described using many indicators such as the number of publications with zero citations, the number of highly cited papers and the number of papers for which the scientist is first or last author. As publication and citation conventions differ considerably across disciplines, it is also important to use additional bibliometric indicators that measure the "relative, internationally field-normalized impact" of publications (van Raan, 2005)—for example, the indicators developed by the Centre for Science and Technology Studies (CWTS; Leiden, The Netherlands) or the Institute for Research Policy Studies of the Hungarian Academy of Sciences (Budapest, Hungary; Glänzel, 2006b, 2008). In addition to bibliometric indicators, every evaluation study should also provide a measure of concentration, such as the Gini coefficient or the Herfindahl index, to assess the distribution of the citations among a scientist's publications (Bornmann *et al*, 2008e; Evans, 2008).

> As a basic principle, it is always prudent to use several indicators to measure research performance...

If the *h* index is used for the evaluation of research performance, it should always be taken into account that, similar to other bibliometric measures, it is dependent on the length of an academic career and the field of study in which the papers are published and cited. For this reason, the index should only be used to compare researchers of a similar age and within the same field of study. At the end of the day, all measurements of research quality should be taken with a grain of salt; it is certainly not possible to describe a scientist's contributions to a given research field with mere numerical values. As Albert Einstein (1879–1955) famously noted: "[n]ot everything that counts is countable, and not everything that's countable counts."

REFERENCES
Anon (2005) Data point. *Science* **309:** 1181
Anon (2006) The impact factor game: it is time to find a better way to assess the scientific literature. *PLoS Med* **3:** e291
Ball P (2005) Index aims for fair ranking of scientists. *Nature* **436:** 900
Banks MG (2006) An extension of the Hirsch index: indexing scientific topics and compounds. *Scientometrics* **69:** 161–168
Batista PD, Campiteli MG, Kinouchi O (2006) Is it possible to compare researchers with different scientific interests? *Scientometrics* **68:** 179–189

Bornmann L, Daniel H-D (2005) Does the *h*-index for ranking of scientists really work? *Scientometrics* **65:** 391–392
Bornmann L, Daniel H-D (2007a) What do we know about the *h* index? *J Am Soc Inf Sci Tec* **58:** 1381–1385
Bornmann L, Daniel H-D (2007b) Convergent validation of peer review decisions using the *h* index. Extent of and reasons for type I and type II errors. *J Informetrics* **1:** 204–213
Bornmann L, Marx W, Schier H, Rahm E, Thor A, Daniel H-D (2008a) Convergent validity of bibliometric Google Scholar data in the field of chemistry—citation counts for papers that were accepted by *Angewandte Chemie International Edition* or rejected but published elsewhere, using Google Scholar, Science Citation Index, Scopus, and Chemical Abstracts. *Journal of Informetrics* [doi:10.1016/j.joi.2008.11.001
Bornmann L, Wallon G, Ledin A (2008b) 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* **17:** 149–156
Bornmann L, Mutz R, Daniel H-D (2008c) 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 Tec* **59:** 830–837
Bornmann L, Mutz R, Daniel H-D, Wallon G, Ledin A (2008d) Are there really two types of *h* index variants? A validation study by using molecular life sciences data. In *Excellence and Emergence. A New Challenge for the Combination of Quantitative and Qualitative Approaches. 10th International Conference on Science and Technology Indicators*, J Gorraiz, E Schiebel (eds), pp 256–258. Vienna, Austria: Austrian Research Centers
Bornmann L, Mutz R, Neuhaus C, Daniel H-D (2008e) Use of citation counts for research evaluation: standards of good practice for analyzing bibliometric data and presenting and interpreting results. *Ethics in Science and Environmental Politics* **8:** 93–102
Braun T, Glänzel W, Schubert A (2005) A Hirsch-type index for journals. *The Scientist* **19:** 8
Campbell P (2008) Escape from the impact factor. *Ethics in Science and Environmental Politics* **8:** 5–7
Csajbók E, Berhidi A, Vasas L, Schubert A (2007) Hirsch-index for countries based on Essential Science Indicators data. *Scientometrics* **73:** 91–117
Egghe L (2006) Theory and practise of the *g*-index. *Scientometrics* **69:** 131–152
Egghe L, Ravichandra Rao IK (2008) The influence of the broadness of a query of a topic on its *h*-index: models and examples of the *h*-index of *n*-grams. *J Am Soc Inf Sci Tec* **59:** 1688–1693
Evans JA (2008) Electronic publication and the narrowing of science and scholarship. *Science* **321:** 395–399
Evidence Ltd (2007) *The Use of Bibliometrics to Measure Research Quality in UK Higher Education Institutions*. London, UK: Universities UK
Garfield E (2006) The history and meaning of the Journal Impact Factor. *JAMA* **295:** 90–93
Glänzel W (2006a) On the *h*-index—a mathematical approach to a new measure

of publication activity and citation impact. *Scientometrics* **67:** 315–321

Glänzel W (2006b) On the opportunities and limitations of the *H*-index. *Science Focus* **1:** 10–11

Glänzel W, Thijs B, Schubert A, Debackere K (2008) Subfield-specific normalized relative indicators and a new generation of relational charts. Methodological foundations illustrated on the assessment of institutional research performance. *Scientometrics* [doi:10.1007/s11192–008–2109–5]

Gracza T, Somoskovi I (2007) Impact factor and/ or Hirsch index? *Orvosi Hetilap* **148:** 849–852

Hirsch JE (2005) An index to quantify an individual's scientific research output. *Proc Natl Acad Sci USA* **102:** 16569–16572

Hirsch JE (2007) Does the *h* index have predictive power? *Proc Natl Acad Sci USA* **104:** 19193–19198

Jacso P (2008) Testing the calculation of a realistic *h*-index in Google Scholar, Scopus, and Web of Science for F.W. Lancaster. *Library Trends* **56:** 784–815

Jin B (2006) *h*-index: an evaluation indicator proposed by scientist. *Science Focus* **1:** 8–9

Jin B (2007) The AR-index: complementing the *h*-index. *ISSI Newsletter* **3:** 6

Jin B, Liang L, Rousseau R, Egghe L (2007) The R- and AR-indices: complementing the *h*-index. *Chin Sci Bull* **52:** 855–863

Joint Committee on Quantitative Assessment of Research (2008) *Citation Statistics. A Report from the International Mathematical Union (IMU) in Cooperation with the International Council of Industrial and Applied Mathematics (ICIAM) and the Institute of Mathematical Statistics (IMS)*. Berlin, Germany: International Mathematical Union

Lehmann S, Jackson A, Lautrup B (2008) A quantitative analysis of indicators of scientific performance. *Scientometrics* **76:** 369–390

Lovegrove BG, Johnson SD (2008) Assessment of research performance in biology: how well do peer review and bibliometry correlate? *Bioscience* **58:** 160–164

Moed HF (2005) *Citation Analysis in Research Evaluation*. Dordrecht, The Netherlands: Springer

Notkins AL (2008) Neutralizing the impact factor culture. *Science* **322:** 191

Petsko GA (2008) Having an impact (factor). *Genome Biol* **9:** 107

Prathap G (2006) Hirsch-type indices for ranking institutions' scientific research output. *Curr Sci* **91:** 1439

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

Rousseau R (2008) Reflections on recent developments of the *h*-index and *h*-type indices. In *Proceedings of WIS 2008, Berlin. Fourth International Conference on Webometrics*. Berlin, Germany: Informetrics and Scientometrics & Ninth COLLNET Meeting

Schreiber M (2007) Self-citation corrections for the Hirsch index. *Europhys Lett* **78:** 30002

Schubert A (2007) Successive *h*-indices. *Scientometrics* **70:** 201–205

Simons K (2008) The misused impact factor. *Science* **322:** 165

van Raan AFJ (2005) Measurement of central aspects of scientific research: performance, interdisciplinarity, structure. *Measurement* **3:** 1–19

van Raan AFJ (2006) Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups. *Scientometrics* **67:** 491–502

Lutz Bornmann (left) and Hans-Dieter Daniel are at the ETH Zurich, Switzerland. Hans-Dieter Daniel is also at the University of Zurich.
E-mails: bornmann@gess.ethz.ch; daniel@evaluation.uzh.ch