The Inflation of Impact Factors of Scientific Journals

Wolfgang G. Stock*[a]

For many years, the scientific community made use of the Journal Impact Factor, created by Garfield and Sher in 1963 and commercially distributed by the Institute of Scientific Information (ISI—today part of Thomson Reuters), to approximate the standing of an academic periodical. Nowadays, we are confronted with a multitude of different impact factors, namely, the (old) ISI Impact Factor, the (new) five-year Thomson Reuters impact factor, the trend line of Elsevier's Scopus, the H index of journals, the SCImago index by the Spanish SCImago Group, and the Eigenfactor score created by Carl Bergstrom. How can we keep an overview? What do these indices measure? What does "standing of a journal" mean: its impact on the scientific community, its perceived reputation, its value, its quality, its prestige, its influence? Is there a kind of toolbox with different infometric tools to describe and evaluate scholarly journals? Is there a leading indicator? And—above all—are journal impact indicators really useful?

Journal Impact Indicators

First of all, we are going to overview main currently used journal impact indicators (see Table 1). All indicators are based on two aspects, namely, on the number of publications of a journal and the number of citations those publications received. In short, publications are marks of a journal's effort or activity whereas citations are a mark of a journal's effect or impact. But is not the main success factor of a journal to be read by a lot of people? In academia, most of the readers are writers as well. And scholars cite—this is an unex-

[a] Prof. Dr. W. G. Stock
Department of Information Science
Heinrich-Heine-University Düsseldorf
Universitätsstr. 1, 40225 Düsseldorf (Germany)
Fax: (+49) 211-81-12917
E-mail: stock@phil-fak.uni-duesseldorf.de

Table 1. Main journal impact indicators.[a]

Indicator	Data Source	Supplier	Fee or free?
Group 1: Simple counting			
total publications in WoS	WoS	JCR, WoS	fee
total citations in WoS	WoS	JCR, WoS	fee
total publications in Scopus	Scopus	Scopus	fee
		SCImago	free
total citations in Scopus	Scopus	Scopus	fee
		SCImago	free
H index (WoS)	WoS	WoS	fee
H index (Scopus)	Scopus	SCImago	free
Group 2: Citation rates (cites per publication)			
citations per publication in WoS	WoS	WoS	fee
citations per publication in Scopus	Scopus	SCImago	free
trend line	Scopus	Scopus	fee
2-year impact factor	WoS	JCR	fee
5-year impact factor	WoS	JCR	fee
Group 3: Eigenvector-oriented indicators (position in journal network)			
eigenfactor score (ES)	JCR	JCR	fee
		Eigenfactor.org	free
article influence score	JCR	JCR	fee
		Eigenfactor.org	free
SJR	Scopus	SCImago	free

[a] WoS: Web of Science; JCR: Journal Citation Reports; SJR: SCImago journal rank.

pressed norm in science—when they write professional articles. When authors do not cite correctly or when they do read, but do not write (and hence do not cite), journal impact indicators have a serious problem.[1] The calculation of indicators presupposes that there are comprehensive digital databases collecting (in the ideal case) all scientific documents and all citations. This ideal case is far away from reality. For example, the British Library holds more than 40000 scientific serials and adds about 800 new journals each year, whereas the two most comprehensive multidisciplinary databases, namely, Elsevier's Scopus and Thomson Reuters' Web of Science (WoS), cover only 16000 (Scopus) and 10000 (WoS) periodicals. Most of all scientometric studies—and nearly all studies concerning the journal impact—are based either on Scopus or on WoS.

Very simple indicators of the journal impact count the number of publications and citations in a certain time frame, using WoS or Scopus data. According to WoS, *ChemPhysChem* has published 2023 documents since its beginning in 2000—and the journal has been cited 23680 times in journals which are covered in WoS. In Scopus, you will find different figures for the citations (3185 for the time span 2000–2007), because Scopus indexes other journals than WoS.

In 2005, Jorge E. Hirsch published a paper on an index to quantify a scientist's research output.[2] A scientist has an index *H* if *H* of his *N* papers have at least *H* citations each and the other *N*−*H* papers have fewer than *H* citations each. The *H* index of a journal follows the same method.[3] It is the number of articles *H* which have at least *H* citations. However, the *H* indexes reported by WoS and Scopus for a particular journal are usually different, which can be explained by the different sets of journals considered in each case.

Journal impact indicators that work with citation rates (cites per publication

over a time period) are more elaborated. Here the average citation rate of a document published in a journal—or the citation rate of a so-called "citable item" (which is very different from all documents, because mainly research articles and review articles come into consideration)—can be calculated. The average citation rate per item of *ChemPhysChem* is 11.71 (if we believe in WoS) or 3.39 (if we believe in SCImago and the Scopus data).

The most popular journal impact indicator is the Journal Impact Factor from ISI (later Thomson Scientific, now Thomson Reuters).[4] It is published in Thomson Reuters' Journal Citation Reports (JCR) database.[5] Garfield's impact factor is a cites-per-publication indicator and is based on two aspects, the number of citations (C) in a certain year (t) to items published in a journal (J) in the previous two years (numerator) and the number of citable items of the journal in the same two years (S) as denominator. If t is the base year and J the journal, then the impact factor (IF) of journal J is calculated using Equation (1):

$$IF(J,t) = C_{(t)}/[S_{(t-1)} + S_{(t-2)}] \qquad (1)$$

For example, if Journal J has published 200 citable items in 2007 ($t-1$) and 250 citable items in 2006 ($t-2$) (i.e., 450 documents in sum), and all the articles of the 2006 and 2007 volumes of Journal J have been cited 1500 times in 2008 (t), then the 2008 impact factor of Journal J will be 1500:450 = 3.333. It is also possible to calculate the IF without the self-citations (i.e., without considering the articles published in Journal J citing items from Journal J), sometimes called "journal incest."

The calculation of the 5-year impact factors follows the same method; the only difference is the time frame. $C'_{(t)}$ counts the citations to items published in the previous five years in the base year t, and the denominator sums up the source articles of the same five years [Eq. (2)]:

$$5y\text{-}IF(J,t) =$$
$$C'_{(t)}/[S_{(t-1)} + S_{(t-2)} + S_{(t-3)} + S_{(t-4)} + S_{(t-5)}] \qquad (2)$$

Scopus publishes the trend line (TL) indicator of journals. It is very easy to calculate but somehow blunt. The trend line is the fraction of the number of all citations (allC) to a journal J in a year t and the number of all papers published in J in the same year [Eq. (3)]:

$$TL(J,t) = allC(t)/S(t) \qquad (3)$$

Since there is no connection between the two factors in the fraction, this indicator seems to be relatively unhelpful regarding its theoretical justification and practical use.

The trend line should not be confused with the Immediacy Index (II) of the JCR, which is a valuable indicator. The II only counts the citations to J in the same year t and is an indicator of the velocity of the information transfer from a journal into the scientific community [Eq. (4)]:

$$II(J,t) = C(t)/S(t) \qquad (4)$$

Following the success of Google's PageRank ranking algorithm, two new indicators have been created, namely, the Eigenfactor score and the SCImago Journal Rank. For the Impact Factor, the H index, the trend line etc., all citations are equal with no regard to their origin. It is a plausible idea that a citation from a high-prestige journal (say, *Nature* or *Science*) should count more than one from a regional industry newsletter. Google's PageRank has an intuitive justification:[6] There is a so-called "random surfer" who starts with a Web page at random and keeps clicking on links (he never hits "back" but eventually starts at random on another page). The probability that this random surfer will visit a single Web page is its PageRank. The probability that the surfer stops and starts with a new page is the damping factor d. If we see the whole WWW as a large link matrix, the PageRank corresponds to the principal eigenvector of the normalized link matrix. Let us assume that the Web pages T_1 to T_n link to the Web page A, that every T_i has a PageRank PR(T_i) and a number of outgoing links $C(T_i)$, and finally that there is a damping factor d of, say, 0.85. Then, the PageRank of A is given by Equation (5):

$$PR(A) = (1-d) + d*[PR(T_i)/C(T_i)$$
$$+ \ldots + PR(T_n)/C(T_n)] \qquad (5)$$

While Google uses the network of links on the Web, the SCImago journal rank indicator[7] and the eigenfactor score[8] use citations in the scientific literature. Eigenfactor works with raw data from JCR, whereas SCImago applies Scopus data. Both new journal impact indicators are Google-like and oriented in the eigenvectors of the journals in the journal citation matrix, but they not only use different raw data but different calculation formulas as well. SCImago is based on the transfer of prestige from a journal to another, where the journal "prestige" is calculated by the number of citations and the prestige of the citing journals. Journal self-citations are included. The eigenfactor metrics gets its theoretical justification by the "random researcher": he goes to a library and selects a journal article at random. From this article, he selects a reference and goes to the cited article. He reads the article and selects a reference—and so on indefinitely. The frequency our random researcher visits a journal, expressed as a percentage, gives the measure of that journal's importance. Journal incest is excluded from the eigenfactor calculation. The Article Influence (AI) score is a measure of the per-article influence of the journal; it is scaled with respect to a mean of 1. *Science*, for example, has an AI score of 17, which means that the journal has 17 times the influence of the mean JCR journal.

Apart from differences due to data sets (Scopus versus WoS) both indicators correlate highly with the classical 2-year ISI impact factor.[9] Therefore, little is won.

Problems of Journal Impact Indicators

All indicators that work with relative frequency measures (i.e., all Group 2 indicators) suffer from serious statistical problems. It is a precondition for calculating average values (in our case: average cites per publication) that there is a Gaussian distribution: most values in the middle, where the average value is locat-

ed, some in the right hand side of the curve and more or less the same amount of values in the left hand side of the distribution (e.g., the distribution of human body heights). In journal informetrics, this is not the case. All known empirical distributions of papers by citation frequency follow two typical shapes: a power-law distribution (with a strong decline of values after the top scorer) or an inverse-logistic distribution (with some top scorers, but a strong decline of values behind those top values).[10] All these distributions end with a "long tail" of few or never cited articles. The distribution of *ChemPhysChem* articles by their citation numbers (Figure 1) is a typical power law. One
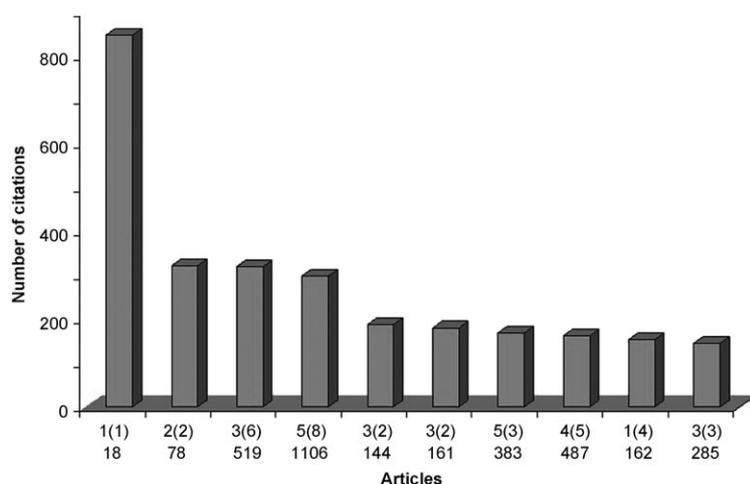
nonparametric und are therefore immune to our statistical problem. **In no case is it possible to use a journal impact factor on the article level to evaluate the influence of an article, an author or an institution.**[12]

What is the unit of publications? And of citations? What does "1" mean? Does it mean that 1 research article = 1 text book = 1 review = 1 letter to the editor? Obviously not.[13] All journal impact factors mentioned do not try to get to the bottom of the publication and citation unit problem. There are approaches to calculate relative impact factors, which are not geared to the number of source articles, but to the number of pages or even to the number of printed letters.[14]

nominator of the formula). Journals with a large amount of "letters" and other non-citable items are overestimated in their impact factors.

- There are differences between the impact factors of journals which publish mainly review articles and those publishing research results (in favour of review journals). Comparisons between both groups are not allowed.
- The representativeness of scientific disciplines (of languages, of countries) is sometimes very different in the databases.
- There are discipline-specific citation behaviors. Therefore, it is not possible to compare journal impacts factors across disciplinary borders.
- Data (in WoS and in Scopus) are not clean. We can notice typing errors and missing citations.

Use of Journal Impact Indicators

Who can profitably use certain sets of journal impact factors? I see four different user groups: 1) For journal editors and publishing houses the impact factors give market research information. Editors are able to approximate "their" journals' standing in comparison to other thematically closely related periodicals. 2) For libraries impact factors are useful for collection development, especially in combination with further indicators, primarily journal prices. Combined indicators, for example, Euros per unit of impact factor or a combination of this indicator with cost-per-use, may be effective for serial selection.[15] 3) Authors get hints on journals in which they can publish their research results (in case there are any authors in academia who do not know "their" appropriate periodicals). 4) Journal scientometrics is a scientific branch of information science and science of science. Since journals play an important role in the scientific communication process, it is necessary for informetrics and scientometrics to study scholarly periodicals.

Is there a leading indicator for journal impact? Unfortunately not. There are some publication–citation-based scores which are different because of the vari-

Figure 1. Distribution of articles in *ChemPhysChem* by their numbers of citations. The articles are identified by volume, issue (in brackets), and first page.

famous article (by Shipway, Katz, and Willner)[11] has been cited about 850 times in WoS, the second and third articles have about 320 cites and so on, ending in a tail of few or never cited articles. The average value of citations per item is 11.71—but this value is located at the beginning of the long tail, so it is not a suitable average. This statistical problem is inherent to all indicators that are based on normalization to the arithmetic mean (citations per publications). These include the Scopus' trend line, the 2-year impact factor, the 5-year impact factor, and the article influence score. The H index, the SCImago journal rank, the eigenfactor score, and the other Group 1 simple counting indicators are

In those cases, we get completely different values for the journal impact.

There are many further problems with the methodology of journal impact factors. I will mention only some examples:

- Not all academic journals are indexed in WoS or Scopus. We do not only miss the articles of those non-WoS and non-Scopus periodicals but their references as well.
- The selection of citable items leads to a bias. Only research papers and reviews count as "citable" (and are included in the denominator of the impact factor formula), but all papers (such as letters to the editors) can be cited (all cites are included in the

ous data sets and the diverse methods of calculation. Is it possible to describe and evaluate a journal only with such a publication–citation-based indicator set? Again, it is not. It is not possible to measure the standing of a scientific journal using only one single dimension.[16] And it is very difficult to evaluate the journal's prestige by data sets of multi-disciplinary databases, because they are not complete. Additionally, we should work with bibliographic and citation data of subject-specific databases, for example, Chemical Abstracts (SciFinder) for chemistry.[17] Furthermore, we are in need of more indicators, for example, download numbers of articles, journal circulation, importance of the publishing house, even surveys of journal reading behaviour and application of the read information in practice.[18] Since every single indicator has some benefits, but some pitfalls as well, we have to accept the rule: Be careful! Handle with caution!

[1] B. Scanlan, *Serials Libr.*, **1987**, *13*, 57–66.
[2] J. E. Hirsch, *Proc. Natl. Acad. Sci. USA*, **2005**, *102*, 16569–16572.
[3] a) T. Braun, W. Glänzel, A. Schubert, *The Scientist*, **2005**, *19*, 8; b) Grupo SCImago, *El Profesional de La Información*, **2006**, *15*, 304–306.
[4] a) E. Garfield, I. H. Sher, *Am. Doc.*, **1963**, *14*, 195–201; b) E. Garfield, *Science* **1972**, *178*, 471–479; c) E. Garfield, *Curr. Comments* **1994**, no. 25, 3–7; d) E. Garfield, *Curr. Comments* **1994**, no. 29, 3–5; e) E. Garfield, *JAMA J. Am. Med. Assoc.* **2006**, *295*, 90–93; f) E. Archambault, V. Larivière, *Scientometrics* **2009**, *79*, 635–649; g) A. J. Flisher, *S. Afr. Med. J.* **2009**, *99*, 226–228.
[5] W. G. Stock, *Password*, **2001**, no. 5, 24–39.
[6] a) S. Brin, L. Page, *Comput. Networks ISDN* **1998**, *30*, 107–117; b) L. Page, *patent no. US 6.285.999*, applied **1998**, granted **2001**.
[7] a) Grupo SCImago, *El Profesional de La Información* **2007**, *16*, 645–646; b) P. Jacsó, *Online* **2009**, *33*, 50–53.
[8] a) C. T. Bergstrom, *Coll. Res. Libr.* **2007**, *68*, 314–316; b) C. T. Bergstrom, J. D. West, M. A. Wiseman, *J. Neurosci.* **2008**, *28*, 11433–11434.
[9] a) M. E. Falagas; V. D. Kouranos, R. Arencibia-Jorge, D. E. Karageorgopoulos, *FASEB J.* **2008**, *22*, 2623–2628; V. D. Kouranos, R. Arencibia-Jorge, D. E. Karageorgopoulos, *FASEB J.* **2008**, *22*, 2623–2628; b) A. Fersht, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 6883–6884; L. Leydesdorff, *J. Am. Soc. Inf. Sci. Technol.* **2009**, *60*, 1327–1336.
[10] W. G. Stock, *J. Am. Soc. Inf. Sci. Technol.* **2006**, *57*, 1126–1129.
[11] A. N. Shipway, E. Katz, W. Willner, *ChemPhysChem* **2000**, *1*, 18–52.
[12] P. Seglen, *Br. Med. J.* **1997**, *314*, 497.
[13] W. G. Stock, *Publikation und Zitat. Die problematische Basis empirischer Wissenschaftsforschung*, University of Applied Sciences, Cologne, **2001**.
[14] a) S. J. Liebowitz, J. C. Palmer, *J. Econ. Lit.* **1984**, *22*, 77–88; b) D. N. Laband, M. J. Piette, *J. Econ. Lit.* **1994**, *32*, 640–666.
[15] a) K. G. Altmann, G. E. Gorman, *Libr. Acquis.-Pract. Th.* **1998**, *22*, 147–159; b) H. K. Chung, *J. Acad. Libr.* **2007**, *33*, 393–402.
[16] S. Haustein, *12th International Conference on Scientometrics and Informetrics (ISSI 2009)*, Doctoral Forum, **2009**.
[17] C. Neuhaus, H. D. Daniel, *Scientometrics* **2009**, *78*, 219–229.
[18] a) C. Schlögl, W. G. Stock, *J. Am. Soc. Inf. Sci. Technol.* **2004**, *55*, 1155–1168; b) C. Schlögl, W. Petschnig, *Libr. Collect. Acquis.* **2005**, *29*, 4–32; c) K. Juchem, C. Schlögl, W. G. Stock, *Information—Wissenschaft und Praxis* **2006**, *57*, 31–37; d) C. Schlögl, W. G. Stock, *J. Doc.* **2008**, *64*, 643–666.

ESSAY

Keeping an overview: Nowadays, we are confronted with a multitude of different impact indicators. This essay discusses some of the problems of citation analysis—and of indicators claiming to give a measure of the standing of a scholarly periodical. Incomplete data sets and the (mostly forgotten) highly skewed distribution of papers by cites (see figure) are two of the most critical points of the journal impact calculation.
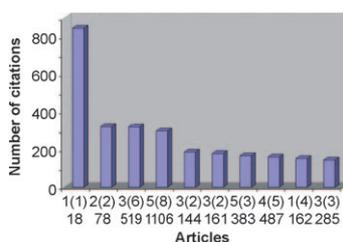
W. G. Stock*

■■ – ■■

The Inflation of Impact Factors of Scientific Journals