A Case Study of the Modified Hirsch Index h_m Accounting for Multiple Coauthors

Michael Schreiber
Institut für Physik, Technische Universität Chemnitz, 09107 Chemnitz, Germany.
E-mail: schreiber@physik.tu-chemnitz.de

J.E. Hirsch (2005) introduced the *h*-index to quantify an individual's scientific research output by the largest number *h* of a scientist's papers, that received at least *h* citations. This so-called Hirsch index can be easily modified to take multiple coauthorship into account by counting the papers fractionally according to (the inverse of) the number of authors. I have worked out 26 empirical cases of physicists to illustrate the effect of this modification. Although the correlation between the original and the modified Hirsch index is relatively strong, the arrangement of the datasets is significantly different depending on whether they are put into order according to the values of either the original or the modified index.

Introduction

The *h*-index was proposed by Hirsch (2005) as an easily determinable estimate of the impact of a scientist's cumulative research contribution. It is defined as the highest number of papers of a scientist that have been cited *h* or more times. Due to its simplicity, it soon became famous and attracted attention, which can be quantified by the high number of about 200 citations which the article (Hirsch, 2005) accumulated within 3 years, thus enhancing Hirsch's *h*-index. Of course, it is always dangerous to reduce the complete scientific output of the researcher to a single number. Nevertheless, it is a fact that the *h*-index is more and more utilized to quantify the visibility, importance, significance, and broad impact of individual scientists, departments, countries, or research fields. Its calculation has even been implemented in the ISI *Web of Science* (*WoS*) provided by Thomson Scientific; however, this automatic determination has to be taken with considerable caution because an undiscriminating *WoS* search often leads to a completely wrong database, usually comprising too many papers due to homographs but sometimes too few papers due to misspellings, transliterated names, name changes, or similar difficulties. This is called the *precision problem*, which is well-known in bibliometrics and scientometrics and has been noted in numerous studies. For the present dataset, it was discussed in detail by Schreiber (2007a).

Advantages and disadvantages of the *h*-index have been discussed by many authors. Already less than 1 year after the introduction of the *h*-index, a review of the research literature on this topic was published (Bornmann & Daniel, 2007). Since then, several variants of the *h*-index have been introduced: for example, the *g*-index, which is sensitive to one or several outstanding, highly cited manuscripts (Egghe, 2006); the *A*-index and the *R*-index, which measure the citation intensity in the *h*-core (i.e., the *h*-defining set of papers) (Jin, Liang, Rousseau, & Egghe, 2007); or the h_s-index, which takes self-citation corrections into account (Schreiber, 2007b). Recently, nine different variants of the *h*-index were compared (Bornmann, Mutz, & Daniel, 2008).

One disadvantage of the *h*-index is its insensitivity to the numbers of coauthors of a given publication, as already noted by Hirsch (2005). He proposed "to normalize *h* by a factor that reflects the average number of coauthors" (p. 16571). This normalization was applied by Batista, Campiteli, Kinouchi, and Martinez (2006) using the mean number of authors of the papers in the *h*-core for the normalization, and the resulting index was labeled h_I; however, Batista et al. already noted that the average is sensitive to extreme values. This means that the influence of single-author publications to one's *h*-index can be strongly reduced. On the other hand, a few papers with a large number of coauthors will lead to an excessively large normalization (Schreiber, 2008a).

In scientometrics, the problem of how to count multi-authored publications has been discussed for a long time (Lindsey, 1980; Price, 1981), assigning credit proportionally to the number of authors, which is usually called *fractional counting* or *adjusted counting*; however, a number of different methods for accrediting publications for several authors have evolved as discussed e.g., by Egghe, Rousseau, & Van Hooydonk (2000). It is widely accepted

Received August 21, 2008; revised January 20, 2008; accepted January 22, 2009

 • Published online 20 March 2009 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/asi.21057

that some kind of discounting should be applied (Burrell & Rousseau, 1995; Egghe et al., 2000; Harsanyi, 1993; Lukovits & Vinkler, 1995; Trueba & Guerrero, 2004) also to the Hirsch index (Burrell, 2007a; Egghe, 2008; Jin et al., 2007; Wan, Hua, & Rousseau, 2007). One difficulty is that different scoring methods can lead to paradoxical effects and yield totally different rankings (Egghe et al., 2000; Van Hooydonk, 1997) so that no unambiguous solution of the "multiple-author problem" (Harsanyi, 1993) exists. But fractional counting is usually preferred since it does not increase the total weight of a single paper (Egghe et al., 2000). Egghe and Rousseau (1990) stated already "that the best way to handle multi-authored papers is to assign credit proportionally" (p. 223).

I recently proposed to modify the h-index by counting the papers fractionally according to (the inverse of) the number of authors yielding the modified index h_m (Schreiber, 2008a). Analyzing the citation records of eight famous physicists, it was shown that this can have a significant influence and lead to a different ranking than that using the original h-index. The same fractional counting of papers has been suggested by Egghe (2008) and applied to two fictitious examples and one empirical case; however, the effect was relatively small because of a large number of single-author papers in his dataset. It is the aim of the present article to demonstrate the effect of the fractionalized counting on the citation records of 26 not-so-prominent physicists. Thus, the obtained observations should be more common for the datasets of "more average" scientists.

The validity of a new index should be analyzed on the basis of empirical data. This is the purpose of the present investigation, which is restricted to the comparison of the original Hirsch index, the normalized index h_I, and the modified index h_m. In principle, most other ways available in the literature to distribute credit among coauthors could be applied to the Hirsch index as well. But these ways are usually more complicated and require assumptions about the relative contribution of different authors. Without specific information about, for example, a particularly high contribution of the first or the last author in the author list, I believe that the fairest way of attributing the credit is to share it among all coauthors as it is applied in the following case study.

The Database, the Computation of the Modified Index h_m, and Its Visualization

Data for the subsequent analysis were compiled in January and February 2007 from the *Science Citation Index* provided by Thomson Scientific in the *WoS*. The 26 citation records were analyzed with respect to the self-citations (Schreiber, 2007a). As specified in that publication, the 26 datasets include the records of all full and associate professors from the Institute of Physics at my university, some recently retired colleagues, and all scientists who have been working as assistants or senior assistants in my group doing their research for their habilitation degree or afterwards. The datasets are labeled A, B, C, ..., Z in conformity with the previous analysis (Schreiber, 2007a).

The same data were utilized for an investigation of the g-index in comparison with the h-index, the A-index, and the R-index (Schreiber, 2008b). As mentioned in the Introduction, the precision problem means that a simple *WoS* search is not sufficient. For the present investigation, reasonably great care was taken to establish the correct database. As detailed earlier (Schreiber, 2007a), homographs yield an enhancement of the h-index in nine cases; in six cases, this was 50% or more, and in one case, there was even a factor of 2.73. For two datasets, the reverse problem was encountered because important publications were missed by the general search in the ISI *WoS*.

The *WoS* allows an automatic arrangement of the publication lists in decreasing order according to the number of citations $c(r)$, where r is the rank attributed to the paper. The h-index is readily read off this list as

$$c(h) \geq h \geq c(h+1) \tag{1}$$

according to Hirsch's (2005) original definition. Note that an anonymous referee as well as some authors (Burrell, 2007b) argue that the original definition demands that the papers beyond the rank h should "have fewer than h citations each." This wording was indeed chosen by Hirsch in the first preprints on the server arXiv:physics (3 Aug 2005). However, this formulation would make the index not quite well-defined in all cases. But in the last version available on the preprint server arXiv:physics (29 Sep 2005) as well as in the final publication (Hirsch, 2005), the phrase has been corrected, requiring now that the papers beyond the rank h should "have $\leq h$ citations each" (p. 16569) with which the second inequality in Equation 1 is in conformance.

In Equation 1, each paper is fully counted for the (trivial) determination of its rank

$$r = \sum_{r'=1}^{r} 1. \tag{2}$$

The upper histogram in Figure 1 shows the citations arranged in this way for Dataset C. The intersection with the white line, which displays the function

$$c(r) = r, \tag{3}$$

yields the h-index. For the dataset in Figure 1, one obtains $h^C = 23$.

If one counts a paper with $a(r)$ authors only fractionally [i.e., by $1/a(r)$], one obtains an effective rank

$$r_{\text{eff}}(r) = \sum_{r'=1}^{r} \frac{1}{a(r')}. \tag{4}$$

This can be utilized to define the modified index h_m as

$$c(r(h_m)) \geq h_m \geq c(r(h_m)+1) \tag{5}$$

DOI: 10.1002/asi

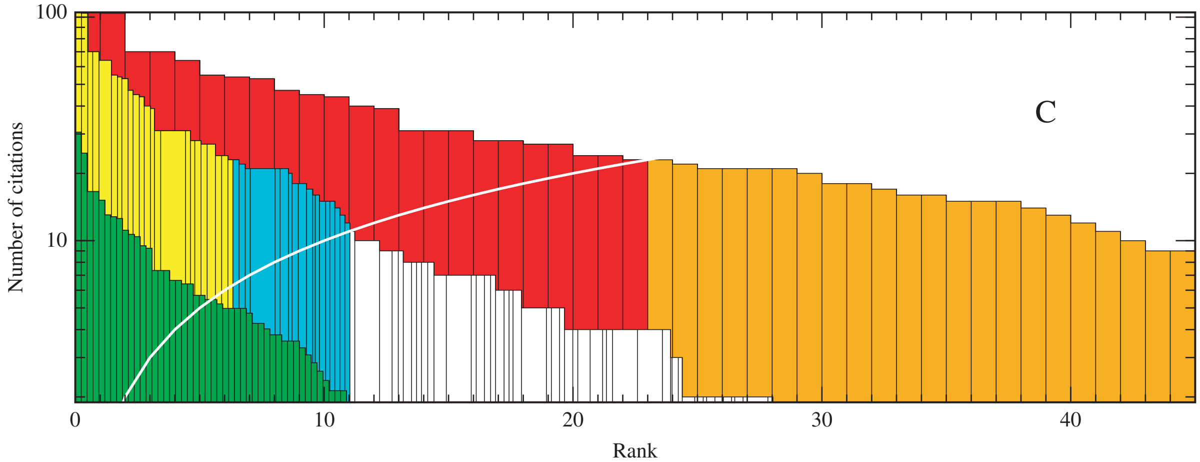

FIG. 1. The citation counts for the papers in Dataset C. The upper histogram with wide bins shows the numbers of citations $c(r)$ versus the rank r which is attributed to each paper by sorting according to $c(r)$, up to the h-index (red/dark grey) and beyond (orange/medium grey). In the middle histogram, the effective rank is used so that the original histogram is compressed towards the left (yellow/light grey for the first h papers, turquoise/medium grey up to the h_m-th paper, and white beyond). In the lower histogram, the normalization with the mean number of authors of the first h papers is used so that the original histogram is compressed to the left as well as downwards (light green/medium grey up to the h_I-index and dark green/dark grey beyond). Note the logarithmic scale for $c(r)$. The thick white line displays the function $c(r) = r$ so that its intersections with the histograms (from top to bottom) yield the h-index, the h_m-index, and the h_I-index, respectively.

where $r(h_m)$ follows from the inverse function $r(r_{\text{eff}})$ of Equation 4. This means that h_m is that effective number of papers which have been cited h_m or more times while the further papers have no more than h_m citations each. It is easy to visualize this definition by plotting the respective histogram with bin widths, which are determined by (the inverse of) the number of authors for each paper, as shown in the middle histogram in Figure 1. This leads to a significant compression of the histogram towards lower values of the rank. Correspondingly, the effective number of papers in the h-core is much smaller than h; for the data in Figure 1, one obtains $r^C_{\text{eff}}(h^C) = 6.33$.

Therefore, beyond h^C, there are papers with more citations than this effective rank. These have to be taken into account for the modified index, as visualized in Figure 1. It is indeed nearly always a considerable number of publications with citation counts between h and h_m which contribute to the h_m-core (i.e., to the h_m-defining set). In the considered example, the h^C_m-core comprises $r(h^C_m) = 41$ publications with an effective rank $h^C_m = r^C_{\text{eff}}(41) = 11.03$. This value also can be read off the intersection of the function $c(r) = r$ with the middle histogram in Figure 1. It means that there are 41 publications with at least 12 citations in the h^C_m-core. The 42nd paper attracted only 11 citations, in agreement with Equation 5.

For a large dataset, the visualization of the citation records as in Figure 1 is probably the easiest way to present the data and to allow an assessment of the influence of multiple coauthors. To demonstrate the calculation of the individual indices in more detail, the citation records of four scientists with a relatively small number of publications are presented in Table 1. There, only the 20 most cited publications are included for each dataset. These are sufficient for determining not only the original Hirsch index but also the modified index. The original Hirsch index is easily read off this table as the highest rank r for which the citation count $c(r)$ is larger than or equal to the rank. The values of the effective rank $r_{\text{eff}}(r)$ in this table are determined according to Equation 4; that is, counting the papers fractionally according to the number of authors for each paper. Consequently, the modified index also can be easily read off this table as the largest effective rank for which the citation count is larger than or equal to this effective rank.

The four datasets in Table 1 represent quite different citation records, Dataset X is characterized by a very high citation count of the first publication, and also in Case V, $c(1)$ is quite large. Datasets V and W show a large tail; that is, the number of citations for the last papers in the table is relatively high, which is reflected in the observation that 6 and 9 papers enter the h_m-core in addition to the 10 and 9 papers in the h-core, respectively. For Dataset X, the number of coauthors is quite large for all publications, yielding of course small effective ranks, which in turn lead to a large increase of the size of the h^X_m-core from $h^X = 8$ to 15 papers, which is, however, accompanied by a relatively small modified index because the citation counts drop quite strongly in this range. An even stronger decrease of the citation counts can be observed for Dataset Y, but in this case, the number of coauthors is relatively small (on average less than 2), so that only one more paper contributes to the h^Y_m-core in addition to the 7 papers in the h^Y-core. It is therefore not surprising that the resulting modified index $h^Y_m = 4.83$ is significantly larger than $h^X_m = 2.95$ and also larger than $h^W_m = 4.33$.

DOI: 10.1002/asi

TABLE 1. The citation records for the 20 most cited papers in Datasets V, W, X, and Y. Besides the number of citations $c(r)$ and the number of authors $a(r)$, the effective rank $r_{\text{eff}}(r)$ as defined in Equation 4 is given. For each case, the citation count of the last paper which contributes to the h-core (i.e., fulfilling Equation 1) is given in boldface and italics, and the citation count of the last paper which enters the h_{m}-core (i.e., which fulfills Equation 5) is given in boldface.

r	$c^{\text{V}}(r)$	$a^{\text{V}}(r)$	$r_{\text{eff}}^{\text{V}}(r)$	$c^{\text{W}}(r)$	$a^{\text{W}}(r)$	$r_{\text{eff}}^{\text{W}}(r)$	$c^{\text{X}}(r)$	$a^{\text{X}}(r)$	$r_{\text{eff}}^{\text{X}}(r)$	$c^{\text{Y}}(r)$	$a^{\text{Y}}(r)$	$r_{\text{eff}}^{\text{Y}}(r)$
1	79	3	0.33	42	4	0.25	204	8	0.13	19	2	0.50
2	34	4	0.58	21	3	0.58	15	3	0.46	14	2	1.00
3	32	4	0.83	16	4	0.83	14	6	0.63	11	2	1.50
4	25	2	1.33	12	8	0.96	11	3	0.96	10	2	2.00
5	16	4	1.58	12	3	1.29	10	8	1.08	9	3	2.33
6	13	4	1.83	10	6	1.46	10	4	1.33	7	1	3.33
7	12	10	1.93	9	3	1.79	9	5	1.53	***7***	1	4.33
8	11	2	2.43	9	5	1.99	***8***	5	1.73	**6**	2	4.83
9	11	3	2.77	***9***	4	2.24	8	5	1.93	5	1	5.83
10	***11***	3	3.10	8	3	2.58	7	8	2.06	5	4	6.08
11	8	3	3.43	8	3	2.91	7	7	2.20	4	3	6.42
12	8	1	4.43	8	6	3.08	6	6	2.37	4	1	7.42
13	8	2	4.93	8	4	3.33	5	4	2.62	4	1	8.42
14	8	4	5.18	7	3	3.66	4	6	2.78	3	1	9.42
15	7	2	5.68	7	8	3.78	**3**	5	2.98	2	2	9.92
16	**7**	1	6.68	7	8	3.91	3	6	3.15	2	2	10.42
17	6	2	7.18	5	4	4.16	3	6	3.32	2	1	11.42
18	6	2	7.68	**5**	6	4.33	3	6	3.48	1	1	12.42
19	5	2	8.18	4	3	4.66	3	6	3.65	1	2	12.92
20	5	3	8.52	4	5	4.86	2	5	3.85	0	2	13.42

Computation of the Normalized Index h_{I} and Its Visualization

For the simple normalization of the h-index, the average number of authors of the first r papers is calculated as the mean

$$\bar{a}(r) = \frac{1}{r}\sum_{r'=1}^{r} a(r') \tag{6}$$

and utilized to determine

$$h_{\text{I}} = \frac{h}{\bar{a}(h)}. \tag{7}$$

If one employs Equation 6 for the (trivial) determination of a normalized rank

$$r_{\text{I}}(r) = \frac{1}{\bar{a}(r)}\sum_{r'=1}^{r} 1 = \frac{r}{\bar{a}(r)} \tag{8}$$

and utilizes respective normalized citation counts

$$c_{\text{I}}(r) = \frac{c(r)}{\bar{a}(h)}, \tag{9}$$

then one can determine the normalized index h_{I} in analogy to Equation 5 from the inequalities

$$c_{\text{I}}(r(h_{\text{I}})) \geq h_{\text{I}} \geq c_{\text{I}}(r(h_{\text{I}}) + 1) \tag{10}$$

where $r(h_{\text{I}})$ follows from the inverse function $r(r_{\text{I}})$ of Equation 8, and we have $r(h_{\text{I}}) = h$. Of course, this complicated calculation is not necessary because the simple definition (Equation 7) is sufficient. But I have made this detour to show that the straightforward normalization of the h-index not only means a scaling of the ranks (Equation 8) similar to Equation 4 but also a scaling of the citation counts (Equation 9). This can be easy visualized: Compare Figure 1 where the lowest histogram is compressed towards the left by the factor $1/\bar{a}^{\text{C}}(h) = 1/4.22 = 0.24$ as well as downwards with the same factor. This factor yields the normalized index $h_{\text{I}}^{\text{C}} = 5.45$, which also can be read off Figure 1 from the intersection of the function $c(r) = r$ with the lower histogram. Effectively, the calculation of the h_{I}-index thus means a fractionalized counting of the citations as well as a fractionalized counting of the papers—in each case, by the mean number of authors. Consequently, the impact is dramatically reduced, as noted by Vinkler (2007). This double normalization is at least questionable.

Results for the Different Indices

The citation counts for five further datasets are presented in Figure 2, and the determination of the indices is visualized.

For Dataset D, it is conspicuous that no single-author publications show up in the h-core and not even in the h_{m}-core, as can be seen in the middle histogram, in which all respective bars are compressed. (In fact, the first single-author paper appears at $r = 113$ with $r_{\text{eff}} = 21.75$ reflected by the wide white bar in the middle histogram at this rank.) This leads to a particularly strong normalization with $\bar{a}(h) = 6.05$, reducing the index from $h^{\text{D}} = 20$ to $h_{\text{I}}^{\text{D}} = 3.31$. But in comparison with Case C in Figure 1, there are many papers with citation counts below, but close to, h^{D}. Therefore, the upper histogram

DOI: 10.1002/asi

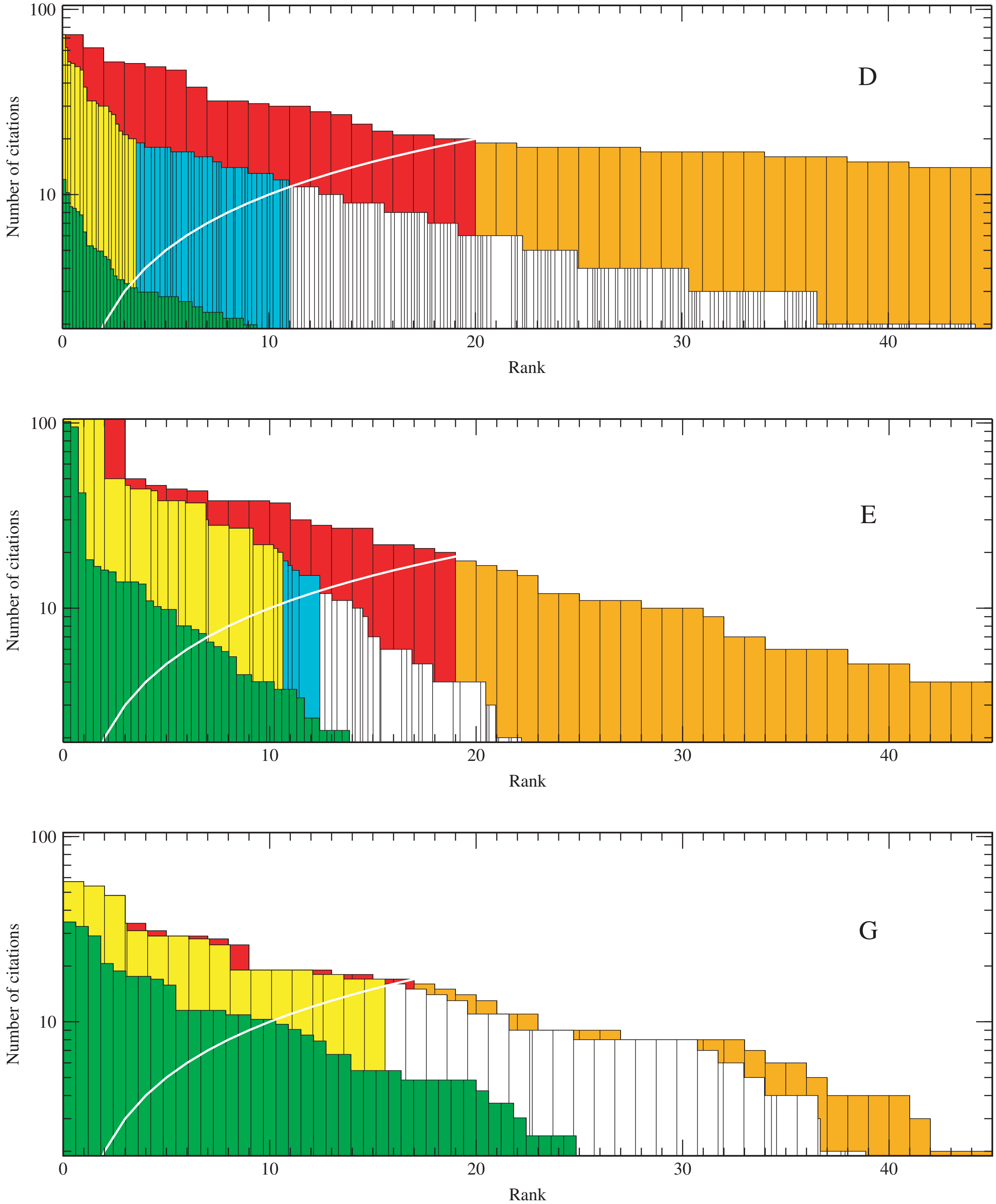

FIG. 2. Same as Figure 1, but for five different datasets: D, E, G, N, and P.

is rather flat beyond this value, which is of course reflected in the likewise quite-flat middle histogram for the fractional counting of papers. Consequently, $h_m^D = 10.97$ is very close to h_m^C.

In contrast, Dataset E comprises a large number of highly cited, single-author publications, so that the histogram for fractional counting remains close to the full histogram at least for small ranks. Thus, it is not surprising that for both variants of the h-index, the resulting values for this dataset are higher than those for Datasets C and D, even though the citation counts decrease quite strongly beyond h^E.

The average number of coauthors in the h-core for Datasets N and P is about the same as for E, so that the reduction between h and h_I is comparable. But the effect on the modified index is quite different. Due to the strong decrease of the citation counts $c^P(r)$, the h_m^P-core includes only one

DOI: 10.1002/asi

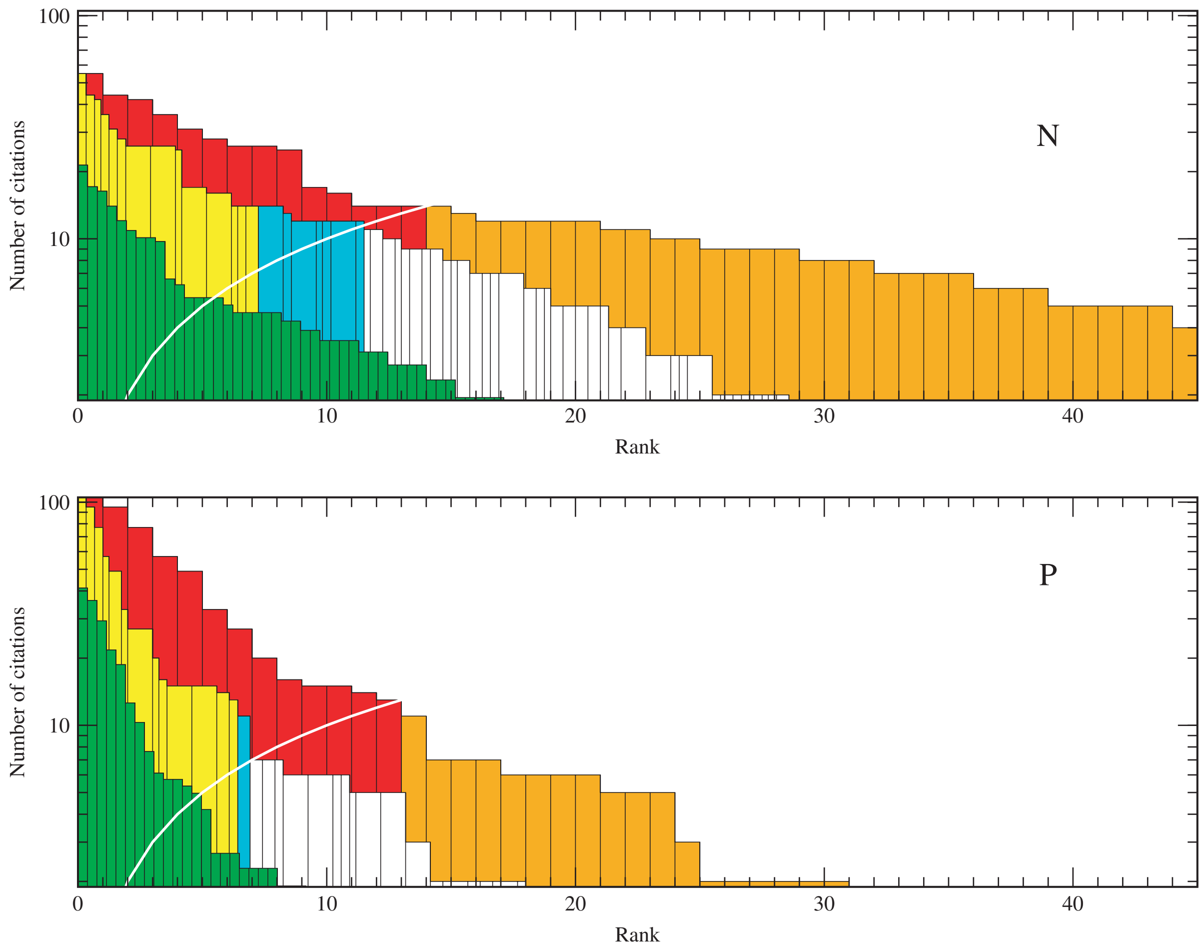

FIG. 2. Continued.

more paper than the h^P-core, leading to a relatively small value of $h_m^P = 6.92$. In contrast, seven papers, three of them single-author publications enter the h_m^N-core in addition to the 14 papers in the h^N-core. Accordingly, the resulting value for $h_m^N = 11.50$ is relatively large. These differences are visualized in Figure 2 by the different widths of the central section in the middle histograms; compare Case P with Case N.

A similar observation can be made comparing the central sections of the middle histograms in Figure 2 for Cases E and D. With 23 publications, the h_m^E-core comprises only four more papers than the h^E-core while for the h_m^D-core, 37 papers have to be added to the 20 publications in the h^D-core. But this difference is counterbalanced by the large number of single-author publications of Author E, so that $h_m^E > h_m^D$, as mentioned earlier.

An extreme case occurs for Dataset G because in this case, 15 of 17 publications in the h^G-core are single-author papers. Consequently, the upper histogram can barely be distinguished from the middle histogram. Moreover, there is no central section in the middle histogram because the h_m-core comprises the same publications as the h-core. Nevertheless, the reduction to $h_I^G = 10.3$ visualized by the compression in the lower histogram is quite strong because one paper with 11 authors has a strong effect on $\overline{a}(h)$. This is an example for the aforementioned fact that the average is sensitive to extreme values.

Effect of the Fractionalized Counting on the Ranking

The resulting values of the indices for the 26 datasets are compiled in Table 2. In Figure 3, the obtained indices are displayed on a logarithmic scale so that the relative changes can be easily visualized.

It is obvious that a strong reduction of the Hirsch index occurs, when the number of coauthors is taken into consideration as quantified by the ratio h_m/h, which is on average $<h_m/h> = 0.58 \pm 0.13$. The effect is very small only for Dataset G, as anticipated from the earlier discussion. Note that in this case, the omission of the 11-author publication would result in expectable reductions to $h^{\overline{G}} = 16$ and $h_m^{\overline{G}} = 15.5$ (no other papers enter the $h^{\overline{G}}$-core or the $h_m^{\overline{G}}$-core), but also in a surprising *increase* to $h_I^{\overline{G}} = 15.06$ because the mean number of authors decreases strongly to $\overline{a}(h) = 1.06$. This is a very strange effect: Neglecting a highly cited paper, in this case with $c^G(4) = 34$, leads to a significantly higher

DOI: 10.1002/asi

TABLE 2. Hirsch index without and with taking multiple coauthorship into account as described in the text. Also given are the average number $\overline{a}(h)$ of authors in the h-core, the compression $r_{\text{eff}}(h)$ of the h-core due to the fractionalized counting and the value $r(h_{\text{m}})$, reflecting the size of the h_{m}-core. The last column shows the order in which the datasets appear after the list is sorted according to the modified index.

Dataset	h	$\overline{a}(h)$	h_{I}	$r_{\text{eff}}(h)$	$r(h_{\text{m}})$	h_{m}	h_{m}/h	$\mathcal{O}(h_{\text{m}})$
A	39	5.87	6.64	7.45	90	16.95	0.43	2
B	27	2.78	9.71	11.36	42	17.81	0.66	1
C	23	4.22	5.45	6.33	41	11.03	0.48	6
D	20	6.05	3.31	3.57	57	10.97	0.55	7
E	19	2.74	6.93	10.64	23	12.43	0.65	4
F	18	4.94	3.64	4.29	39	9.32	0.52	8
G	17	1.65	10.30	15.59	17	15.59	0.92	3
H	16	5.00	3.20	4.36	27	5.81	0.36	21
I	15	5.27	2.85	4.01	28	6.93	0.46	16
J	15	4.80	3.13	3.72	21	5.67	0.38	22
K	14	4.14	3.38	3.96	25	7.38	0.53	13
L	14	3.21	4.36	5.06	22	8.38	0.60	10
M	14	4.43	3.16	5.56	23	7.14	0.51	15
N	14	2.57	5.45	7.25	21	11.50	0.82	5
O	13	3.38	3.85	4.68	22	7.80	0.60	11
P	13	2.62	4.96	6.42	14	6.92	0.53	17
Q	13	2.77	4.69	6.00	20	8.58	0.66	9
R	12	3.75	3.20	4.81	17	7.59	0.63	12
S	12	2.83	4.24	4.92	17	7.33	0.61	14
T	10	3.50	2.86	3.03	21	5.98	0.60	20
U	10	3.30	3.03	3.58	18	6.75	0.68	18
V	10	3.90	2.56	3.10	16	6.68	0.67	19
W	9	4.44	2.03	2.24	18	4.33	0.48	24
X	8	5.25	1.52	1.73	15	2.98	0.37	25
Y	7	1.86	3.76	4.33	8	4.83	0.69	23
Z	5	3.00	1.67	1.83	6	2.83	0.57	26

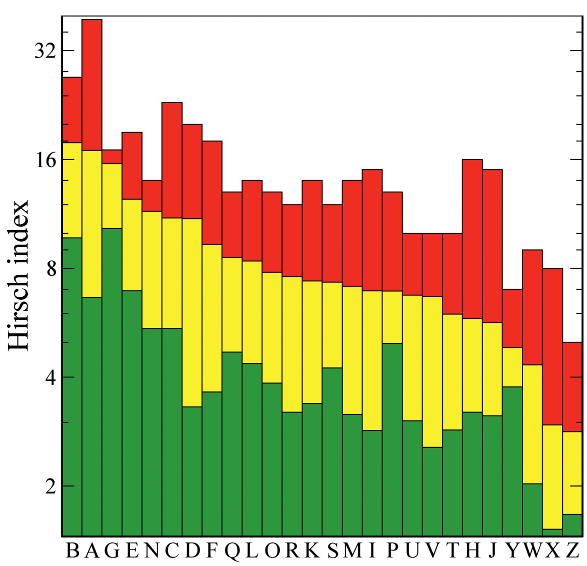

FIG. 3. Hirsch indices for the 26 investigated datasets. From top to bottom: original Hirsch index h (red/medium grey), modified index h_{m} due to fractional counting of the papers according to the number of authors (yellow/light grey), and index h_{I} determined by normalization of the h-index with the mean number of authors (green/dark grey). The datasets are ordered according to the modified index h_{m}, as indicated at the horizontal axis, where the letters are not in alphabetical order in contrast to the sequence in Table 2 determined by the original index h. Note the logarithmic scale for the h values.

h_{I}-index. For the original index, such strange behavior cannot occur nor is it possible for the modified index h_{m}. This is certainly an extreme case, but it is not unique. A closer inspection of the data for Case X in Table 1 shows that deleting the most cited paper (which incidentally has eight authors) from this dataset would lead to a similar effect: In this case, the Hirsch index would not change because another paper with $c(8) = 8$ citations enters the $h^{\overline{\text{X}}}$-core, but this has only five authors, so that the mean number of authors in the core decreases to $\overline{a}(h) = 4.88$. As a consequence, the normalized index increases to $h_{\text{I}}^{\overline{\text{X}}} = 1.64$, which is the same strange behavior as for the earlier example in Case G. On the other hand, as expected, the modified index decreases somewhat—in this case to $h_{\text{m}}^{\overline{\text{X}}} = 2.86$.

The relatively small changes for Dataset N were discussed in the previous section. Likewise, Dataset Y is characterized by a very small average number of coauthors and a corresponding small effect for h_{I}^{Y}. But the reduction of the modified index as compared to the original index is not so small because, again, a strongly decreasing citation record beyond $h^{\text{Y}} = 7$ allows only one publication to enter the h_{m}^{Y}-core in addition to the h^{Y}-core, as discussed previously.

The values $r_{\text{eff}}(h)$ in Table 2 reflect how strongly the h-core is compressed by the fractionalized counting of the papers. Of course, $r_{\text{eff}} > h_{\text{I}}$ in all cases because the average (Equation

DOI: 10.1002/asi

4) of the inverse numbers of authors is always larger than the inverse of the average (Equation 6) (also see Figures 1 and 2).

The values $r(h_m)$, which also are given in Table 2, show how many papers contribute to the h_m-core and thus demonstrate how many more papers beyond the h-core have to be taken into account. Here, the strong or weak decrease of the citation records in dependence on the rank as discussed earlier is significant. Comparing $r(h_m)$ with h gives an indication of how much more severe the precision problem becomes for the determination of the modified index because a significantly larger number of publications has to be evaluated.

The last column in Table 2 shows the rank which the datasets would hold after arranging them according to the modified index. Of course, the smaller changes should not be overinterpreted, but there are some very large rearranging effects, for example, Colleagues N and Q move forward nine and eight places, respectively, in the h_m-sorted list. On the other hand, Scientists H and J fall back 13 and 12 places, respectively. These rearrangements are illustrated in Figure 3, where the h values yield relatively low bars for Cases N and Q in the left part of the diagram and relatively high bars for Cases H and J in the right part.

The observations can be quantified by calculating Spearman's rank-order correlation coefficient. The value $\kappa(h,h_m) = 0.755$ shows that there is of course quite some correlation, which is not surprising because the ratio h_m/h is not so different for most datasets. The correlation $\kappa(h,h_I) = 0.597$ is significantly weaker. The correlation $\kappa(h_m,h_I) = 0.795$ is strongest. It is not unexpected that such significant correlations between the indices exist because the consideration of the multiple coauthorship does of course not lead to a completely different valuation of the citation records. That means that the indices might be called redundant, and one might be tempted to consider just the weakest correlation as an incremental contribution of the h_I-index compared with the h-index. However, in my opinion, it is important that some kind of discounting is only fair, especially when a large number of coauthors has contributed to a publication. Therefore, it appears to be appropriate to modify the Hirsch index, although the modified index is strongly correlated with the original index. Note that the correlation between the h-index and other variants such as the g-index and the R-index for the same datasets as in the present investigation is much higher (Schreiber, 2008b). The value $\kappa(h,h_m) = 0.755$ is of the same order as the correlation between the Hirsch index and the total number of papers or the total number of cited papers (Schreiber, 2008b).

Conclusions

The modified Hirsch index h_m (Schreiber, 2008a) was introduced to account for multiple coauthors in a reasonable way. It stands to reason that it is necessary to test the validity of a new index thoroughly on the basis of empirical data before using such an index for comparison. The present case study provides such an analysis. Of course, the values of the modified index are always smaller than those of the original index, and the reduction depends on the number of coauthors. Naturally, scientists with many single-author publications become more prominent in the h_m-sorted list. This is the main effect of the modification. But the individual citation records also can have a strong influence on the modified index because they determine how many more papers contribute to the h_m-core in addition to the h-core. Authors with a rather flat frequency function $c(r)$ of citations are favored by the modified index, which in my opinion is appropriate.

Of course, the precision problem increases for the determination of the modified index in comparison with the calculation of the original Hirsch index because more papers have to be taken into account. This might be considered a disadvantage in contrast to the simple normalization of the h-index with the mean number of authors. However, that normalization appears to be unreasonably strong because it effectively means that not only the papers are fractionally counted but also that the citations are fractionalized. The strange effect which a single paper with a large number of authors can have on the normalized index h_I, as observed for Datasets G and X, also can be elicited in the following reversed way (Schreiber, 2008c): If a publication with many authors enters the h-core because its number of citations has increased, then this can lead to a decrease of the h_I-index—certainly a peculiar behavior for an index that is supposed to measure the impact of the publications in terms of the numbers of citations. For the modified index, such a problem does not occur.

The fractional counting of papers is not the only way to allocate the credit to several authors of a manuscript. Another straightforward fractional crediting system is to divide the number of citations by the number of authors for each paper (Egghe, 2008). In this way, the citations are shared between the coauthors. However, to determine a Hirsch-type index, this counting requires that the citation records are rearranged according to the fractional citation counts. Moreover, it leads to a fundamental problem when datasets are aggregated (e.g., when one determines the combined index of several people such as all scientists in an institute). For example, a publication with two authors from that institute would contribute two times one half to their h_m-indices, if the citation count is large. But it would be fully taken into account twice when the citations are fractionalized, provided that the number of citations is sufficiently large. This is a methodological problem which also occurs for the original Hirsch index h as well as for the normalized index h_I. In contrast, the total weight is preserved for the modified index, as it should be (Egghe et al., 2000).

In summary, there are three reasons, in my opinion, that the modified index h_m is more appropriate than the h_I-index (i.e., the straightforward normalization of the original Hirsch index h by the mean number of coauthors in the h-core). (a) For h_I, not only the fractional paper count but also the fractional citation count is utilized, which yields a double normalization and is thus excessive. For the modified index, only the rank is fractionalized. (b) The total weight of a publication is preserved in the aggregation of citation records for

DOI: 10.1002/asi

different datasets (e.g., when combining the citation records of several scientists in an institute or of several institutes in a university, or of several universities in a country). In contrast, the weight is not preserved for h_{I}. (c) Adding a highly cited publication to a dataset always increases the modified index, but it can decrease the h_{I}-index, if the number of authors is above average. In conclusion, there are at least three methodological advantages of the modified index h_{m} in comparison with the normalized index h_{I}.

More complicated countings also have been proposed (Egghe et al., 2000; Wan et al., 2007) taking into account that the order of the names in the author list might give an indication about their relative contributions to the whole research. Whether this is indeed the case depends on the customs in the field, and also may be different for groups in the same field. Therefore, in my opinion, it is most appropriate to share the credit equally among all authors, as long as one does not have good information about their relative contributions. Ideally, information obtained directly from the authors about their individual contributions should be used for determining the impact (Vinkler, 1993).

From the point of view of a scientist who cites a publication, the number of coauthors of that publication is usually irrelevant. Accordingly, the value of a citation and thus the impact of a paper should be taken into account independently of the number of authors. This is exactly what happens in the calculation of the modified index because as discussed earlier, the total weight of a publication and thus the value of its citation is preserved for the modified index.

In conclusion, I believe that the modified index is the fairest way of appropriately taking multiple authorship into consideration if one attempts to quantify the impact of a scientist's cumulative research contribution (Hirsch, 2005) by a single number. Whether such an assessment is reasonable on the whole remains a matter of controversy, and I close with a word of caution (which has been attributed to A. Einstein): "Not everything that counts can be counted. And not everything that can be counted counts."

References

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

Bornmann, L., & Daniel, H.-D. (2007). What do we know about the h index? Journal of the American Society for Information Science and Technology, 58, 1381–1385.

Bornmann, L., Mutz, R., & 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, 59, 830–837.

Burrell, Q., & Rousseau, R. (1995). Fractional counts for authorship attribution: A numerical study. Journal of the American Society for Information Science, 46, 97–102.

Burrell, Q.L. (2007a). Should the h-index be discounted? In W. Glänzel, A. Schubert, & B. Schlemmer (Eds.), The multidimensional world of Tibor Braun (pp. 65–68). Leuven, Belgium: ISSI.

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

Egghe, L. (2006). An improvement of the h-index: The g-index. ISSI Newsletter, 2, 8–9.

Egghe, L. (2008). 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, 59, 1608–1616.

Egghe, L., & Rousseau, R. (1990). Introduction to informetrics. Quantitative methods in library, documentation and information science. Amsterdam: Elsevier.

Egghe, L., Rousseau, R., & Van Hooydonk, G. (2000). Methods for accrediting publications to authors or countries: Consequences for evaluation studies. Journal of the American Society for Information Science and Technology, 51, 145–157.

Harsanyi, M.A. (1993). Multiple authors, multiple problems—Bibliometrics and the study of scholarly collaboration: A literature review. Library and Information Science Research, 15, 325–354.

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

Jin, B.H., Liang, L.M., Rousseau, R., & Egghe, L. (2007). The R- and AR-indices: Complementing the h-index. Chinese Science Bulletin, 52, 855–863.

Lindsey, D. (1980). Production and citation measures in the sociology of science: The problem of multiple authorship. Social Studies of Science, 10, 145–162.

Lukovits, I., & Vinkler, P. (1995). Correct credit distribution: A model for sharing credit among coauthors. Social Indicators Research, 36, 91–98.

Price, D.J. de Solla. (1981). Multiple authorship. Science, 212, 986.

Schreiber, M. (2007a). A case study of the Hirsch index for 26 non-prominent physicists. Annalen der Physik (Leipzig), 16, 640–652.

Schreiber, M. (2007b). Self-citation corrections for the Hirsch index. Europhysics Letters, 78, 30002.

Schreiber, M. (2008a). To share the fame in a fair way, h_{m} modifies h for multi-authored manuscripts. New Journal of Physics, 10, 040201.

Schreiber, M. (2008b). 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, 59, 1513–1522.

Schreiber, M. (2008c). A modification of the h-index: The h_{m}-index accounts for multi-authored manuscripts. Journal of Informetrics, 2, 211–216.

Trueba, F.J., & Guerrero, H. (2004). A robust formula to credit authors for their publications. Scientometrics, 60, 181–204.

Van Hooydonk, G. (1997). Fractional counting of multiple-authored publications: Consequences for the impact of authors. Journal of the American Society for Information Science, 48, 944–945.

Vinkler, P. (1993). Research contribution, authorship and team cooperativeness. Scientometrics, 26, 213–230.

Vinkler, P. (2007). Eminence of scientists in the light of the h-index and other scientometric indicators. Journal of Information Science, 33, 481–491.

Wan, J., Hua, P., & Rousseau, R. (2007). The pure h-index: Calculating an author's h-index by taking co-authors into account. Collnet Journal of Scientometrics and Information Management, 1, 1–5.