EMBO reports

correspondence

A proposal for calculating weighted citations based on author rank

A recent article in *EMBO reports* by Bornmann & Daniel (2009) commented that "the *h* index […] is already regarded as the counterpart to the [impact factor]". Indeed, the *h* index (Hirsch, 2005) is increasingly being used to evaluate the achievements of individual scientists, and major citation databases—such as the Scientific Citation Index (SCI; Thomson Reuters, New York, NY, USA) and Scopus (Elsevier, Amsterdam, The Netherlands)—already list by default the *h* index and total citations of every published scientist.

However, this use confuses two distinct concepts: the citation number for a paper and that for an author, which differ because in a paper with multiple authors, their contributions are hardly equal and not all of them should take full credit. Nevertheless, routinely every author in a paper claims all citations as his or her own. Although the author rank is evident in the byline of a publication, it is invisible in citation numbers. For example, SCI and Scopus both disregard author rank when computing the total citation number and *h* index for a scientist. Indeed, multiple authorship is considered to be damaging to the credit system and the situation is becoming more severe as the average author number per paper continues to increase (Greene, 2007; Kennedy, 2003).

Ten years ago, *Nature* introduced a policy advising authors to include a statement about their contributions for each paper (Campbell, 1999). This policy, increasingly being adopted by other scientific journals, is doubtlessly useful and necessary. However, this information becomes invisible when citation numbers are concerned and it is completely qualitative; author contributions should be quantified.

To address this, I propose a quantitative scheme to calculate co-author weight coefficients. Consider a paper of five authors with the last being the corresponding author. Weight coefficients *c* for the first and corresponding authors are 1 for both. Contributions of the second, third, and fourth authors are proportional to 4, 3 and 2, respectively, hence coefficients being 4/9, 3/9 and 2/9, respectively, where $9=4+3+2$. Similarly, for the *k*th author in a paper with *n* authors, $c(k,n)=2(n-k+1)/(n+1)(n-2)$, $n\geq 4$, $2\leq k\leq n-1$ (a special case is $c(2,3)=0.7$ based on the extrapolation of $c(2,n)$, $n\geq 4$). By this definition, except the first and corresponding authors, the sum of weights for the remaining authors is 1. Weighted citation numbers, calculated by multiplying regular citations by weight coefficients, remain the same as regular citations for the first and corresponding authors, but decrease linearly for authors with increasing rank.

The *h* index is based on total citation numbers, which disregard author rank. Therefore, we define *w*, the weighted *h* index, based

Table 1 | A comparison between two researchers with the same total citation numbers and *h* indices, but different weighted citations and *w*, the weighted *h* index

	Researcher 1					Researcher 2				
	Paper number	Author number	Author rank	Regular citation	Weighted citation	Paper number	Author number	Author rank	Regular citation	Weighted citation
	1	6	C*	122	122	1	5	3	106	35.3
	2	7	C	86	86	2	4	C	83	83.0
	3	8	C	44	44	3	4	3	69	27.6
	4	6	C	38	38	4	9	7	33	2.8
	5	8	7	26	1.9	5	10	C	28	28.0
	6	8	C	22	22	6	6	4	24	5.1
	7	7	C	19	19	7	7	5	16	2.4
	8	4	C	15	15	8	5	3	16	5.3
	9	3	2	12	8.4	9	8	4	12	2.2
	10	5	C	8	8	10	4	3	5	2.0
Σcitation				392	364.3				392	193.7
h or *w*				$h=9$	$w=8.3$				$h=9$	$w=5.3$
cac†					0.93					0.49

*Corresponding author; †coefficient of author contribution, which is total weighted citation/total regular citation.

on weighed citations. Let the integer part of w be denoted by $[w]$. An author is said to have the index w if $[w]$ of his or her N papers have at least w weighted citations each and remaining $(N–[w])$ papers have less than w weighted citations each. The h index is a natural number, whereas the w index is a real number.

Recently, Sekercioglu proposed that the kth ranked co-author is considered to contribute $1/k$ as much as the first author (Sekercioglu, 2008), highlighting an earlier proposal (Hagen, 2009; Hodge & Greenberg, 1981). In reality, the corresponding author is usually the last author and takes full credit, and therefore should not be considered to contribute $1/k$ of the first author. Furthermore, a critical flaw of this scheme is the hyperbolic author weight distribution—that is, weights initially decay quickly, but then become almost constant—whereas ideally they follow a linear distribution whereby weights are directly proportional to author ranks.

In summary, when the total citation number or the h index is used to evaluate the scientific impact of a scientist, an underlying assumption is that this researcher takes full credit for all his or her papers. However, this assumption is frequently invalid in papers with multiple authors. The quantitative scheme proposed in this Correspondence can be used to calculate weighted citation numbers and weighted h index, which remain the same as regular citations for the first and corresponding authors, but decrease linearly for authors with increasing rank. Table 1 shows an example: the first researcher is the corresponding author on most of his or her papers, whereas the second is mainly a contributing author; however, their total citation numbers and h indices are the same. By contrast, the weighted citation number and the w index of the former are higher than those of the latter, which is closer to common sense.

REFERENCES
Bornmann L, Daniel HD (2009) The state of *h* index research. *EMBO Rep* **10:** 2–6
Campbell P (1999) Policy on papers' contributors. *Nature* **399:** 393
Greene M (2007) The demise of the lone author. *Nature* **450:** 1165
Hagen NT (2009) Credit for coauthors. *Science* **323:** 583
Hirsch JE (2005) An index to quantify an individual's scientific research output. *Proc Natl Acad Sci USA* **102:** 16569–16572
Hodge SE, Greenberg DA (1981) Publication credit. *Science* **213:** 950
Kennedy D (2003) Multiple authors, multiple problems. *Science* **301:** 733
Sekercioglu CH (2008) Quantifying coauthor contributions. *Science* **322:** 371

Chun-Ting Zhang is in the Department of Physics at Tianjin University, China.
E-mail: ctzhang@tju.edu.cn

doi:10.1038/embor.2009.74