ISSN 1019-3316, Herald of the Russian Academy of Sciences, 2012, Vol. 82, No. 5, pp. 403–405. © Pleiades Publishing, Ltd., 2012.
Original Russian Text © O.V. Mikhailov, 2012, published in Vestnik Rossiiskoi Akademii Nauk, 2012, Vol. 82, No. 9, pp. 829–832.

Point of View

Recently, Russian science studies, particularly scientometrics, have increasingly often discussed issues associated with bibliometric indicators of the scientific productivity of Russian researchers. Scientists are concerned with how fully information about the number of the most frequently cited publications is reflected, which sciences occupy the leading positions in the world and in our country, the extent to which our scientific works are included in world databases, and so on. In analyzing citation specifics, the author of the article published below proposes to introduce a new citation index for researchers.

DOI: 10.1134/S1019331612040119

A New Citation Index for Researchers

O. V. Mikhailov*

Over the past several years, the problem of increasing the citation level of published articles has been "pain in the neck" for the overwhelming majority of scientific publications, primarily scientific journals, all over the world. Researchers, including Russian ones, have also become concerned with increasing their citation level. Although personal citation per se, no matter how high it is, does not prove the significance and value of scientific works or even the demand for them on the part of a certain part of the scientific community [1–4], it is nevertheless a useful criterion that characterizes the level of a researcher's scientific activity.

Generally, citation consists of at least two components: a researcher's personal citation, i.e., the total number of references to articles contained in a respective database (Russian Science Citation Index, Scopus, Web of Science, and others) and the citation of scientific publications (mainly periodicals) that published this researcher's works; the latter is characterized by the publication's impact factor I_F (the ratio of the number of references made to all articles in the publication over the two-year period that preceded the survey year to the total number of articles published in it over the same period) of the date of the publication of a specific work by this researcher [5–7]. At present, both domestic and Western scientometrics, as a rule, prefer to use the first of the above components as basic, largely because of the simplicity of calculation: one should count only the total number of references to the researcher's works included in the respective database, and it will at once become clear who is who.

Even a superficial analysis of the specificity of a researcher's citation shows the following important circumstances.

First, one and the same work may be cited in an exclusively authoritative publication, such as, for example, *Nature*; in the "shoptalk" *Journal of Experimental and Theoretical Physics*; and in a little-known edition, for instance, in a *Herald* or *Proceedings* of a research institution. Formally speaking, we have three references here. However, do we have the right to view them as equivalent to one another, and should we account for the scientific authoritativeness of the edition that provides this reference? There is no clear answer to this question thus far. In addition, citation may be due to various reasons, including those that have no connection with science [8]. Hence, the real "value" of each reference to any publication should not be recognized as identical even on purely formal grounds.

Second, in a number of cases, for one reason or another, researchers make references to their own earlier publications, but databases do not distinguish between self-citation and citation by other authors (at any rate, thus far). Meanwhile, obviously, the value of the citation of a work by a researcher who is not its coauthor and has no connections altogether with its authors and the institution where they work is undoubtedly higher than that of citation by its coauthors or, even more so, self-citation. However, no existing citation-associated database accounts for this circumstance, and references made by "third parties" and self-citation are formally of equal value. If so, these databases include self-citation into the calculation of the currently "voguish" h-index, or the Hirsch index (it is always an integer and equals the number of such h articles each of which was cited at least h times). For any author, the Hirsch index may not exceed the total number of articles published by this author; note that the minimal number of references to a researcher's works with an h-index is h^2.

However, there is a subtlety here. For example, a researcher who is the author or a coauthor of a single article referenced 100 times by other authors, has an h-index of 1. The author or a coauthor of 100 articles,

* Oleg Vasil'evich Mikhailov, Dr. Sci. (Chem.), is chief researcher and a professor in the Department of Analytical Chemistry, Certification, and Management of the Kazan National Research Technological University.

each of which was referenced at least once, will enjoy the same index. Note that, if the author referenced one of his or her 100 articles ten times; another one, five times; still another, three times; ten of the articles were referenced by him or her only once; and no references were made to the rest, his or her Hirsch index will be 3, although the total number of references per 100 articles (28), as well as their value, is substantially lower than in the former two cases. If a researcher has one article with nine references to it, two articles with eight references, three articles with seven references, four articles with six references, five articles with five references, six articles with four references, and seven articles with three references, then, with his or her total 140 references, the *h*-index will be 5. Thus, there is no direct correlation between a researcher's *h*-index and the total number of references to his or her works. The *h*-index often provides a distorted assessment of the significance of a researcher's works. The most obvious situation of this type is the short career of a researcher and a small number of works published by him or her, notwithstanding high citation after his or her death. For example, the *h*-index of the outstanding French mathematician *E.* Galois, one of the founders of the modern theory of algebraic equations, who was killed in a duel at the age of 20, is only 2 and will remain the same forever, no matter how many times his works (unfortunately, they are only two) will be cited in the future. If A. Einstein, one of the greatest physicists ever (apropos, in 1929, one of American Indian peoples named him Chief of Great Relativity—a surprising title for a scientist), had suddenly left for the other world in early 1906, his *h*-index would rather have stopped at 5 despite the extreme significance of the five articles on photoelectric effect interpretation that were published in 1905 and later made him a Nobel prizewinner. These are cases of understating the real significance of a scientist's works, but there are also examples of the opposite situation. For instance, the late and unlamented "People's Academician" T.D. Lysenko, who openly forced other scientists to refer to his works, as well as the Marxism–Leninism classics and "leaders of nations" of the 20th century Stalin and Mao Tse-tung have *h*-indexes of many tens, if not hundreds. This does not mean, however, that the contribution of each of them to any scientific field surpasses many times that of an average researcher. Thus, the *h*-index should hardly be viewed as an adequate reflection of a researcher's real citation either.

Third, it is no secret that the value of a publication substantially depends on the authoritativeness of the edition. It is well known that the most scientifically prestigious and demanded (in any case, if we judge from the citation level of the materials published) are journals that publish review articles and, to a somewhat smaller degree, original, extraordinary results. However, citation databases do not account for this, and a reference to the above-mentioned journal *Nature* is valued no higher than a reference to any other journal. No account for this is found in *h*-index values as well.

At the beginning of this article, we mentioned two components of citation. Logically, the real citation index should reflect both components rather than one of them, as is the case at present. Let us consider them in turn.

Let us assume as the initial postulate that the value of different references to one and the same publication varies and depends on the edition in which it saw the light of day, on the type of the reference (in particular, whether it is citation or self-citation), and on the number of the coauthors of the publication cited. The quantitative measure of an edition's authoritativeness is, as is assumed, its impact factor I_F. This approach implies that, with regard to citation, the averaged significance of each *i*-th article published in a journal is $(I_F)_i$; hence, the averaged significance of any of its parts and, consequently, each reference made to any publication from this journal in any edition (including in the journal itself) may be set equal to $(I_F)_i$. Within this concept, the value of a reference does not depend on what is specifically cited: a short message, an original article, or a scientific review.

We may single out at least five types of citations of scientific works (publications):

• a citation by noncoauthors of the paper in question who have never been coauthors of the given researcher and his or her coauthors;

• a citation by noncoauthors of the publication who have never been coauthors of the given researcher but who were coauthors of at least one of the other coauthors;

• a citation by noncoauthors of the publication but who are coauthors of the given researcher in other publications (irrespective of the date when they saw the light of day);

• a citation by a coauthor of the publication except for the researcher for whom the individual rating of scientific activity is being established; and

• a citation by the coauthor for whom the individual rating of scientific activity is being established (self-citation).

No doubt, it is necessary to account for references of all five categories but, at the same time, to particularize their different weights: it is obvious that it is more difficult for any researcher to obtain the first of the above references, while the fifth one is the simplest to get. This may be done by introducing into consideration the so-called "value coefficient" of a reference (ϕ_i). In this case, the value of a reference is determined as the product of the impact factor of a scientific edition $(I_F)_i$ by ϕ_i. The specific values of the latter coefficient are a particular question, but, in our opinion, the most rational in its solution is the following approach: to assume that it is 1 for references of the first category as the most valuable and from 0 to 1 for the rest (in particular, 0.9 for references of the second category, 0.75 for references of the third category, 0.5 for references of the fourth category,

and 0.25 for references of the fifth category). In the first (so to speak, default) approximation, the $(I_F)_i\phi_i$ value for a specific article should be shared equally among its N_i coauthors (although, by mutual agreement between the coauthors, it is possible to divide it according to the real contribution of each of them to the creation of the article). With account for the fact that different citation levels have historically developed in different scientific branches, it is necessary to introduce the correction factor Φ into the above formula. The lowest citation is in mathematics: mathematicians traditionally prefer not to refer either to the works of colleagues or even to their own works. Hence, the Φ value for mathematics may be taken as a reference point equal to 1, and for other sciences, it may be defined as the quotients of dividing the average citation frequency in mathematics by the average citation frequency in the respective scientific branch. Determining specific Φ values for a respective scientific branch is also a particular problem, analogous to that of defining ϕ_i. With account for the above, we may introduce the OC (own citation) parameter that characterizes the citation of a specific researcher by the number of references to his or her publications in scientific editions:

$$\mathrm{OC} = \sum_{i=1}^{n} \Phi(\phi_j I_{F_j})/N_j .$$

In the same way, it is possible to introduce the EC (edition citation) parameter. As was mentioned, with regard to citation, the averaged significance of each article published in a journal is (I_F); consequently, the $(I_F)_j/N_j$ value will occur for the researcher's EC parameter. From the analogy of the OC parameter, it is also necessary to account for the different degrees of citation in different scientific branches, i.e., $\Phi(I_F)_j/N_j$ will occur in the expression for the EC parameter. It is obvious, however, that the value of an article in any journal and the value of a reference to it are far from being the same. In this context, it is necessary to introduce an a coefficient in the formula for edition citation, which would reflect the degree of EC significance compared to OC significance:

$$\mathrm{EC} = \sum_{j=1}^{k} a\Phi(I_{F_j})/N_j .$$

Within our concept, OC and EC are additive values, and the summary citation (SC) index will be the sum of these two parameters. Obviously, the ratio between OC and EC is very considerable, at least no less than two orders of magnitude. Since the OC and EC formulas incorporate the ratios of the impact factor to the number of coauthors and the Φ factor, and they differ only in the former having the ϕ_j factor, which reflects the type of a reference, and the latter having the a coefficient, it is possible to assume in the first approximation that the a coefficient is 100. As a result, to calculate SC, we obtain the expression

$$\mathrm{SC} = \sum_{j=1}^{n} \Phi(\phi_j I_{F_j})/N_j + 100\sum_{j=1}^{k} \Phi(I_{F_j})/N_j .$$

Theoretically, it is possible to insert in the OC, EC, and SC formulas the values of $(I_F)_i$ and $(I_F)_j$ for the respective scientific editions that they have in the year of calculating these parameters and the $(I_F)_i$ and $(I_F)_j$ values for respective scientific editions each of which had in the year of publication and in the year of their citing specific works by the researcher. It is evident that, to calculate the SC index within the first of the two variants, it is necessary to know the values of the respective impact factors only for one year, while the other requires that at least two years should be taken into consideration (in reality, it is sometimes necessary to account for dozens of different years). In addition, the scientific capital of many patriarchs of our science (of those who are at present 70 years of age or older) contains works that were both published and cited before 1970, when the very notion of the impact factor did not exist. There are no data on the citation of scientific journals and other scientific editions prior to 1970; even the Institute for Scientific Information (ISI) has not yet realized the retrospective view on the citation of scientific editions, let alone other national and international systems that monitor citations. Nevertheless, the second of the above variants of calculating the EC seems more objective, although it is much more difficult to implement in practice.

ACKNOWLEDGMENTS

This article was supported by the Russian Foundation for Basic Research, project no. 10-06-00056.

REFERENCES

1. O. V. Mikhailov, "A Scientist's Citation Index: Is It the Key Criterion of the Quality of the Scientist's Scientific Work?," Naukovedenie, No. 1 (2000).
2. O. V. Mikhailov, "On Objective Criteria of Assessing the Quality of Scientific Work," Naukovedenie, No. 3 (2002).
3. O. V. Mikhailov, "Splendors and Miseries of the 'Citation Index,'" Her. Russ. Acad. Sci. **74** (6), 627 (2004).
4. O. V. Mikhailov, "Citation Index of Russian Scientific Journals: Ways and Reserves of Growth," in *Science Studies 2010: A Collection of Scientific Works* (INION, RAN, Tsentr Nauchn.-Inform. Issled. po Nauke, Obrazovaniyu i Tekhnologiyam, Moscow, 2010) [in Russian].
5. http://ru.wikipedia.org/wiki/%D0%98%D0%BC%D0%BF%D0%B0%D0%BA%D1%82-%D1%84%D0%B0%D0%BA%D1%82%D0%BE%D1%80
6. V. A. Markusova, "Citation of Russian Publications in the World's Scientific Literature," Vest. Ross. Akad. Nauk **73** (4) (2003).
7. A. E. Varshavskii, V. V. Ivanov, and V. A. Markusova, "Adequate Assessment of Scientific Output," Her. Russ. Acad. Sci. **81** (4), 358 (2011).
8. O. V. Mikhailov, "Criteria and Parameters of an Objective Assessment of the Quality of Scientific Activity," Her. Russ. Acad. Sci. **81** (4), 415 (2011).