Proper use of citation counting: the impact factor was created as a metric only for journals, not for individual researchers

ALEXANDRU T. BALABAN
Texas A&M University Galveston, 200 Seawolf Parkway, Galveston, TX 775551, USA

Journal impact factors (JIF), proposed by Garfield for helping librarians to select which journals to subscribe to, have been mistakenly employed (according to the JIF values of journals where their papers are published) for assessing scientific merits of authors working in pure science, This faulty use of a metric that was introduced for a different purpose should be stopped according to a recent declaration signed by many scientists, journal editors, and publishers (DoRA – The San Francisco Declaration on Research Assessment). For assessing and ranking scientists or academic institutions, there are other metrics, such as the Hirsch index *h* which is discussed in the present article, and is favorably mentioned in DoRA. The main ideas of the 18 recommendations of DoRA are briefly discussed.

Keywords: Journal impact factors, Scholarly journals, Hirsch index, Individual researchers

1. Introduction: Science, Journal Impact Factor (JIF), and Science of Science

According to *Webster's New Collegiate Dictionary*, the definition of ***science*** is "knowledge attained through study or practice," or "knowledge covering general truths of the operation of general laws, especially as obtained and tested through scientific methods and concerned with the physical world." Science refers to a system of acquiring knowledge. This system uses observation and experimentation to describe and explain natural phenomena. The term science also refers to the organized body of knowledge people have gained using that system. Less formally, the word science often describes any systematic field of study or the knowledge gained from it.

Pure science is the study of the physical and natural world and phenomena (especially by using systematic observation and experiment), whereas applied science consists in using results of research to fulfill human needs. Just as the "proof of the pudding is in the eating", results of applied science prove the correctness of the pure science that led to the application. Fields of science are commonly classified along two major lines: (i) Natural sciences, the study of the natural world, and (ii) Social sciences, the systematic study of human behavior and society.

It is astonishing how the pseudoscience of astrology and horoscopes can still be present in today's newspapers and magazines! An explanation may be that the zodiac, like the weather, offers ways of opening conversations between strangers.

About half a century ago, Derek De Solla Price in Cambridge, England, and Vassili Vassilievich Nalimov in Moscow, USSR, pondered about how to evaluate pure science, coining terms such as "science of science" or "scientometrics", respectively. However, it was only after Eugene Garfield in USA founded in 1960 the Institute for Scientific Information in Philadelphia (ISI, now part of Thomson-Reuters) and launched the *Science Citation Index* that this field could be established on a firm basis. After Tibor Braun in Budapest, Hungary, launched the journal *Scientometrics*, many scientists in natural and social sciences became aware of this field.

Scientific journals, where results of pure science are published after peer review, could now be divided into about 14,000 "mainstream journals" taken into account by Garfield's *Science Citation Index* and *Current Contents* on one hand, and the numerous remaining journals on the other hand. For helping libraries to select what journals to subscribe to, Garfield proposed an index called journal impact factor (JIF): the per cent ratio between the number of citations in one year and the total number of papers published during the preceding two years [1, 2]. Every year, impact factors for mainstream journals

are published by Thomson-Reuters under the title *Journal of Citation Reports*. Prestigious journals such as *Chemical Reviews*, *Nature*, and *Science* have the largest JIF values, but there is a bias according to the various sciences: biomedical journals have on the average higher JIF values than journals for physics, chemistry or technology, and with mathematics with even lower JIF values.

Unfortunately, a misuse of these JIF indices has started to occur in that JIF values were taken into account when judging the merits of a researcher, a university, or a research institute. Warnings were ignored that one should judge the merits of an author or a scientific paper by the corresponding number of citations, and not by the average citations of the journal. In turn, this malpractice caused a series of inconveniences for individuals, institutions, and scientific journals.

2. Scientometric Indicators for Researchers and Research Institutions. The Hirsch Index

For evaluating the performance of an applied scientist, the criteria are fairly simple – the way the technology works, or the revenue brought by the procedure, device, or patents However, pure science results in publications, among which articles in peer-reviewed journals are the most dynamic ones (books and book chapters are usually published after the author has reached a certain age or position). Citations of publications are the most common criteria which help in assessing the relative importance of papers. The secondary details about self-citations, negative citations, co-authorship, differences between various scientific fields, etc. have been much discussed in the scientometric literature.

In 2005, Jorge E. Hirsch proposed a simple numerical index *h* for measuring individual scientific research outputs: a scientist with *h* publications, each of which has been cited in other publications at least *h* times, has index *h* [3]. With a list of publications ordered according to the decreasing number of citations for each publication, the number *h* is at the intersection between the linearly ascending line of the ordinal numbers for publications starting with 1, and the descending curve for the numbers of citations for each publication. At present one can find Hirsch indices *h* free from Google Scholar (in which case publications include papers published in peer-reviewed journals, books, and book chapters, but there are missing publications) or from SciFinder (Web of Science – Thomson-Reuters), PubMed, Scopus, etc. when publications are limited only to papers published in peer-reviewed journals. Despite various criticisms and attempts for improvements, the index *h* has survived owing to its simplicity and ready availability.

Modeled after the "individual *h* index", one can develop similar indices for institutions, countries, and even journals [4]. The present author has published several papers about JIF, index *h*, and possible improvements of these scientometric indicators [5.6]. Contrary to the established opinion that a paper should be more highly appreciated if it is published in a journal with high JIF (of course, such journals have a wider audience and are more likely to be cited), my argument [6] was that if it was cited *despite its lower visibility*, this meant that it was more valuable than a "run-of-the-mill" paper with high visibility.

3. DoRA – The San Francisco Declaration on Research Assessment

Traditionally, journal publishers recovered publication expenses via the high cost of subscriptions. However, with on-line access, libraries are restricting or canceling at present their subscriptions. Journals published by scientific societies such as the American Chemical Society cost less than those published by private publishing companies such as Elsevier, Taylor and Francis, Wiley, Academic Press, etc. Moreover, electronic journals are much less expensive than journals that have also a printed edition. At present there is an extensive discussion and controversy about the "open journals" containing results financed by grants originating in public funding, when the authors should cover the publication expenses. Most publishers of mainstream journals flag their JIF values in commercial promotions, in order to attract authors and libraries or to overcome the less promising competitors, especially when the JIF is higher than 1.

As a reaction against the misuse of the journal impact factor for assessing research by individuals or academic institutions [7], on May 17, 2013 a group of scientists, scientific journal editors, and publishers disclosed a set of 18 recommendations referred to as the *San Francisco Declaration on Research Assessment* which was published in several leading journals [8]. After enumerating several deficiencies of the journal impact factor, the recommendations are explicitly directed to funding agencies and academic institutions, and are meant to refer to peer-reviewed

research articles but may be extended to other research outputs such as datasets. In funding, appointment and promotion considerations for individual researchers, the research should be assessed on its own merits rather than on the basis of the journal in which the paper was published.

The 1st *general recommendation* is as follows: "1. Do not use journal-based metrics, such as Journal Impact Factors, as a surrogate measure of the quality of individual research articles, to assess individual scientist's contributions, or in hiring, promotion, or funding decisions."

The next two recommendations (2 and 3) are addressed to *funding agencies*, stressing the fact that, especially for early-stage investigators, the scientific content of the paper is much more important than publication metrics or the identity of the journal in which it was published.

A similar content is addressed to *institutions* in the next two recommendations (4 and 5).

For *publishers*, the 6th and the next three recommendations apply: "6. Greatly reduce emphasis on the journal impact factor as a promotional tool, ideally by ceasing to promote the impact factor, or by presenting the metric in the context of a variety of journal-based metrics (e.g., 5-year impact factor, EigenFactor, SCImago, *h*-index, editorial and publication times, etc.) that provide a richer view of journal performance." Then recommendations 7, 8, 9, and 10 are for encouraging a shift toward assessment based on the scientific content of articles rather than publication metrics of the journal in which it was published; calling for so-called "open references". allowing readers to have access to the bibliography of a paper without purchasing the journal or the paper; removing limitations to the number of references; mandating citations of primary literature, rather than of reviews, in order to give credit to the group(s) who first reported a finding.

For *organizations that supply metrics*, recommendations 11, 12, 13, and 14 call for transparency by providing data and methods to calculate all metrics; by allowing unrestricted reuse of data, where possible; and by accounting for difference in article types (reviews versus research articles) and different subject areas when comparing metrics.

Finally, for *researchers*, recommendations 15 to 18 apply: make assessments based on scientific content rather than publication metrics when decisions are involved about funding, hiring, awarding tenure or promotion; in order to give credit where credit is due, cite literature in which observations are first reported rather than reviews.

4. How the Thomson Reuters Company, Which Produces Journal Impact Factors, Reacted to the San Francisco Declaration on Research Assessment

In a statement published shortly after the publication of the *San Francisco Declaration on Research Assessment,* the Thomson Reuters Company provided information on how the Journal Impact Factor (JIF) is calculated, for what purpose it was devised, what are the uses to which the information provided by JIF should be applied, and what should not be done with JIF [9]. One must note that according to this statement the JIF does not measure the quality of individual articles but it does correlate to the reputation of the journal in its field. This reputation is based by the fact that scientists and scholars contribute their time and effort as authors, reviewers, and editors to ensure the quality of journals, and as a result, a comparison among journals may be made based on cited references irrespective of the volume of published items and of the research field.

The important conclusion is that Thomson Reuters, which produces the comprehensive citation index (*Web of Science*) based on approximately 14,000 "mainstream" scholarly journals and the *Journal Citation Report* that contains JIF among other essential information, is in complete agreement with DoRA in that one should use JIF for the purpose it was created, and not for assessing research performance for an individual, a department, or an institution. In the preceding section more details have been presented.

In some European countries during the last two decades, several mistaken tendencies in judging the performance of individual researchers for promotion, tenure, or grant awards have become manifest:

(i) It is an error to judge performance by the number of publications. To counter this tendency, citation-based methods have become fashionable, but the wrong indicator (JIF) was used, instead of citations for the authors' individual articles.

(ii) Counteracting the proliferation of obscure journals that accept hastily-written articles without proper peer review (containing "me-too science" and often written in poor English) had the well-intentioned result of taking into account the quality of journal. However, this idea misfired when lower thresholds for JIF values were decreed, as indicated in an Editorial written by the Editor-in-Chief, Bruce Alberts, in the issue of *Science* that published DoRA [7].

(iii) In some countries like Roumania, during the last two decades, in order to reduce the low-quality research papers published in ephemeral journals such as Annals of obscure "universities" a threshold of JIF > 0.5 was decreed for some types of grant awards. Since JIF varies year after year, the result was that sometimes even the best Roumanian Academy's journals such as *Revue Roumaine de Chimie* or some society journals like *Revista de Chimie Bucuresti* (both having JIF around 0.5 due to good editorial policies) do not attain this threshold and thus in a vicious circle with positive feedback, have become depleted of contributions that would bring citations. This problem is quite complicated by political factors (brain drain, low quota for research budget, loss of industries leading to unemployment, etc.).

(iv) Three years ago, the other journal (*Nature*) that competes with *Science* for the top place among all-encompassing scholarly journals, published an "opinion" selection of six short papers on "how to improve the use of metrics". Of relevance to the present discussion is the first paper of the six, by T. Braun [10], which criticizes the misuse of numbers in scientometrics by persons that do not know the basics of this new sience.

5. Conclusions

Because the journal impact factor (JIF, initially designed for evaluating scholarly journals on the basis of citations during one year of all papers published in a journal during the preceding two years) was misused by including it in the evaluation of individuals or institutions, a group of scientists, journal editors and publishers has recently issued the *San Francisco Declaration on Research Assessment* with recommendations for using JIF appropriately. In a statement published shortly thereafter, the producer of JIF, Thomson Reuters Company, underscored the fact that JIF is not to be used in assessing individual performance in research, in agreement with DoRA. The metric for assessing the performance of individual researchers should be based on citations of papers published by these individuals, irrespectively of the JIF for the journals in which they appeared.

References

[1] E. Garfield, *Citation analysis as a tool in journal evaluation*, Science, **178**, 471(1972).

[2] T. Braun (Ed.) *The Impact Factor of Scientific and Scholarly Journals: Its Use and Misuse in Research Evaluation*, Budapest, Akadémiai Kiado, Scientometric Guidebook Series, **2**, 2007.

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

[4] T. Braun, W. Glänzel, A. Schubert, *A Hirsch-type index for journals*, Scientometrics 2006, **69**,169, Reprinted in T. Braun (Ed.) *The Hirsch-index for Evaluating Science and Scientists. Its Uses and Misuses*. Budapest, Akadémiai Kiado, Scientometric Guidebook Series, vol. 3, 2006.

[5] A. T. Balaban, *How should citations to articles in high- and low-impact journals be evaluated, or what is a citation worth ?* Scientometrics, 1996, **37**, 495 (1996).

[6] A. T. Balaban, *Positive and negative aspects of citation indices and journal impact factors,* Scientometrics **92**, 241(2012).

[7] B. Alberts, Impact factor distortion, *Science*, **340**, 787 (2013).

[8] M. Way, S. A. Ahmad, *The San Francisco Declaration on Research Assessment*, Journal of Cell Science, **126**, 1(2013). Link: www.ascb.org/SFdeclaration.html

[9] *Thomson Reuters statement regarding the San Francisco Declaration on Research Assessment*. Thomson Reuters Research Analytics (2013) May.

[10] T. Braun, *Get experts on board.* Nature, **465**, 870 (2010).

Corresponding author: balabana@tamug.edu