

W-Index: A Weighted Index for Evaluating Research Impact

Xindong Wu^{1,2}

¹Hefei University of Technology, Hefei, China

²Mininglamp Academy of Sciences, Beijing, China

Email: xwu@hfut.edu.cn

How to cite this paper: Wu, X. (2021) W-Index: A Weighted Index for Evaluating Research Impact. *Open Journal of Applied Sciences*, 11, 149-156.
<https://doi.org/10.4236/ojapps.2021.111010>

Received: December 14, 2020

Accepted: January 26, 2021

Published: January 29, 2021

Copyright © 2021 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Academic evaluations such as tenure/promotion applications and society fellowship nominations rely heavily on bibliometric measures of each candidate's research impact, including their research citations. This article first reviews existing evaluation criteria such as the h-index and *q*-most-citations, and then proposes a weighted w-index which minimizes shortcomings in existing single-number measures. The w-index consists of three factors—3 most cited first-author publications, 3 most cited publications as the corresponding/last author, and 3 additional most cited publications as a co-author, but does not allow double counting of these publications.

Keywords

Research Impact, Bibliometric Measures, Citations, Publications, Author Rank

1. Popular Citation Indices

Research seeks to advance knowledge on a particular problem with evidence. Research publications aim to explain researchers' ideas or study the state-of-the-art on a particular research topic, and make them accessible to others. With knowledge advancement in mind, a textbook or newspaper report does not count as a research publication, while a research monograph or a survey paper counts. An innovative idea or a state-of-the-art review can generally be evaluated by the number of citations that the research publication (or publication for short) has attracted. A researcher's research quality can generally be evaluated by three dimensions: productivity, competitiveness, and impact. Productivity refers to numbers of research publications, competitiveness is concerned about the publication forums such as first-rate journals and top-ranked conferences, and im-

impact has much to do with citations. How to evaluate a publication forum's competitiveness has been a subjective and controversial topic, and we will avoid getting into this topic in this article.

Below are some popular single-number measures that have been widely used.

- $N_{p,n}$: the total number of publications published over n years.

$N_{p,n}$ is a simple criterion to encourage productivity, but does not measure impact or competitiveness. Having a good $N_{p,n}$ is a necessary condition at many institutions for promotion/tenure applications, and different institutions have different $N_{p,n}$ thresholds.

- $N_{c,total}$: the total number of citations.

This is a simple indication of total impact, which might not really represent the individual's own impact if the person is just a co-author with many co-authors on several highly cited publications.

- $N_{>c}$: the number of publications that have each been cited for at least C times.

$N_{>c}$ favors seniority and large research groups. A common C value is 100—only if a publication has been cited at least 100 times, can one claim that the publication has attracted visible attention in the research community.

- h -index [1] [2]: a researcher's citation index is h if the person has at least h publications that each has been cited at least h times.

Many variants of the h index have been proposed in recent years, including \bar{h} ("hbar") [3]: an individual's \bar{h} ("hbar") is the number of papers of an individual that each has a citation count larger than or equal to the \bar{h} of all coauthors of each paper.

The h -index and its variants all favor productivity and collaborative citations.

- N_{q-most} : the sum average of the number of citations to each of the q most cited publications (e.g., $q = 3$).

Like the h -index, N_{q-most} does not distinguish primary, senior and secondary co-authors.

The above criteria totally ignore author rank which is explicitly given on the byline of every publication. This ignorance can be dangerous to the academic credit system, as all co-authors can hardly contribute equally to a publication, and the average author number per paper has kept increasing [4] along with the emerging importance of research collaborations. These single-number measures also cause serious confusion when comparing individuals with different backgrounds. For example, a lab managing director or a dean of a small college may publish together with many researchers even in different disciplines.

- Fractional credit [5]: each co-author is given the same credit, hence $1/N$ for a publication if the number of co-authors is N .

This fractional credit has the same issues as the h -index and ignores author rank, but discourages co-authorship.

- $N_{FirstA,c}$: the number of first-authored publications that have each been cited for at least C times.

$N_{FirstA,c}$ does not encourage co-authors but emphasizes the importance of primary authorship in research influence. Suppose someone has a high h -index but

a low $N_{FirstA,c}$ this person is probably a good advisor at a good college, but not a top researcher.

- $1/k$ credit [6], also referred to as the harmonic authorship credit [5]. The k th ranked co-author is given $1/k$ credit of the first author and all co-authors' contributions are normalized to one.

This $1/k$ credit is similar to the weighted citations in [7], the Ab-index [8], and the fair ranking [9], which give the same credit to the first author and the corresponding author, or provide a bonus to the corresponding author.

The $1/k$ credit calculus can be adjusted if the co-author contributions are clearly declared in the publication, such as some co-authors have contributed equally, or all co-authors have taken an alphabetical order on co-author names.

In addition to the above simple criteria, there have also been other efforts such as:

- complicated approaches such as the author matrix [10] which would involve an individual assessment process for each publication,
- efforts on identifying article types (such as methodology studies) that would attract higher citations [11],
- algorithms to capture co-cited publications [12],
- quantification methods of authors' contributions and eligibility for authorship [13],
- rankings of co-authors in research groups [14], and
- dynamic allocation [15] that incorporates additional mechanisms and functions.

2. Common Misunderstandings on Popular Citation Indices

Reasonable people can disagree with others on any single criterion, as there is no single criterion that satisfies everyone's situation. Below are some statements that have appeared in various academic discussions when the author of this article was involved as an academic leader in both the US and China.

2.1. Putting Students as First Authors as the Norm

When a PhD student co-authors a publication with their advisor, it is common practice that the student drives a majority of the work, and the advisor provides advice and all necessary support, hence the student generally deserves being the first author. The PhD student is supposed to be creative and proactive in research activities. Also, if technical contributions are approximately equal, the student would have done most of the implementation work, therefore also deserves the first authorship, in comparison with the advisor or other more senior co-authors. However, there are instances where the student should not be the first author in a joint publication with the advisor. For example, if the central idea has come from the advisor, then it is not an honest approach for the student to take over the first-author credit. Alternatively, if a publication or project involves contributions from multiple students with the overall design from the ad-

visor, it is also more appropriate for the advisor to serve as the first author.

First-authored publications can enhance a student's job marketability, however advisors should not stop producing first-authored publications and count the student's joint publications as the advisor's full credit for this reason. If the advisor has a PhD degree, the citations of the advisor's PhD work should be examined in the same way.

2.2. Alphabetical Ordering on Author Names as Common Practice

In some disciplines such as mathematics and biology, many publications have their co-authors listed in alphabetical ordering, though the author of this article has found non-alphabetical ordering of the coauthors in every discipline. Alphabetical ordering generally indicates that

- Each co-author has made approximately equal contributions in one way or another, hence non-first authors cannot claim more credit than others.
- None of the non-first authors can take full credit.

2.3. Last Co-Author as the Senior and Most Important Co-Author

This is problematic. Last co-authors are most senior individuals in some disciplines or last-added co-authors in other disciplines, but might have made the least technical contributions across disciplines. They are often lab managing directors and/or grant holders in the first case, or have provided last-minute assistance in the final preparation stage of a publication.

2.4. Corresponding Author as the Most Important Author

This is also problematic. Historically, students wrote papers and then changed their affiliations because of graduation. As there did not exist email facilities, the students did not have any permanent corresponding addresses, and their advisors had to act as the corresponding authors. This indicates that a historical corresponding author simply meant a person who could collect offprints and possible reader feedback after paper publications. There should be no more credit to take for a corresponding author than the first author.

With today's email facilities and the World Wide Web, many journals such as several IEEE Transactions no longer indicate corresponding authors, hence there is little importance of a corresponding author.

2.5. Co-Authors for Promoting Collaborations

In 1958, McConnell [16] argued against more than 3 co-authors for each non-monographic treatment. While different research fields might need different levels of collaborations, Green [4] stated that multiple co-authorship endangers the author credit system. There are special cases to make, but we should not include everyone in a publication's authorship who belongs to the same lab or has participated in a group meeting.

It is clearly inappropriate for each co-author to claim the whole credit of a

co-authored publication, whether the co-authors come from the same affiliation or different affiliations.

There are “productive” co-authors who are proactive in asking about others’ ongoing research, making all possible suggestions, hence becoming a co-author of so-many publications, such as an 8th co-author on a 9-author publication for so many publications that come from different affiliations. Such productive co-authors are generally senior in their profession, exist in all disciplines, and always have a high h-index. Whether their co-authors have taken advantage of their senior status in paper publications is still an open question, and their h-index requires a further investigation.

3. W-Index: A Weighted Index for Impact Evaluation

Based on the single-number criteria analyzed in Section 1, the w-index consists of the following factors:

1) W_f : 3 most cited first-author publications, each with the number-of-citations * 100% points;

2) W_l : 3 most cited publications as the corresponding author, each with the number-of-citations * 50% points; (If a publication does not have an explicitly indicated corresponding author, we can use the last author as the corresponding last author.)

3) W_r : any 3 other most cited publications, each with $1/k * 100\%$ points where the candidate is the k th co-author of this publication. These 3 other publications can be selected to maximize the total number of points, hence first-authored publications take priority.

- No double counting, meaning that if a publication has been counted in W_f (with 100% credit) it will not be counted again in W_l or W_r .

What happens if someone does not have 3 first-author publications or 3 last-author publications? Such an individual’s research impact should be evaluated on a case by case basis. If the individual’s publications have always had an alphabetical ordering on co-author names, then each co-author should get the same points for each publication, and the weights on W_f , W_l and W_r should be adjusted accordingly.

Take the author of this article as an example. As of December 12, 2020, his 10 most cited publications at Google Scholar

(<https://scholar.google.com/citations?user=X8sHmqIAAAAJ&hl=en&oi=ao&pagesize=10>) are listed below.

- 1) 5280 (number of citations), 1 (rank of author), 14 (total number of authors)
- 2) 2873, 1, 4
- 3) 1219, 3, 4
- 4) 950, 4, 4
- 5) 870, 1, 2
- 6) 765, 2, 2
- 7) 747, 4, 4
- 8) 673, 2, 2
- 9) 580, 3, 4

10) 571, 1, 3

His w-index is as follows:

- Step 1: from 3 most cited first-author publications, each with number-of-citations * 100% points (Publications 1, 2, 5), $W_f = (5280 + 2873 + 830) \times 100\% = 8933$ points.
- Step 2: from 3 most cited publications as the last co-author excluding publications in Step 1 (Publications 4, 6, 7), each with number-of-citations * 50% points: $W_1 = (950 + 765 + 747) \times 50\% = 1231$ points.
- Step 3: 3 other publications from top-10-most-cited publications excluding publications in Steps 1 and 2 (Publications 3, 8, 10), each with $1/k * 100\%$ points: $W_r = (1219/3 + 673/2 + 571/1) \times 100\% = 1314$ points. Note that Publication 10 is chosen here over Publication 9 as Publication 10 gets a larger number of points.
- Step 4: Xindong Wu's total points in the w-index: $W = W_f + W_1 + W_r = 11,478$.

The author of this article has collected citation numbers several times over several years from Google Scholar, for 12 well-published researchers from the US, China, Britain, Canada, and Australia, to analyze their professional standing (such as fellowships with international societies) and national/international recognition (such as memberships in national academies) with regard to their citation indices. The observation is that the w-index provides a better tool to rank researchers' research impact. It can more accurately predict academy and society inductions than any single-number measures.

4. Concluding Remarks

We have not discussed self-citations in this article for two reasons. First, counting self-citations is a rather tedious process, and most indexing agencies such as the Web of Science and Google Scholar do not provide such a mechanism for their own reasons. Second, the author of this article checked the citation data mentioned in Section 3, and self-citations do not really play a significant role. If a researcher's self-citations have reached a threshold, the person might have ethical concerns for investigation, like journal self-citations handled by the annual Thomas JCR Report.

If the w-index is difficult to calculate in some institutions, the author of this article recommends that the three most cited first-authored publications be used as the primary measure and the h-index as a secondary criterion.

Citations of representative publications or most cited publications play a more important role than the total number of citations. The w-index shares this principle with the h-index and the "highly-cited researchers" by Clarivate Analytics.

Citations should not be the only criterion for evaluating research impact in every situation. But before finding a better criterion for a specific situation, citations, especially those on representative publications are generally a good starting point. The w-index presented in this article was designed to serve this purpose.

Acknowledgements

This work was supported by the National Key Research and Development Program of China, under grant 2016YFB1000901, the National Natural Science Foundation of China under grant 91746209, and the Program for Changjiang Scholars and Innovative Research Team in University (PCSIRT) of the Ministry of Education, China, under grant IRT17R32.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- [1] Hirsch, J.E. (2005) An Index to Quantify an Individual's Scientific Research Output. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 16569-16572. <https://doi.org/10.1073/pnas.0507655102>
- [2] Hirsch, J.E. (2007) Does the *h*-Index Have Predictive Power? *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 19193-19198. <https://doi.org/10.1073/pnas.0707962104>
- [3] Hirsch, J.E. (2010) An Index to Quantify an Individual's Scientific Research Output That Takes into Account the Effect of Multiple Coauthorship. *Scientometrics*, **85**, 741-754. <https://doi.org/10.1007/s11192-010-0193-9>
- [4] Green, M. (2007) The Demise of the Lone Author. *Nature*, **450**, 1165. <https://doi.org/10.1038/4501165a>
- [5] Hagen, N.T. (2008) Harmonic Allocation of Authorship Credit: Source-Level Correction of Bibliometric Bias Assures Accurate Publication and Citation Analysis. *PLoS ONE*, **3**, e4021. <https://doi.org/10.1371/journal.pone.0004021>
- [6] Sekercioglu, C.H. (2008) Quantifying Coauthor Contributions. *Science*, **322**, 371. <https://doi.org/10.1126/science.322.5900.371a>
- [7] Zhang, C.T. (2009) A Proposal for Calculating Weighted Citations Based on Author Rank. *EMBO Reports*, **10**, 416-417. <https://doi.org/10.1038/embor.2009.74>
- [8] Biswal, A.K. (2013) An Absolute Index (*Ab-Index*) to Measure a Researcher's Useful Contributions and Productivity. *PLoS ONE*, **8**, e84334. <https://doi.org/10.1371/journal.pone.0084334>
- [9] Vavryčuk, V. (2018) Fair Ranking of Researchers and Research Teams. *PLoS ONE*, **13**, e0195509. <https://doi.org/10.1371/journal.pone.0195509>
- [10] Chien, T.W., Wang, H.Y., Kan, W.C. and Su, S.B. (2019) Whether Article Types of a Scholarly Journal Are Different in Cited Metrics Using Cluster Analysis of MeSH Terms to Display: A Bibliometric Analysis. *Medicine*, **98**, e17631. <https://doi.org/10.1097/MD.0000000000017631>
- [11] Clement, T.P. (2014) Authorship Matrix: A Rational Approach to Quantify Individual Contributions and Responsibilities in Multi-Author Scientific Articles. *Science and Engineering Ethics*, **20**, 345-361. <https://doi.org/10.1007/s11948-013-9454-3>
- [12] Shen, H.W. and Barabási, A.-L. (2014) Collective Credit Allocation in Science. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, 12325-12330. <https://doi.org/10.1073/pnas.1401992111>

- [13] Ivaniš, A., Hren, D., Sambunjak, D., Marušić, M. and Marušić, A. (2008) Quantification of Authors' Contributions and Eligibility for Authorship: Randomized Study in a General Medical Journal. *Journal of General Internal Medicine*, **23**, 1303-1310. <https://doi.org/10.1007/s11606-008-0599-8>
- [14] Ausloos, M. (2013) A Scientometrics Law about Co-Authors and Their Ranking: The Co-Author Core. *Scientometrics*, **95**, 895-909. <https://doi.org/10.1007/s11192-012-0936-x>
- [15] Bao, P. and Zhai, C.X. (2017) Dynamic Credit Allocation in Scientific Literature. *Scientometrics*, **112**, 595-606. <https://doi.org/10.1007/s11192-017-2335-9>
- [16] McConnell, D. (1958) Quantifying Coauthor Contributions. *Science*, **128**, 1157.