Credibility on Scholar Performance Evaluation Using Google Scholar and ResearchGate

YOSHIYASU TAKEFUJI

Abstract: Scholar performance evaluation plays a key role in management science and engineering. Scholar evaluation using Google Scholar and ResearchGate can serve as an indispensable scouter for evaluating scholar performance. Both tools to quantitatively evaluate scholars can be used to support evidence-based decision making in administration and human resources. However, both tools must be used together for complementing accurate scholar evaluation. This author shows examples of fatal drawbacks in Google Scholar and ResearchGate, respectively. Scopus and Publons, used as default scholar performance, are affected by publisher-bias selection of journals and conferences. The author recommends scholar performance evaluation using both tools such as Google Scholar and ResearchGate together with Scopus and Publons.

Keywords: scouter, SNS, impact factor, author impact metrics, scholar evaluation

INTRODUCTION

Measuring the performance of scholars and researchers is closely tied to issues of educational ethics. The fairness approach assumes that people should not be discriminated against and should be treated equally regardless of their status. This article identifies credibility in the current evaluation metrics and proposes a possible solution.

According to Wikipedia and academic study,[1] Dragon Ball has become one of the most successful manga and anime series in the world. Dragon Ball was sold in over forty countries and the anime broadcast in more than eighty countries. The forty-two paperback volumes are estimated to have sold more than 160 million copies in Japan and 250 to 300 million copies worldwide. Dragon Ball became the second-best-selling manga series, behind only One Piece.

YOSHIYASU TAKEFUJI – Faculty of Data Science, Musashino University, Tokyo, Japan, https://orcid.org/0000-0002-1826-742X

 Journal of Scholarly Publishing 2023
doi: 10.3138/jsp-2022-0076

Dragon Ball is a Japanese media franchise created by Akira Toriyama in 1984. The ultimate science-fiction martial arts manga or comics and one of the best-selling series of all time, sold in more than forty countries in the world. Since 1998, this animated television series has been distributed worldwide through Cartoon Network and has become a hit. Dragon Ball is the story of a young warrior named Son Goku. He is a strange boy with a tail who begins an adventure to become stronger and learns of the existence of the Dragon Balls, which, if he collects all seven, will grant him any wish.

In Dragon Ball, an introduced scouter is a wearable, all-purpose computer that Frieza's army uses.[2] Scouters are mainly used to measure the power levels of the animated characters in Dragon Ball.

Inspired by the concept of the scouter in Dragon Ball, this article proposes a scouter that uses two tools together to measure scholar performance: Google Scholar scores and ResearchGate (RG) scores.[3] According to Graduate Institute Geneva, author impact metrics are indicators that can be used to assess the impact of a scholarly article on a single author, a group of multiple authors such as a laboratory, or a collection of groups such as an institution.[4] They are based on the number of citations and the number of publications. The most recognized author metrics is the *h*-index. The *g*-index and the i10-index are alternatives to the *h*-index.[5]

There are two types of scholar performance evaluations: publisher's metrics, such as Scopus and Publons, and social networking service (SNS) metrics, such as Google Scholar and RG. This article shows the drawbacks of individual metrics using actual examples.

In Scopus, owned by Elsevier, the metrics overview includes the number of documents by author, citations, and *h*-index.[6] Scopus is the largest abstract and citation database of peer-reviewed literature: scientific journals, books, and conferences. However, the problem of Scopus lies in that literature is selected only by Scopus or Elsevier. The Web of Science (WOS), owned by Clarivate, is a website that provides subscription-based access to multiple databases that provide comprehensive citation data.[7] WOS is now using Publons.[8] Publons is a free resource for the global, multi-disciplinary scholarly research community. However, Clarivate is owned by private-equity firms Onex Corporation and Baring Private Equity Asia. Owners or profit organizations may tend to affect their business and influence author metrics for their merits, especially biased selection of journals and conferences. Clarivate provides the impact factors of

 doi: 10.3138/jsp-2022-0076

their selected journals. Journal Citation Reports™ provides journal intelligence that highlights the value and contribution of a journal through a rich array of transparent data and metrics, including the Journal Impact Factor™ (JIF).[9]

Martín-Martín et al. investigated and compared coverage via citations. They concluded that WOS is the smallest and Google Scholar is still the most comprehensive source. Google Scholar is the best choice in almost all subject areas for those needing the most comprehensive citation counts but not needing complete lists of citing sources.[10]

Singh et al. studied articles and citations and concluded that Google Scholar can be a useful tool for locating open-access full-text versions of close to about half of the scientific articles of the world, which has special significance for under-developed and developing countries.[11] This means that researchers in under-developed or developing countries cannot conduct scientific research at commercial publisher sites such as Scopus and WOS.

The following shows an example of discrepancies between Scopus and Publons.

For example, the evaluation of the same researcher (the author of this manuscript) is completely different as follows:

Scopus: publications: 203, citations: 2608, *h*-index: 25
Publons: publications:171, cited: 1736, *h*-index: 18
Google Scholar: citations: 6402, *h*-index: 33, i10-index: 94
RG: publications: 680, citations: 4269, RG Score: 50.47; *h*-index: 30

This *h*-index differences were created simply by selecting journals and conferences. This is called publisher's bias due to journal and conference selection. In Scopus and Publons, users are not allowed to add journals and conferences.

Since Google Scholar and RG are not publishers, they are able to select publications without the influence of publishers or publisher owners. However, Google Scholar and RG have credibility problems with correctable drawbacks. This article investigates what drawbacks on scholar performance are in their systems.

Google Scholar was launched on November 20, 2004. Google Scholar is a freely accessible web search engine over the Internet that indexes the

 doi: 10.3138/jsp-2022-0076

full text and metadata of scholarly literature in a variety of publication formats and disciplines. Google Scholar can report an *h*-index and i10-index with a scholar's given name. The *h*-index is the largest number *h* such that at least *h* articles in a scholar's publication were cited at least *h* times each. The *h*-index is a single measure of the cumulative impact of an author's scholarly output and accomplishments, measuring quantity with quality by comparing publications and citations. The i10-index indicates the number of academic publications an author has written that have been cited by at least ten sources. Google Scholar uses automated software, known as 'robots' or 'crawlers,' to fetch scholar files to include in the search results.[12]

According to RG, RG was founded in 2008. The RG Score measures scientific reputation based on how a scholar's work is received by a scholar's peers. RG believes that researchers are the best judges of each other's work and that all a scholar's research, published or not, deserves credit. With this in mind, the RG Score is calculated based on any contribution of a scholar's share on RG or added to a scholar's profile, such as published articles, preprints, unpublished research, projects, questions, and answers, respectively. RG had 20 million users as of March 2022. The RG Score is relative; it can go up or down depending on the activity and scores of other RG members.

The open-access policies adopted by Google Scholar and RG are very useful for researchers around the world, but the existing system has several shortcomings. This article discloses some shortcomings of Google Scholar and RG with actual examples.

In this article, the drawbacks of Google Scholar and RG are summarized, respectively. Using both Google Scholar and RG with removing the mentioned drawbacks can serve as a useful and indispensable scouter for measuring scholar performance. However, the current peer-review system should be fixed since we do not have enough research evaluators to be highly valued and trained.[13]

METHODOLOGY

RG has also been criticized for failing to provide safeguards against 'the dark side of academic writing,' including such phenomena as fake publishers, 'ghost journals,' publishers with 'predatory' publication fees, and fake impact ratings.[14] There is no available information on how much the ghost-journals problem has been alleviated by RG. In addition to the

 doi: 10.3138/jsp-2022-0076

predatory problem, RG does not disclose its algorithm for calculating RG Scores. Due to many criticisms of RG, the metric on scholar performance has been changing.

This article briefly depicts two critical drawbacks of services provided by RG. Skills and expertise information of researchers can be obtained in RG Scores. However, highly cited researchers have no association with their skills and expertise. In other words, skills and expertise can be picked by a researcher where they have nothing to do with published articles. The verification of skills and expertise is needed for achieving reliable researcher performance. The first problem lies in that RG Score has no relationship with skills and expertise as a provided service for which skills and expertise should be verified.

The RG Score has a questions-and-answers score. The questions-and-answers scores give unreliable author impact metrics. For example, Peter Breuer has an RG Score of 181.48, which is the highest score in Logic Programming (skills and expertise) in the world with the following: Research items: 250, Total Research Interest: 1199, Citations: 1518, Recommendations: 6881, Reads: 157,684, Answers: 5359. However, for Peter Breuer, Google Scholar shows the following: Citations: 1582, *h*-index: 21, i10-index: 44. Google Scholar indicates that Peter Breuer is not a highly successful researcher. The questions-and-answers score increased his RG Score dramatically. However, Google Scholar does not support Peter Breuer is the best logic programming researcher in the world.

A more extreme case can be found in artificial intelligence of skills and expertise. Joachim Pimiskern has the highest RG Score in the world: 312.53 in artificial intelligence with Research items: 0, Total Research Interest: 0, Citations: 0, Recommendations: 4628, Reads: 77,306, Answers: 3504. However, Joachim Pimiskern has no articles published with the highest RG Score in artificial intelligence in the world. Joachim Pimiskern has no account in Google Scholar because of no articles being published. The second problem of RG Score lies in the questions-and-answers scores. In questions-and-answers scores, there is no expert review or expert peer review.

Although questions-and-answers scores should be removed from the RG Score, the RG score incorporates the JIF to evaluate individual researchers while Google Scholar does not. The higher the impact-factor journal articles, the higher the RG Score obtained. As long as questions-and-answers

 doi: 10.3138/jsp-2022-0076

scores are removed from the RG Scores, it will be a useful scouter for evaluating researcher performance using the RG Score.

Google Scholar automatically crawls and collects authors' publication information on the Internet. In other words, the same citation information can be counted as multiple citations with different domain names. In the current Google Scholar, the number of different domain names citing the same article is equivalent to the number of different citations against the same article. It is recommended that Google Scholar should fix multiple-citation-counting problems with domain names.

In the general sciences, *Science* and *Nature* have been the two most prestigious journals in the world with a high impact factor. Google Scholar metrics use the *h*5-index and *h*5-median, respectively, for calculating journal impact metrics. However, there is no scientific explanation for why the *h*5-index is used. The *h*5-index is the *h*-index for articles published in the last five complete years. It is the largest number *h* such that *h* articles published in between 2015 and 2019 have at least *h* citations each. The *h*5-median is based on *h*5-index but instead measures the median (or middle) value of citations for the *h* number of citations. However, Google Scholar does not provide a new service incorporating the Journal impact factor based on the h5-index or the h5-median for calculating author impact metrics. There is no scientific explanation for why Google Scholar uses the *h*5-index.

Table 1 shows the top ten journals based on the *h*5-index and the *h*5-median. There is no document on why the *h*5-index was chosen in Google Scholar for ranking journals and conferences. In other words, using a different *h*X-index can change the rankings, where X is a positive integer.

DISCUSSION

RG Scores should be corrected in two problems: a no-association problem between 'skills and expertise' and published articles and a questions-and-answers-score problem. The algorithm for calculating RG scores is currently not disclosed in public. RG needs transparency on the RG Scores algorithm. However, the RG Score incorporates the journal impact metrics while Google Scholar only shows the *h*5-index and the *h*5-median without incorporating the journal impact metrics. Using both the RG Score and the Google Scholar score can serve as a useful and indispensable scouter for measuring scholar performance.

 doi: 10.3138/jsp-2022-0076

TABLE 1. Publication with *h*5-Index and *h*5-Median

Publication	h5-index	h5-median
1. *Nature*	444	667
2. *The New England Journal of Medicine*	432	780
3. *Science*	401	614
4. IEEE/CVF Conference on Computer Vision and Pattern Recognition	389	627
5. *The Lancet*	354	635
6. *Advanced Materials*	312	418
7. *Nature Communications*	307	428
8. *Cell*	300	505
9. International Conference on Learning Representations	286	533
10. *Neural Information Processing Systems*	278	436

Petersen et al. showed the discrepancies in scientific authority and media visibility.[15] In other words, reliable scientist assessment tools play a key role in measuring scholar performance. In using reliable scientist assessment tools, the UN's Sustainable Development Goals (SDGs) should be also considered in the future, since there is no algorithm available in assessing scholar performance with embedding SDGs scores.

Li et al. examined the long-term impact of co-authorship with established, highly cited scientists on the careers of junior researchers in four scientific disciplines.[16] Their result suggests that accurate scholar scores can be used for selecting the advisors of their dissertation.

However, three national scientific academies, the French Academy of Sciences, the German Leopoldina, and the UK Royal Society, are issuing a joint statement on how to make sure research evaluation is done well.[17] The statement stated that the current peer-review system is broken so that research evaluators to be highly valued and trained are needed.[18] The current scholar assessment tools cannot easily resolve the problem. Open peer review may alleviate the problem. However, an editorial has addressed the pros and cons of open peer review.[19] The current (open or closed) peer-review problem cannot be easily resolved.

It is crucial for students to select research advisors or to choose majors based on big data technologies.[20] Google Scholar and RG alone cannot be used for accurate scholar evaluation, so it is necessary to use both tools such as Google Scholar and RG to complement each other.

As of 10 September 2022, RG has changed its scoring metric due to a lack of openness in scoring individual researchers. In other words, the openness of scoring researchers is one of the most important issues in scholar quality evaluation. RG and Google Scholar still have many problems, such as relying on researchers to add publications to their own accounts, although some of them are automatically added. The selection of publications and their reliability play an important role in evaluating individual researchers. None of the existing systems for measuring the performance of scholars and researchers fulfils any of the important aspects, such as the selection of publications and metric openness with the selection of journals and conferences.

Koltun et al. gave a strong statement on the *h*-index such that the *h*-index is no longer an effective correlate of scientific reputation.[21] They proposed fractional allocation measures such as h-frac. However, the calculation of scientific reputation is controversial. In other words, we do not know how to scientifically calculate scientific reputation.

Smith strongly criticized peer review, showing the flawed process at the heart of science and journals.[22] However, as far as we know, there is no alternative to peer review.

The open-access policies adopted by Google Scholar and RG are very useful for researchers around the world such as developed, underdeveloped, or developing countries, but the existing system has several shortcomings. This article disclosed some shortcomings of Google Scholar and RG with actual examples. This means that we must cautiously use open-access sites.

Commercial publishers' sites such as Scopus and WOS, as Martín-Martin et al. noted, charge a fee called a paywall and have a limited selection of citations. The open access policies adopted by Google Scholar and RG are useful but currently have some drawbacks. It may be a long time before the world moves to open science and open citation. Therefore, until reliable open-citation sites are available to the public, it is necessary to carefully use open-access sites such as Google Scholar and RG in combination with commercial publisher sites such as Scopus and WOS.

 doi: 10.3138/jsp-2022-0076

CONCLUSION

This article showed the drawbacks of publisher's metrics, such as Scopus and Publons, and SNS metrics, such as Google Scholar and RG. RG needs transparency in calculating the RG Scores. Recently, RG has changed how this score is formulated. In Google Scholar, the automated crawling and collecting authors' information over the Internet should be fixed because of how multiple citations are counted against the same article. However, using both SNS tools, such as Google Scholar and RG, can serve as a useful and indispensable scouter for measuring scholar performance together with Scopus and Publons. We must solve the (open or closed) peer-review problem for reliable scholar scores. A no-bias tool quantitatively evaluating scholars should be used for supporting evidence-based decision making using big data technologies. Until reliable open citation sites are available to the public, it is necessary to carefully use open-access sites, such as Google Scholar and RG, in combination with commercial publisher sites, such as Scopus and WOS.

YOSHIYASU TAKEFUJI graduated from Electrical Engineering, Keio University with a BS (1978), an MS (1980), and a PhD (1983). He has served as an assistant professor at the University of South Florida (1983–1985), an associate professor at the University of South Carolina (1985–1988) and Case Western Reserve University (1988–1996: tenured in 1992), as a tenured professor and emeritus professor at Keio University (1992–2021), as a professor at Musashino University (2021–present), and as a docent professor at Jyvaskyla University. He has authored forty-one books and more than 870 scientific articles and has supervised more than forty-one PhD students. He has published 211 articles in *Science*, one in *Nature*, and one in the *New England Journal of Medicine*. Based on publications, he was selected in 2021 and 2022 by Times-Higher Education as a reputation participant representing peers in both his discipline and country for World University Rankings.

NOTES

1. Wikipedia, s.v., 'Dragon Ball,' accessed April 21, 2023, https://en.wikipedia.org/wiki/Dragon_Ball; José Andreés Santiago Iglesia, 'Dragon Ball Popularity in Spain Compared to Current Delocalized Models of Consumption,' *Mutual Images Journal*, no. 2 (2017): 110–36, https://doi.org/10.32926/2017.2.san.drago.
2. Maxime Derian, 'Design and Distribution of Detachable Digital Prostheses,' in *Cognitive Prosthetics* (London: ISTE Press, 2018), 29–57; Fandom, 'Scouter,' accessed April 19, 2023, https://dragonball.fandom.com/wiki/Scouter.

 doi: 10.3138/jsp-2022-0076

3. Alberto Martín-Martín, Mike Thelwall, Enrique Orduna-Malea, and Emilio Delgado López-Cózar, 'Google Scholar, Microsoft Academic, Scopus, Dimensions, Web of Science, and opnCitations' COCI: A Multidisciplinary Comparison of Coverage via Citations,' *Scientometrics* 126, no. 1 (2021): 871–906, https://doi.org/10.1007/s11192-080-03690-4; Wikipedia, s.v., 'Google Scholar,' accessed February 26, 2023, https://en.wikipedia.org/wiki/Google_Scholar; Josef Schmied, 'Popular Digital Knowledge Dissemination Platforms: Evaluating the Pragmatic Professional Credibility from Wikipedia to Academica.edu and ResearchGate,' *Journal of Pragmatics* 180 (July 2021): 187–202, https://doi.org/10.1016/j.pragma.2021.04.027; Wikipedia, s.v., 'ResearchGate,' accessed February 26, 2023, https://en.wikipedia.org/wiki/ResearchGate.
4. Graduate Institute Geneva, 'What Is the Author Impact?,' accessed February 26, 2023, https://libguides.graduateinstitute.ch/metrics/author_impact.
5. Graduate Institute Geneva, 'What Is the Author Impact?'
6. Elsevier, 'About Scopus,' accessed February 26, 2023, https://www.elsevier.com/solutions/scopus; Yrjo Lappalainen and Nikesh Narayanan, 'Harvesting Publication Data to the Institutional Repository from Scopus, Web of Science, Dimensions and Unpaywall Using a Custom R Script,' *Journal of Academic Librarianship*, published ahead of print 21 December 2022, https://doi.org/10.1016/j.acalib.2022.102653; Elsevier, 'About Elsevier,' accessed February 26, 2023, https://www.elsevier.com/about.
7. Ying Chen, Thorsten Koch, Nazgul Zakiyeva, Kailiang Liu, Zhiton Xu, Chunhouh Chen, Junji Nakano, and Keisuke Honda, 'Article's Scientific Prestige: Measuring the Impact of Individual Articles in the Web of Science,' *Journal of Informetrics* 17, no. 1 (2023): 101379, https://doi.org/10.1016/j.joi.2023.101379; Wikipedia, s.v., 'Web of Science,' accessed February 26, 2023, https://en.wikipedia.org/wiki/Web_of_Science; Clarivate, 'Web of Science,' accessed February 26, 2023, https://clarivate.com/products/scientific-and-academic-research/research-discovery-and-workflow-solutions/webofscience-platform/.
8. Publons, 'Publons, Track More of Your Research Impact,' accessed February 26, 2023, https://publons.com/about/home.
9. Clarivate, 'Journal Citation Reports,' accessed February 26, 2023, https://clarivate.com/products/scientific-and-academic-research/research-analytics-evaluation-and-management-solutions/journal-citation-reports/.
10. Martín-Martín et al., 'Google Scholar, Microsoft Academic, Scopus.'
11. Vivek Kumar Singh, Satya Swarup Srichandan, Rajesh Piryani, Anurag Kanaujia, and Sujit Bhattacharya, 'Google Scholar as a Pointer to Open Full-Text Sources of Research Articles: A Useful Tool for Researchers in Regions with Poor Access

 doi: 10.3138/jsp-2022-0076

to Scientific Literature,' *African Journal of Science, Technology, Innovation and Development*, published ahead of print, 7 October 2022, https://doi.org/10.1080/20421338.2022.2124689.

12. Google, 'Google Scholar, Inclusion Guidelines for Webmasters Crawl Guidelines,' accessed February 26, 2023 https://scholar.google.com/intl/en/scholar/inclusion.html#crawl.
13. Mary Kay Gugerty and Dean Karlan,'Ten Reasons Not to Measure Impact—and What to Do Instead,' *Stanford Social Innovation Review* 16, no. 3 (2018): 41–47, https://doi.org/10.48558/2A2K-0K07; Jeff C. Clements, 'Don't Be a Prig in Peer Review,' *Nature* 585, no. 472 (2020), https://doi.org/10.1038/d41586-020-02512-0; 'Pros and Cons of Open Peer Review,' editorial, *Nature Neuoscience* 2 (1999): 197–98, https://doi.org/10.1038/6295.
14. Aamir Raoof Memon, 'ResearchGate Is No Longer Reliable: Leniency towards Ghost Journals May Decrease Its Impact on the Scientific Community,' *Journal of the Pakistan Medical Association* 66, no. 12 (2016): 1643–47.
15. Alexander Michael Petersen, Emmanuel M. Vincent, and Anthony LeRoy Westerling, 'Discrepancy in Scientific Authority and media Visibility of Climate Change Scientists and Contrarians,' *Nature Communications* 10 (2019): 3502, https://doi.org/10.1038/s41467-019-09959-4.
16. Waihua Li, Tomaso Aste, Fabio Caccioli, and Giacomo Livan, 'Easly Coauthorship with Top Scientists Predicts Success in Academic Careers,' *Nature Communications* 10 (2019): 5170, https://doi.org/10.1038/s41467-019-13130-4.
17. Richard Catlow, 'Peer Reviewers Need More Nurturing,' *Nature* 552, no. 293 (2017), https://doi.org/10.1038/d41586-017-08289-z.
18. Catlow, 'Peer Reviewers.'
19. 'Pros and Cons.'
20. Wenhua Yu, Lei Zhao, Xiangyu He, Jiacheng, Zhou, Tong Cheng, Changzhao Xue, and Fan Yang, 'A New Evaluation System for Scholars and majors Based on Big-Data Techniques,' in *2016 IEEE International Parallel and Distributed Processing Symposium Workshops, Chicago* IL (New York: IEEE, 2016), 1477–80.
21. Vladlen Koltun and David Hafner, 'The h-Index Is No Longer an Effective Correlate of Scientific Reputation,' *PLoS One* 16, no. 6 (2021): e0253397, https://doi.org/10.1371/journal.pone.0253397.
22. Richard Smith, 'Peer Review: A Flawed Process at the Heart of Science and Journals,' *Journal of the Royal Society of Medicine* 99, no. 4 (2006): 178–82, https://doi.org/10.1258/jrms.99.4.178.