

Multiple versions of the *h*-index: cautionary use for formal academic purposes

Jaime A. Teixeira da Silva¹ · Judit Dobránszki²

Received: 9 October 2017 / Published online: 20 February 2018
© Akadémiai Kiadó, Budapest, Hungary 2018

Abstract The *h*-index, or Hirsch index, named after Jorge E. Hirsch, is one of the few author-based metrics currently available that offers a perspective of the productivity and citation impact of a scientist, researcher, or academic. There are four tools most commonly used to calculate the *h*-index, all of which depend on separate databases: Scopus, Web of Knowledge, Google Scholar, and ResearchGate. Using the *h*-index of the authors of this paper derived from these four sources, it is abundantly clear that scores vary widely and that it is unclear which of these sources is a reliable or accurate source of information, for any purpose. As the use and application of author-based metrics increases, including for official academic purposes, it is becoming increasingly important to know which source of the *h*-index is most accurate, and thus valid. Although this is not a review of the *h*-index, some perspectives are provided of the *h*-index-related literature to place this case study within a wider context of the weaknesses and criticisms of using the *h*-index as a metric to evaluate scientific outcome.

Keywords Accuracy · Author-based metrics · Creditability · Databases · Google Scholar · Scopus · Web of Science

Abbreviations

ABM Author-based metric
GS Google Scholar

Jaime A. Teixeira da Silva and Judit Dobránszki are equal contributors and co-corresponding authors.

✉ Jaime A. Teixeira da Silva
jaimetex@yahoo.com

✉ Judit Dobránszki
dobranszki@freemail.hu

¹ Miki-cho, Takamatsu, Japan

² Research Institute of Nyíregyháza, IAREF, University of Debrecen, Nyíregyháza, P.O. Box 12, 4400, Hungary

<i>h</i> -index	Hirsch index
SRA	Scientist, researcher, or academic
WoS	Web of Science

Dear *Scientometrics* Editors,

Some background about the *h*-index

The *h*-index, or Hirsch index is an author-based metric (ABM) for scientific output. As Hirsch (2005) defined: “a scientist has index h if h of his/her N papers have at least h citations each, and the other $(N_p - h)$ papers have $\leq h$ citations each”. One of its greatest benefits is that the *h*-index serves as an indicator that is able to simultaneously measure both the productivity and the impact of a scientist, researcher, or academic (SRA), as it includes the quantity and quality of their publishing activity (Hirsch 2005; Costas and Bordons 2007; Bornmann and Daniel 2009). It is also used to measure the scientific output of research groups and even countries (Bornmann and Daniel 2009). Thus, an SRA with an *h*-index of 100 has written 100 papers that have each accumulated at least 100 citations, therefore only considering highly cited papers. However, as for any academic indicator or metric, the *h*-index has some limitations: (1) the maximum value of the *h*-index is, as is evident from its definition, determined by the maximum number of papers published, but that must meet a minimum citation threshold; (2) there are differences in the *h*-index between disciplines (Glänzel and Persson 2005), so the achievements of SRAs from different disciplines cannot be compared (Hirsch 2005); (3) the *h*-index does not indicate the real citation number of papers (\neq the real number of cited papers; Costas and Bordons 2007); (4) the *h*-index depends on the citation tool used (i.e., bibliographic management tools or citation managers, such as Endnote, Mendeley, RefWorks, and Zotero) and the length of the SRA’s carrier (Hirsch 2005; Costas and Bordons 2007); (5) one of the most important limitations of using and comparing *h*-index lies in database coverage (Bar-Ilan 2008; Adriaansee and Rensleigh 2013; Halevi et al. 2017).

Using a search of three of the most commonly used academic databases, Google’s Google Scholar (GS), Elsevier’s Scopus, Clarivate Analytics’ Web of Science (WoS, previously Web of Knowledge), as well as the academic social networking site (ASNS) ResearchGate (RG), we highlight next some of the most pertinent positive aspects and criticisms of this metric based on actual *h*-index scores for both authors.

Authors’ *h*-indexes, and use of the *h*-index for formal academic purposes

Not too much information exists on the use of the *h*-index for formal academic purposes. However, as article-level metrics have become increasingly recognized over the last few years over journal-based metrics, such as Clarivate Analytics’ (formerly the Intellectual Property and Science business of Thomson Reuters) journal impact factor (JIF; see criticisms in Teixeira da Silva 2017), for measuring the scholarly activity of SRAs, the role of the *h*-index has increased, and is now used for tenure promotion, academic ranking, prizes, salaries or to obtain research grants (Glänzel and Persson 2005; Bornmann and Daniel 2009; Svider et al. 2014; Lippi and Mattiuzzi 2017; Popova et al. 2017; Saraykar et al. 2017).

The first author of this paper was recently (August, 2017) invited to serve as a reviewer for research projects, primarily in the plant sciences. One of the prerequisites was to provide the *h*-index based on Scopus and/or WoS for the last 5 years (2012–2016). Accessing the *h*-index using GS was possible, and using a non-updated account,¹ an *h*-index of 38 was revealed for this time period, or 42 overall. The second author of this paper has a GS-based *h*-index of 14 for this time period, or 15 overall.² The RG-based scores for the first and second authors are 36 (30 excluding self-citations) and 13, respectively (12 excluding self-citations) (Table 1). The GS and RG *h*-indexes are free and open to the public to view, although to view the RG *h*-index, one has to be a registered user and login to see one's own and others' *h*-indexes. In contrast to GS and RG, the Scopus and WoS *h*-indexes lie behind proprietary paywalls: Scopus is an abstract and citation database owned by Elsevier Ltd.³ while WoS is owned by Clarivate Analytics.⁴

For the first author, only 25 publications are listed in the WoS Core Collection for the 1975–2018 time period while for the second author, 49 publications are listed. In reality, there are 972 publications for the first author, as listed accurately on RG⁵ and 100⁶ for the second author, over the same time period. The WoS-based *h*-index for the first and second author is 8 and 9, respectively. Thus, WoS, at least for these two authors, does not represent an accurate portrayal of the real publication status of SRAs and thus the WoS-based *h*-index should not be used.

Scopus lists six individuals (17 name variations among them) for the first author and one individual (four name variations for it) for the second author. Integrating the first author's six "identities", the combined Scopus *h*-index for the first author is 29 (24 excluding self-citations) while that for the second author is 11 (10 excluding self-citations) for the 2002–2017 period and including papers published in 1998–2017 and 1993–2017, by the first and second author, respectively. That indicates that 408 and 66 documents were taken into account in the Scopus citation reports for the first and second author, respectively. The number of documents considered in Scopus, as for WoS, differs significantly from reality, considering only 42 and 66% of the first and second authors' publications, respectively. As for WoS, the Scopus-based *h*-index does not represent an accurate portrayal of the real publication status of these two authors and should not be used by SRAs until at least the 95th percentile of publications is used.

We conclude that, based on these four databases, that the *h*-index for these two SRAs is completely different. In such a case, which *h*-index should be used for "official" purposes, if requested? It appears that the GS-based *h*-index reflects a wider scope of publications and not only those limited to Elsevier- or Clarivate Analytics-related data, but more widely

¹ <https://scholar.google.com/citations?user=dWCOqCkAAAAJ&hl=en>.

² <https://scholar.google.com/citations?user=L85nCpoAAAAJ&hl=en>.

³ <https://www.elsevier.com/solutions/scopus>.

⁴ <https://clarivate.com/products/web-of-science/>.

⁵ https://www.researchgate.net/profile/Jaime_Teixeira_Da_Silva (this represents the most accurate and up-to-date database of publications for the first author, curated by the first author).

⁶ https://www.researchgate.net/profile/Judit_Dobranszki (this represents the most accurate and up-to-date database of publications for the first author, curated by the second author).

⁷ <http://clarivate.libguides.com/webofscienceplatform/coverage> (coverage: 138 million records (journals, books, and proceedings)).

⁸ https://www.elsevier.com/__data/assets/pdf_file/0007/69451/0597-Scopus-Content-Coverage-Guide-US-LETTER-v4-HI-singles-no-ticks.pdf (coverage: 69 million core records + 147 million non-core records or documents).

Table 1 Summary of *h*-indexes for both authors of this paper based on four most popular academic databases: Google Scholar (GS), Scopus, Web of Science (WoS), and ResearchGate (RG)

	GS		Scopus		WoS		RG	
	All	ESC	All	ESC	All	ESC	All	ESC
First author	42	NA	29	24	8	NA	36	30
Second author	15	NA	11	10	9	NA	13	12

Two values are listed, for all citations, and excluding self-citations (ESC)

NA not available on these databases (the authors assume that the parameter “all citations” also includes self-citations, but does not issue a separate *h*-index excluding self-citations). All data accurate on October 6, 2017

to any publicly available source on the internet. If so, then a request to provide WoS⁷- or Scopus⁸-based *h*-indexes for official academic purposes—especially if that score is then used to select appropriate referees—is unrealistic simply because it reflects a fragment of the publishing profile—at least for both authors of this paper—and may exclude other papers that have narrower readership or fewer citations, but that might represent other important aspects of their scholarly record. However, readers are cautioned that GS carries errors, such as the inclusion of duplicate entries (Adriaansee and Rensleigh 2013), or inaccuracies as exemplified below. An argument that could be made is that if all applicants are judged based on exactly the same *h*-index source, then *h*-index values/scores are comparable, and the relative value reflects an accurate SRA-to-SRA comparison, as occurs with the RG Score, but to iron out differences and discrepancies between them, a median of all four *h*-index values/scores could be used (so an *h*-index of 32.5 and 12 for the first and second author, respectively). Ultimately, any data (in this case, the *h*-index) derived from limited, flawed or faulty databases will provide an inaccurate, or misleading, metric.

Curiously, a study by Bar-Ilan (2008), which assessed the *h*-index of 40 highly cited Israeli researchers showed widely varying values/scores across GS, WoS and Scopus, but several of the SRAs had, except for two researchers, WoS and Scopus *h*-index values that did not differ by more than three units, suggesting that those databases were far more alike almost a decade ago than now, as shown by the small analysis in Table 1, where WoS- and Scopus-based *h*-index values differ considerably. Bar-Ilan (2008) did not provide a suitable explanation as to why the *h*-index values varied so widely for the same SRA, although it can likely be ascribed to differences in the source of data in the different databases and also to the information selected from those databases to derive the *h*-index.

Recommendations, limitations and conclusions

Any index is as good as its destined use as well as on the ultimate purpose or end-use. Thus, indexes that are limited, faulty, short-sighted or unrepresentative of an SRA’s academic and publishing profiles should never be used. There is prestige and even reward associated with being a highly cited researcher, leading to such a status quo title by Clarivate Analytics (see some criticisms in Teixeira da Silva and Bernès 2017), but this is based exclusively on the JIF, i.e., a journal-based metric used erroneously as an ABM. The *h*-index is an ABM, and currently four main/popular sources or databases provide an *h*-index score for an SRA (there may be more): Google’s GS, Elsevier’s Scopus, Clarivate Analytics’ WoS, and RG, the latter rising rapidly in prominence among SRAs. However, as demonstrated in this case study, the *h*-indexes for the two SRAs (authors) of this paper

differ widely, making the choice of source of the h -index for any official purpose confusing and unreliable. Next follow some possible reasons for these discrepancies, as well as other issues, topics and limitations of our study:

- (a) how is data curated for any of these databases, and is it reliant on user-based updates, or automatic feeds?
- (b) while GS provides citation links for articles that are available on the internet, Scopus and WoS use their own databases, so there are always proprietary and pay-to-view limitations;
- (c) SRAs are increasingly using ASNSs to expand exposure to their research and publications, including GS, RG, Academia.edu, and others such as Mendeley (owned by Elsevier), Loop (owned by Frontiers), and ORCID, in which the SRA is expected to maintain the accuracy of their own publishing profiles. However, one or more of these ASNSs are known not only to automatically establish a profile for the SRA, absent their explicit permission, while in other cases, SRAs might not have the time, or the patience, to maintain several databases accurate, making the h -index, at least for GS, a source of potentially incorrect, inaccurate and unreliable information (Halevi et al. 2017). As an example, the GS profile of Leonid Schneider,⁹ a science journalist and science watchdog (Teixeira da Silva 2016), reflected until recently when a public complaint was made on his blog on September 7, 2017, an incorrect and inaccurate portrayal of his publishing profile as a result of the inclusion of other authors with the same name L. Schneider). Therefore, the GS-based h -index for Leonid Schneider had been grossly inflated with a score of 14. Schneider has since then cleaned up his GS profile, reflecting 50% of the originally displayed h -index. This indicates how poorly curated or outdated GS profiles can easily mislead other SRAs. Unlike the two authors of this paper (Table 1), the GS-based h -index for Schneider and his RG-based h -index score are identical, at 7, which may reflect a truncated scientific career, and a relatively small publications record.
- (d) The gap between WoS and GS may be closing, but is not closed yet, given a recent collaboration that was established between WoS and GS,¹⁰ although it is unclear at this time if this will affect the accuracy of data that contributes to the h -index. Moreover, the discrepancy between WoS and Scopus is not solved but can be explained by differences in databases.
- (e) Our small analysis, which only assesses the data for the authors of this paper, is based on a sample size of 2, so inferences should not be extrapolated toward the entire SRA pool globally. However, since caution is being recommended, it would be worthwhile, for bibliometric specialists, to conduct a wide-scale meta-data analysis to determine if the discrepancies between these four h -indexes apply more widely to a larger pool of SRAs, reaching the same conclusions, or indicating that our results represent outliers.

⁹ <https://scholar.google.com/citations?user=StDfMFQAAAAJ&hl=en> (approximately 35 papers previously listed, with one duplicate, now showing 15 after correction; compare to 20 papers on RG account: https://www.researchgate.net/profile/Leonid_Schneider). He was accused of fraud on Twitter, claiming initially that he would not clean up his GS profile: <https://twitter.com/schneiderleonid/status/906080994758930432> (“I’ve better things to do than daily check stuff Google Scholar adds to my profile. Now that I was accused of fraud, I sure won’t clean it up”); <https://twitter.com/schneiderleonid/status/906064401274470400> (“I won’t be deleting false papers, since I never use Google Scholar as my official publications list reference. Let’s see how this continues.”).

¹⁰ <http://wokinfo.com/googlescholar/>.

- (f) Self-citations may have the ability to inflate the h -index, leading some to claim that the self-citation-free h -index should be used rather than the full (i.e., inclusive of self-citations) h -index, although this assumption may assume incorrectly, or falsely, that self-cites are used inappropriately, or out of context. Flatt et al. (2017) developed a self-citation index that reports on the level of self-citation with the objective of curbing this practice.
- (g) The exclusive use of the h -index may over-value some SRAs while under-valuing others (Costas and Bordons 2007). It is for this reason, to undercut the importance and weighting that any single factor or metric has on an SRA's performance profile, that the Global Science Factor was envisioned (Teixeira da Silva 2013).
- (h) Even though the Scopus- and RG-based h -indexes, at least according to the data in Table 1, show similar or comparable “all citations” : “self-citation” ratios, it is unclear how RG calculates its h -index, i.e., is the calculation proprietary, as for the RG Score?
- (i) Should retractions or retracted papers count towards an SRA's h -index? Ultimately, retracted papers that continue to be cited result in an unfair portrayal of that SRA (Teixeira da Silva and Bornemann-Cimenti 2017; Teixeira da Silva and Dobránszki 2017), and if used to calculate the h -index of authors of retracted papers, may be offering an unfair, or misleading, perspective of those SRAs.

Compliance with ethical standards

Conflict of interest The authors declared that they have no conflict of interest.

References

- Adriaanse, S. L., & Rensleigh, C. (2013). Web of Science, Scopus and Google Scholar: A content comprehensiveness comparison. *The Electronic Library*, 31(6), 727–744. <https://doi.org/10.1108/EL-12-2011-0174>.
- Bar-Ilan, J. (2008). Which h -index?—A comparison of WoS, Scopus and Google Scholar. *Scientometrics*, 74(2), 257–271. <https://doi.org/10.1007/s11192-008-0216-y>.
- Bornmann, L., & Daniel, H.-D. (2009). The state of h index research. Is the h index the ideal way to measure research performance? *EMBO Reports*, 10(1), 2–6. <https://doi.org/10.1038/embor.2008.233>.
- Costas, R., & Bordons, M. (2007). The h -index: Advantages, limitations and its relation with other bibliometric indicators at the micro level. *Journal of Informetrics*, 1, 193–203. <https://doi.org/10.1016/j.joi.2007.02.001>.
- Flatt, J. W., Blassime, A., & Vayena, E. (2017). Improving the measurement of scientific success by reporting a self-citation index. *Publications*, 5, 20. <https://doi.org/10.3390/publications5030020>.
- Glänzel, W., & Persson, O. (2005). H -index for price medallists. *ISSI Newsletter*, 1(4), 15–18.
- Halevi, G., Moed, H., & Bar-Ilan, J. (2017). Suitability of Google Scholar as a source of scientific information and as a source of data for scientific evaluation—Review of the literature. *Journal of Informetrics*, 11(3), 823–834. <https://doi.org/10.1016/j.joi.2017.06.005>.
- Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences USA*, 102(46), 16569–16572. <https://doi.org/10.1073/pnas.0507655102>.
- Lippi, G., & Mattiuzzi, C. (2017). Scientist impact factor (SIF): A new metric for improving scientists' evaluation? *Annals of Translational Medicine*, 5(15), 303. <https://doi.org/10.21037/atm.2017.06.24>.
- Popova, O., Romanov, D., Drozdov, A., & Gerashchenko, A. (2017). Citation-based criteria of the significance of the research activity of scientific teams. *Scientometrics*, 112(3), 1179–1202. <https://doi.org/10.1007/s11192-017-2427-6>.
- Saraykar, S., Saleh, A., & Selek, S. (2017). The association between NIMH funding and h -index in psychiatry. *Academic Psychiatry*, 41, 455–459. <https://doi.org/10.1007/s40596-016-0654-4>.
- Svider, P. F., Husain, Q., Folbe, A. J., Couldwell, W. T., Liu, J. K., & Eloy, J. A. (2014). Assessing national institutes of health funding and scholarly impact in neurological surgery. *Journal of Neurosurgery*, 120(1), 191–196. <https://doi.org/10.3171/2013.8.JNS13938>.

- Teixeira da Silva, J. A. (2013). The global science factor v. 1.1: A new system for measuring and quantifying quality in science. *The Asian and Australasian Journal of Plant Science and Biotechnology*, 7(Special Issue 1), 92–101.
- Teixeira da Silva, J. A. (2016). Science watchdogs. *Academic Journal of Interdisciplinary Studies*, 5(3), 13–15. <https://doi.org/10.5901/ajis.2016.v5n3p13>.
- Teixeira da Silva, J. A. (2017). The journal impact factor (JIF): Science publishing's miscalculating metric. *Academic Questions*, 30(4), 433–441.
- Teixeira da Silva, J. A., & Bernès, S. (2017). Clarivate analytics: Continued *omnia vanitas* impact factor culture. *Science and Engineering Ethics*. <https://doi.org/10.1007/s11948-017-9873-7>. (in press).
- Teixeira da Silva, J. A., & Bornemann-Cimenti, H. (2017). Why do some retracted papers continue to be cited? *Scientometrics*, 110(1), 365–370. <https://doi.org/10.1007/s11192-016-2178-9>.
- Teixeira da Silva, J. A., & Dobránszki, J. (2017). Highly cited retracted papers. *Scientometrics*, 110(3), 1653–1661. <https://doi.org/10.1007/s11192-016-2227-4>.