Journal of Scientometric Res. 2019; 8(3):125-130
http://www.jscires.org

Research Article

In Search of a Silver Bullet: Evaluating Researchers' Performance in Bosnia and Herzegovina

Haris Memisevic[1,*], Arnela Pasalic[1], Edin Mujkanovic[2], Mahira Memisevic[3]

[1]University of Sarajevo, Faculty of Educational Sciences, Sarajevo, BOSNIA AND HERZEGOVINA.
[2]Herzegovina University, Department of Special Education, Mostar, BOSNIA AND HERZEGOVINA.
[3]University of Sarajevo, Faculty of Science, Sarajevo, BOSNIA AND HERZEGOVINA.

ABSTRACT
Background: Evaluating academic production and researchers' impact has become a common practice in many areas of academic life. Researchers are being evaluated for many purposes such as getting employment, promotion, tenure and winning grants. Achieving full objectivity of the evaluation is a rather difficult, if not the impossible task. The goal of the present paper was to evaluate research performance of scholars from Bosnia and Herzegovina (BiH and to propose a single number that captures several scientometric indices. Methods: We took data from 303 scholars from 4 public universities in BiH on their number of citations and *h*-indexes derived from four databases/services: Web of Science, Scopus, Google Scholar and Research Gate. In addition to this, we performed a Principal Component Analysis of number of citations and *h*-indexes from these indices to come up with a single number that best captures the scientific impact of the researchers. Results: The results of this study indicate a strong relationship between all indices of scholarly achievement as measured through citations and *h*-indexes. Principal component analysis has shown that it is possible to obtain a single number that captures researchers' scientific impact. Conclusion: Many metrics can be useful in evaluating researchers' scientific impact. As the researchers in BiH have a low scientific production, universities in BiH need to adapt a strategy to stimulate the increase in their scientific productivity.

Keywords: Academic performance, Evaluation, *h*-index, Principal component analysis, Bosnia and Herzegovina.

Correspondence

Haris Memisevic

University of Sarajevo, Faculty of Educational Sciences-71000 Sarajevo, BOSNIA AND HERZEGOVINA.

Email: hmemisevic@gmail.com

Received: 21-08-2019;
Revised: 11-11-2019;
Accepted: 23-12-2019.
DOI: 10.5530/jscires.8.3.27

INTRODUCTION

Academic workers are constantly evaluated in many areas of their everyday jobs. These areas can roughly be divided into three segments: teaching, administrative duties and research, all of which are subject to evaluation. Teaching and teaching effectiveness are most frequently evaluated through student's evaluation of teaching.[1] Although important and valuable, students' evaluations are mostly subjective and do not represent a legitimate measure of teaching quality.[2] On the other hand, performing administrative duties can be assessed more objectively through the number of hours academics serve on faculty committees, membership on college boards, tutorship of students, chairing a department, etc.[3] Administrative duties are usually not evaluated for the purposes of getting tenure or grant. Lastly, evaluating scholars' contribution in research or their scientific impact, on the face value, appears to be a very objective procedure. However, it is not a simple task as it might look. This article deals with ways of quantifying scientific impact, as it is the most important pillar of academic work required for career advancement.[4] Quantifying scientific contribution of scholars has become an important task not just in academia but also in public at large. Whether it is employment at the university, tenure, promotion, grant, or any other purpose, various committees evaluate scholars' scientific impact on a regular basis. Measuring academic accomplishments has become a "fact of scientific life".[5] Researchers themselves are aware that without valuable scientific contribution their chances of winning a project or getting tenure are significantly reduced.[6] Successful academic career is dependent on many factors including ingenuity, longevity, publications and even luck.[7,8]

Copyright
© The Author(s). 2019 This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

But what is the best way to measure the scientific impact of scholars? Unfortunately there are no clear-cut answers that would be universal and applicable throughout the world. One way to measure scientific impact is through the number of citations that certain publication gets.[9] One of the most widely used methods for evaluating scientific output is the *h*-index, which is the number of publications that have as many number of citations.[10] Although not without flaws,[11,12] *h*-index is a commonly accepted, single-number, indicator of scientific output. *h*-index has been well received in the scientific community and has been thus widely accepted.[13] Many authors have asked how meaningful is the *h*-index as a measure of individual researcher's scientific output.[14] Even if we accept the *h*-index as a "gold-standard" in measuring scientific impact, we still might ask the question which *h*-index is the best and will be used for the evaluation purposes. Whether it will be yet another "gold-standard" – Web of Science's (WOS) *h*-index, or maybe some other *h*-index will be more useful? Studies show that it matters which citation database will be used for calculating the *h*-index as some scientific fields might be over or under represented in certain databases.[15] For example, there are several objections to using WOS as a sole criterion for evaluating scientific performance. Some of these objections are its limited coverage for the Social Sciences and Humanities, the lack of inclusion of journals in languages other than English and lack of citations from sources not covered by WOS,[16] to name just a few.

Due to its wider coverage, especially in the area of Social Sciences and Humanities, a Scopus *h*-index might be used as a viable alternative to WOS. However, it is important to note that there are limitations in the content of Scopus database.[17] Another limitation for both, WOS and Scopus is that they are subscription based and are not freely available to researchers. Some authors have also criticized WOS and Scopus for their lack of comprehensiveness.[18] This brings us to yet another alternative and that is Google Scholar (GS). GS is one of the largest, if not the largest, scientific database covering a large number of journals and academic web documents from the Google search engine.[19] Researchers can freely create their Google Scholar Profiles, which in turn, automatically retrieves the papers they authored or the authors can manually add the papers to their profiles. GS automatically retrieves the number of citations for these papers and calculates the *h*-index. Some authors have reported that the accuracy of GS services is often inconsistent.[20] Last alternative service that we mention is the Research Gate (RG). RG is a social network for researchers where they can list their publications and interact with other scientists throughout the world. It is acknowledged that scholars have interest to read articles on social networks[21] and thus the interest for the service such as RG is growing in popularity.[22] Another benefit of RG is that it is quite easy to use, it is free for users and offers several numerical indices as a measure for the authors scientific contribution, including the number of citations and *h*-index. RG in creating these indices combines bibliometric and altmetric data to create a more comprehensive measure for researchers.[23] Overall, commercially available databases, WOS and Scopus, are more accurate than GS and RG, as the authors cannot themselves add papers to their profiles. One of the shortcomings of WOS and Scopus is that there can be a long time lag between publication of the paper and its visibility in these bases.[6] In GS and RG, these publications can be almost instantly visible.

Bosnia and Herzegovina (BiH) is a country on a scientific periphery with a small scientific production. Scientific periphery is the term used for countries that lag behind in scientific production.[24] Universities in BiH are not represented in Academic Ranking of World University list and the only global ranking list that covers universities in BiH is created by Webometrics. According to Webometrics, July 2019 Edition, the best ranked university in BiH is the University of Sarajevo with a world rank of 1669. This relatively poor ranking of BiH universities has led universities to urge their staff to become more "visible". Thus, in order to increase scientific visibility of research conducted in BiH and of researchers themselves, universities in BiH have recommended their staff to create their Google Scholar Profiles. There is a lack of scientometric studies examining scholars and their scientific output from BiH and thus the need for one such study.

Research objectives

The goal of the present paper is to evaluate research performance of scholars from BiH and to compare the number of citations and *h*-indexes from four (WOS, Scopus, GS and RG) databases at the researcher's level. Another goal was to perform a principal component analysis (PCA) of these measures in order to find major dimensions of these indices. PCA is a multivariate technique that analyses data of inter-correlated variables with a goal of extracting important information from the data and to represent it as new variables or principal components.[25]

METHODS

Participants

The final sample for this study comprised 303 researchers from four largest public universities in BiH, namely University of Sarajevo, University of Tuzla, University of Banja Luka and University of Mostar. Descriptive data on the researchers in presented in Table 1. We first extracted data from Google Scholar for the first 100 researchers from each of these Institutions- 400 in total. We then checked whether these researchers had a Research Gate (RG) profile and only those with a RG account were kept for further analysis. This left us with a sample of 323 full profiles. Out of these, 20 cases represented significant multivariate outliers according to Mahalanobis distances and were thus excluded from the further analysis.

Procedure

For each of the 303 scholars, we took the following data: number of WOS publications, number of WOS citations, WOS *h*-index, number of Scopus citations, Scopus *h*-index.,

Table 1: Descriptive data of the sample of researchers from Bosnia and Herzegovina.

University	Males (n %)	Females (*n* %)	Total (*n*)
Sarajevo	36 51.4	34 48.6	70
Banja Luka	56 69.1	25 30.9	81
Tuzla	45 59.2	31 40.8	76
Mostar	44 57.9	32 42.1	66
Total	181 59.7	122 40.3	303

Table 2: Descriptive data for number of citations and *h*-indexes of researchers in Bosnia and Herzegovina.

Indice	First quartile	Median	Third quartile
WOS citations	4	19	70
WOS_h_index	1	2	4
Scopus citations	12	49	139
Scopus_h_index	2	3	6
GS citations	66	141	296
GS_h_index	4	6	9
ResearchGate cit.	34	86	203
ResearchGate_h_index	3	5	7

Note. GS citations- Google Scholar Citations; GS_h_index- Google scholar h index.

number of Google Scholar citations, Google Scholar *h*-index, number of RG citations and RG *h*-index. Data were extracted on 15th and 16th of August 2019.

Statistical analysis

We provided descriptive data for all indices and performed a correlation analysis between all the metric indices. As the data were not normally distributed, we used Spearman rank-order correlations. We then performed a principal component analysis and the extracted factor was again correlated with the metric indices. An alpha level was set to .05.

RESULTS

We first present the numbers of citations and *h*-indexes for all researchers. These results are shown in Table 2.

Table 2 clearly shows the well-established pattern of citation distribution, with largest number of citations and *h*-index for GS and smallest for WOS.

We next calculated the correlation between all the metric indices. As the distribution of citation counts was skewed, we used Spearman rank-order correlations (Table 3).

As can be seen from Table 3. all correlations are high and statistically significant at .001 level. The highest overall correlation was between WOS number of citations and WOS *h*-index (r=.94) and the smallest was between GS number of citations and WOS *h*-index (r=.69). We then performed a Principal Component Analysis (PCA) to see whether there is

Table 3: Correlation between number of citations and h indexes of Web of Science, Scopus, Google Scholar and ResearchGate.

	GS	GS_h_ index	WOS	WOS_h_ index	Scopus	Scopus_h_ index	RG	RG_h_ index
GS	1	--	--	--	--	--	--	--
GS_h_ index	0.89	1	--	--	--	--	--	--
WOS	0.71	0.67	1	--	--	--	--	--
WOS_h_ index	0.69	0.70	0.94	1	--	--	--	--
Scopus	0.82	0.75	0.84	0.78	1	--	--	--
Scopus_h_ index	0.78	0.80	0.79	0.80	0.93	1	--	--
RG	0.92	0.83	0.77	0.75	0.86	0.82	1	--
RG_h_ index	0.86	0.88	0.75	0.78	0.82	0.84	0.93	1

Note. GS - Google Scholar Citations; GS_h_index- Google scholar h index.

Table 4: Spearman rank-order correlations between first principal component and citation indices.

Indice	Prin1
Google Scholar	0.92
Google Scholar *h*-index	0.91
WOS	0.86
WOS_h_index	0.86
Scopus	0.92
Scopus_h_index	0.92
RG	0.94
RG_h_index	0.95

an underlying factor that can capture scores from all the indices. According to the PCA, the first extracted factor explained 84.1% of variance in the scores. The second factor explained 6% of the variance and the third factor explained 5%. The first component appears to capture most variation in citations and *h*-index scores. In Table 4. We present Spearman rank-order correlations between PCA first factor and metric indices.

As can be seen from Table 4, all correlations between the extracted factor and metric indices were highly correlated and statistically significant at .001 level. All correlations were close to or larger than 0.9.

It seems possible to obtain a single number that captures most of the variance for citations and *h*-indexes obtained from various databases. The formula for getting the single number representing all the indices:

X =

0.0015333944 • Scholar

+0.0108265905 • h_indexSCH

+0.0031350675 • WOS

+0.1269951052 • WOS_h_index

+0.002429375 • Scopus

+0.1163433515 • Scopus_h_index

+0.002121713 • RG

+0.110819443 • RG_h_index

+−3.312167833

This formula is probably sample specific, but in this study it captured a massive 84.1% of the variance in scores.

DISCUSSION

The goal of the present study was to examine the relationship between number of citations and *h*-indexes obtained from four different databases: WOS, Scopus, GS and RG. The total number of citations and *h*-index were highest for Google Scholar and smallest for WOS. This finding is in line with existing studies.[26,27] Although we found strong and highly significant correlations between all different metric indices, they seem to differ among each other in their size. For example, there was a stronger relation between RG *h*-index and GS *h*-index, than between RG *h*-index and WOS *h*-index. These results have important implications for evaluating scholars in BIH and also for other countries with small scientific production. The main question is: How justified is it to use WOS when its median *h*-index is 2? Here it is important to note that this median number is retrieved from 303 of the most prolific researchers and that this number would probably be closer to 0 if we analyzed all researchers from these 4 largest public universities in BiH. Thus, we believe it would be more informative to evaluate researchers from developing countries through a GS metric either as a standalone procedure or as a complementary method. The similar conclusion was reached by other authors as well.[28] Despite these limitations, the results of this study can be very useful and informative to universities for tenure and/or promotion purposes. If a researcher is applying for a position and has a WOS *h*-index of 5, than his/her *h*-index is in the top 25% of the researchers at the public universities in BiH. Other studies have also suggested similar methodology in the researcher's evaluation process.[29] However, there needs to be a control for the discipline the author is coming from, as it may affect the number of citations. As an example, the field of medicine is getting more citations than the field of mathematics.[30]

Another goal of this study was to determine whether a single number can capture researcher's scientific impact. This was tested through the principal components analysis and the results indicated that 84.1% of the variance in citations and *h*-indexes scores were explained by a single component. Likewise, the principal component was significantly related to all metric indices. We suggested a formula that captures most of the variance in the metric scores. Although the formula is sample specific, it can be used as a single number indicator of a scholar's scientific impact and could potentially be used for scientometric purposes. Hypothetically it would be ideal if a scientific impact could be accurately measured with a single number. However, that ideal is hard, if not impossible, to reach. Many authors have postulated that scientific impact is a multidimensional construct that cannot be properly expressed through a single number and that it is always wise to use several indicators to measure research performance.[9,14] Although we agree with this position that several indicators can better illustrate scholar's research output, we believe it is even more important that these numeric indicators truly represent scholar's output. The analysis of excluded outliers in our study can best illustrate this point. Although, most of the highly significant outliers were excluded because they had significantly higher number of citations and significantly higher *h*-index (mostly of authors from University of Sarajevo), some outliers were excluded as they had a large discrepancy between number of citations and *h*-index. For example, the author has 500 WOS citations but his/her WOS *h*-index is 2. The article that received almost all citations was the study of international group in which our author was one of the several (usually more than 10) co-authors of the article and which was published in a Journal with a high impact factor such as *Nature* of *The Lancet*. What is the real scientific impact of that author and can he/she be evaluated at all? Another example of discrepancy in scores deals with the number of articles published in certain journals. For example, an author publishes papers mostly (or exclusively) in Journals covered by Scopus and not by WOS and the Scopus *h*-index is 14 and WOS *h*-index is 3. Again, how shall we evaluate this author? If the components of our proposed formula truly represented scholar's scientific impact we believe it could be than used as a single-number indicator of scientific impact. However, several above-mentioned examples discourage us from offering a scientometric panacea at the individual researcher's level. Therefore, besides number of citations and *h*-indexes obtained from several databases, in the formula for a scientometric "silver-bullet", we need to include and answer the following: 1. Is the author first or corresponding author of the paper? 2. In how many articles is our author first or corresponding author and how many citations these publications received? 3. How many scholars co-authored the article and how many times these authors have cited the paper (auto-citations)? 4. What is the scientific field the author/s belongs to? These and probably other questions need to be taken into considerations when evaluating researchers' scientific impact. Of course, the different weight in the formula needs to be given to each of these questions.

Some of these limitations of *h*-index have been noted earlier and numerous other measures have been put forward.[13] Again, most of the objections to *h*-index are justified, but many more inaccuracies regarding measures of scientific impact stem from academic misconduct through inappropriate authorship. This topic has received much scientific attention in the recent years, but problems with honorary or ghost authorship still persist.[31-33] For example, there were some authors who had high *h*-indexes in many databases, but were not the first or corresponding authors in any of the cited papers. In such cases, are we evaluating these authors networking skills or their scholarly impact? Although it is possible that these authors were valuable and legitimate team members, without whose contribution such articles would have not be written, it still, nevertheless, opens a lot of space for doubt. One way to, at least partially, circumvent this problem is through the use of the so-called Zerem-score or Z-score.[34] This measure offers overall scientific score as the sum of author score and author citation score. It also discourages expansion of paper's author list, so the authors whose contribution was minor or non-significant will not be included.

Another point that needs to be further explored is the variation in subject specific scores. For example, is this, "silver bullet" measure a better fit for Natural scientists than for social scientists? This question, along with the in-depth investigation of the authors who represent significant outliers (mentioned above) will provide us with a clearer profile of research and researchers in BiH.

Concluding Remarks

There are many objective obstacles in evaluating scholars' scientific impact or research performance. In this article we provided the "shape of science" in BiH through the analysis of researchers from 4 largest public universities. We also performed a Principal Component Analysis (PCA) of the metric indices and showed the way to come up with a formula to represent scientific contribution through a single number. Extreme outliers were excluded from the analysis and we explained the criteria for the exclusion. Given the relatively poor research output, Universities in BiH need to take some strategic steps in order to increase the scientific productivity of their staff. This can be achieved by offering different incentives to academics, such as reducing teaching loads, increasing number of grants etc.[35] The authorities in BiH can also help also by increasing their investments in universities. These strategies are known to result in greater scientific production.[36] Future studies should include other measures in PCA as well, such as number of articles in WOS, Scopus and RG. PCA can be used as a good method of extracting common variance from many metric indices and perhaps offer us a formula through which we will get a single number representation of the scientific impact.

ACKNOWLEDGEMENT

None.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ABBREVIATIONS

BiH: Bosnia and Herzegovina; WOS: Web of Science; GS: Google Scholar; RG: Research Gate.

REFERENCES

1. Uttl B, White CA, Gonzalez DW. Meta-analysis of faculty's teaching effectiveness: Student evaluation of teaching ratings and student learning are not related. Studies in Educational Evaluation. 2017;54:22-42.
2. Hornstein HA. Student evaluations of teaching are an inadequate assessment tool for evaluating faculty performance. Cogent Education. 2017;4(1):1304016.
3. García-Gallego A, Georgantzís N, Martín-Montaner J, Pérez-Amaral T. (How) Do research and administrative duties affect university professors' teaching?. Applied Economics. 2015;47(45):4868-83.
4. Cadez S, Dimovski V, Zaman GM. Research, teaching and performance evaluation in academia: The salience of quality. Studies in Higher Education. 2017;42(8):1455-73.
5. Lane J. Let's make science metrics more scientific. Nature. 2010;464(7288):488.
6. Memisevic H, Taljic I, Hadziomerovic AM. Making Use of *h*-index: The Shape of Science at the University of Sarajevo. Acta Informatica Medica. 2017;25(3):187.
7. Petersen AM, Wang F, Stanley HE. Methods for measuring the citations and productivity of scientists across time and discipline. Physical Review E. 2010;81(3):036114.
8. Sinatra R, Wang D, Deville P, Song C, Barabási AL. Quantifying the evolution of individual scientific impact. Science. 2016;354(6312):5239.
9. Bollen J, DeSompel HV, Hagberg A, Chute R. A Principal Component Analysis of 39 Scientific Impact Measures. Plos One. 2009;4(6):e6022.
10. Hirsch JE. An index to quantify an individual's scientific research output. Proceedings of the National academy of Sciences. 2005;102(46):16569-72.
11. Purvis A. The *h*-index: Playing the numbers game. Trends in Ecology and Evolution. 2006;21(8):422.
12. Yong A. Critique of Hirsch's citation index: A combinatorial Fermi problem. Notices of the AMS. 2014;61(9):1040-50.
13. Bornmann L, Daniel HD. What do we know about the *h*-index?. Journal of the American Society for Information Science and Technology. 2007;58(9):1381-5.
14. Bornmann L, Daniel HD. The state of *h*-index research. EMBO Reports. 2009;10(1):2-6.
15. Bar-Ilan J. Which *h*-index?—A comparison of WoS, Scopus and Google Scholar. Scientometrics. 2008;74(2):257-71.
16. Harzing AW, DerWal R. A Google Scholar *h*-index for journals: An alternative metric to measure journal impact in economics and business. Journal of the American Society for Information Science and Technology. 2009;60(1):41-6.
17. Jacsó P. The pros and cons of computing the *h*-index using Scopus. Online Information Review. 2008;32(4):524-35.
18. Ghazavi R, Taheri B, Ashrafi-rizi H, *et al*. Article Quality Indicator: Proposing a New Indicator for Measuring Article Quality in Scopus and Web of Science. Journal of Scientometric Research. 2019;8(1):9-17.
19. Aguillo IF. Is Google Scholar useful for bibliometrics? A webometric analysis. Scientometrics. 2012;91(2):343-51.
20. Falagas ME, Pitsouni EI, Malietzis GA, Pappas G. Comparison of PubMed, Scopus, web of science and Google scholar: strengths and weaknesses. The FASEB Journal. 2008;22(2):338-42.
21. Asefeh A, Mahboobe H. Correlation between the Articles Citations in Web of Science (WoS) and the Readership Rate in Mendeley and Research Gate (RG). Journal of Scientometric Research. 2018;7(3).
22. Ovadia S. ResearchGate and Academia. edu: Academic social networks. Behavioral and Social Sciences Librarian. 2014;33(3):165-9.
23. Yu MC, Wu YCJ, Alhalabi W, Kao HY, Wu WH. Research Gate: An effective altmetric indicator for active researchers?. Computers in Human Behavior.

2016;55:1001-6.

24. Marušic A, Marušic M. Small scientific journals from small countries: Breaking from a vicious circle of inadequacy. Croat Med J. 1999;40(4):508-14.
25. Abdi H, Williams LJ. Principal component analysis. Wiley Interdisciplinary Reviews: Computational Statistics. 2010;2(4):433-59.
26. Martín-Martín A, Orduna-Malea E, Thelwall M, López-Cózar ED. Google Scholar, Web of Science and Scopus: A systematic comparison of citations in 252 subject categories. Journal of Informetrics. 2018;12(4):1160-77.
27. Thelwall M, Kousha K. ResearchGate versus Google Scholar: Which finds more early citations?. Scientometrics. 2017;112(2):1125-31.
28. Onyancha OB, Ocholla DN. Assessing researchers' performance in developing countries: Is Google Scholar an alternative?. Mousaion. 2009;27(1):43-64.
29. Lacasse JR, Hodge DR, Bean KF. Evaluating the productivity of social work scholars using the *h*-index. Research on Social Work Practice. 2011;21(5):599-607.
30. Eck NJV, Waltman L, Raan AFV, Klautz RJ, Peul WC. Citation analysis may severely underestimate the impact of clinical research as compared to basic research. PloS One. 2013;8(4):e62395.
31. Wislar JS, Flanagin A, Fontanarosa PB, DeAngelis CD. Honorary and ghost authorship in high impact biomedical journals: A cross sectional survey. BMJ. 2011;343:d6128.
32. Gasparyan AY, Ayvazyan L, Kitas GD. Authorship problems in scholarly journals: considerations for authors, peer reviewers and editors. Rheumatology International. 2013;33(2):277-84.
33. Rivera H. Inappropriate authorship and kinship in research evaluation. *Journal of Korean medical science*. 2018;33(13):e105.
34. Zerem E. The ranking of scientists based on scientific publications assessment. Journal of Biomedical Informatics. 2017;75:107-9.
35. Miller JC, Coble KH, Lusk JL. Evaluating top faculty researchers and the incentives that motivate them. Scientometrics. 2013;97(3):519-33.
36. Payne AA, Siow A. Does federal research funding increase university research output?. Advances in Economic Analysis and Policy. 2003;3(1).