On the Measurement of Scientific Performance: Do We Really Need to Take the Distribution of Citations into Account?

Nuno Crespo
Associate Prof., Department of Economics, ISCTE - Instituto Universitário de Lisboa, BRU-ISCTE (Business Research Unit), Lisboa, Portugal.
ORCID iD: https://orcid.org/0000-0002-6546-6325

Nadia Simoes
Associate Prof., Department of Economics, ISCTE - Instituto Universitário de Lisboa, BRU-ISCTE (Business Research Unit), Lisboa, Portugal.
Corresponding Author: nadia.simoes@iscte-iul.pt
ORCID iD: https://orcid.org/0000-0002-4334-4482

Abstract

Despite the drawbacks already pointed out and the wide set of variants suggested to overcome some of them, the h-index is the most used measure to conduct the author-level scientific evaluation. Simplicity is probably its main advantage and the reason for its popularity. The main goal of the current study is to propose an index that is directly linked and almost perfectly correlated with the h-index but is even simpler to obtain. Our index does not require any form of calculation. It disregards the distribution of citations among the papers of the author. Instead, only the total number of papers and the total number of citations of the author is used to obtain the final score. The score of each author can be seen in the table we present in this study. Although much simpler than the h-index (and other measures that account for citation distribution), our index produces the same general conclusions. We illustrate the application of this index with a sample of economists.

Introduction

Almost 15 years after its proposal by Jorge Hirsch (Hirsch, 2005), the perception of the h-index impact is so powerful that sometimes it is seen as the beginning of a new era in bibliometric analysis (Todeschini & Baccini, 2016). It is probably the only bibliometric index that most researchers know (Wildgaard, Schneider, & Larsen, 2014). The h-index was intensively used not only to evaluate author's performance but also the performance of other units of interest, such as journals (Braun, Glänzel, & Schubert, 2006; Franceschini & Maisano, 2010), universities/research centers (Van Raan, 2006), or countries (Csajbók, Berhidi, Vasas, & Schubert, 2007).

An author has an h-index h if h of his/her papers have at least h citations each and the remaining papers have no more than h citations each. Despite its simplicity, the h-index suffers from several well-known shortcomings (Alonso, Cabrerizo, Herrera-Viedmac, & Herrera, 2009; Egghe, 2010). These shortcomings have led to the emergence of many alternative metrics, each of them trying to solve a particular limitation. Among many others, these alternative measures include, for example, the g-index, the m-index, the w-index, the hg-index,

or the Euclidean index (see, for instance, Simoes & Crespo, 2020a). In their extensive survey of author-level bibliometric indicators, Bornmann, Mutz, Hug & Daniel (2011) conclude that many of the metrics proposed to overcome the shortcomings of the h-index (see, for instance, Schubert & Schubert, 2019; Wildgaard, 2019) are highly correlated with it. Therefore it provides "redundant contributions."

Let us discuss five of the most significant drawbacks of the h-index. First, it usually does not include the complete list of papers of the author under analysis. All the author's papers that, until the moment of the evaluation, have less than h citations (including, of course, the uncited papers) are irrelevant, i.e., they do not contribute to the scientific merit of the author under analysis. Second, it produces many ties, namely in the case of authors with medium-low levels of scientific performance. While this is not one of the most stressed shortcomings of the h-index, it is, in our perspective, one of the most critical when the purpose of the evaluation is to produce a ranking of authors. Using a sample of 472 economists from the departments of the economics of the top 10 world universities, Crespo & Simoes (2019) find that 324 authors have an h-index ranging between 0 and 20. To overcome this problem, they introduce a procedure "differentiating the authors with equal h-indices using the papers and the citations that are more than those strictly necessary to achieve that score" (ibid, p. 2497). Third, as mentioned, for example by Fenner, Harris, Levene & Bar-Ilan (2018), once a paper belongs to the h-core (i.e., the h relevant papers), additional citations do not receive credit. This is usually known as the "excess citations" problem and is probably the most discussed shortcomings of the h-index. Let us suppose an author with an h-index of 14 and two different scenarios regarding (for example) the number of citations of the most cited paper: (i) 1000 citations; (ii) 20 citations. Although case (i) corresponds to a situation in which the scientific impact of the most cited paper is much higher (the number of citations is 50 times greater in that case), this does not impact the h-index of the author. Fourth, as almost all the alternative metrics, it does not consider the academic age of the authors (i.e., number of years since the first paper), focusing on the accumulated scientific production. However, this procedure conduces in many cases to unfair comparisons and, therefore, unfair decisions. Which of the following situations corresponds to an author with higher scientific merit: (i) author A with an academic age of 30 years and $h = 20$; (ii) author B with 15 years of academic age and $h = 19$? The answer is author A using the h-index as a reference.

Nevertheless, if we take the academic age into account, the answer can be different. Considering the *m quotient* suggested by Hirsch (2005), corresponding to the ratio between h and the academic age (i.e., the intensity of scientific production, as designated by Simoes and Crespo, 2020b), the answer is B. Fifth, and it does not consider the number of authors of each paper (Hirsch, 2019) and their role in its conception and production. The solution to this problem implies the definition of a co-authorship weighting scheme. It is corresponding to the percentage of the total credit that should be given to that specific author (see Todeschini & Baccini, 2016 for a general discussion on this topic and Marušić, Bošnjak & Jerončić, 2011 for a discussion on the differences in co-authorship patterns across scientific fields). However, while identifying the number of authors is easy to obtain, each author's role in a given paper is hard or even impossible to identify. Some of the solutions that have been adopted include: (i) the use of the byline as a source of information, assuming that the sequence of the names has informative value, which is not consensual since many different rules can be followed (Tscharntke, Hochberg, Rand, Resh & Krauss, 2007); (ii) giving a more significant credit to

the corresponding author; (iii) the identification of specific roles in the paper such as scientific leadership (Hirsch, 2019) or scientific influence (Simoes & Crespo, 2020b). These approaches are essential inputs to adjust the h-index (Egghe, 2008; Abbas, 2011; Wan, Hua & Rousseau, 2007).

Many of the new metrics introduced to overcome the specific limitations of the h-index are valuable contributions to the literature. However, it is important to emphasize that, when compared with the h-index, these measures usually introduce additional levels of complexity. Nevertheless, the simplicity in calculation and interpretation is probably the main reason justifying the tremendous popularity and widespread use of the h-index (Leydesdorff, Bornmann, & Opthof, 2019).

In the present study, we follow a different approach. We propose an index directly linked and almost perfectly correlated with the h-index but is even simpler to obtain. The author's total number of papers and citations is enough to obtain the final performance score, without considering the citations' distribution among the author's papers. Retaining the measurement principles inherent to the h-index but severing the link between papers and citations (i.e., ignoring the distribution of citations), we build a table (presented in this study as Supplementary Material) showing these final scores for each range of values in terms of papers and citations.

In the next section, we introduce our index. Following, we illustrate the application of the measure with a sample of economists working in the world's top universities. The last section concludes.

The new index

We introduce a new measure for assessing scientific performance. It is closely linked with the standard h-index and can be obtained in two simple steps. First, we calculate: (i) the maximum h the author could reach with his/her number of papers ($h_i^{POT}(P)$); (ii) the maximum h the author could reach with his/her total number of citations ($h_i^{POT}(C)$). Following the "rules" adopted by the h-index, these terms correspond to:

$$h_i^{POT}(P) = p_i \tag{1}$$

and

$$h_i^{POT}(C) = int(\sqrt{c_i}) \tag{2}$$

in which p and c represent the total number of papers and the total number of citations, respectively. $h_i^{POT}(C)$ is rounded to the lower value.

Taking into account that in the h-index the conditions related to papers and citations must coexist, in the second step we obtain the h-potential index (h_i^{POT}) as follows:

$$h_i^{POT} = \min\left(h_i^{POT}(P), h_i^{POT}(C)\right). \tag{3}$$

It is immediate to verify that $h_i \leq h_i^{POT}$.

The scientific performance score can be seen in the Table presented as Supplementary Material. Table 1 shows a small part of that table, illustrating the application of our measure. Let us consider the case of author A with 10 papers and 26 citations. This information is not enough to obtain the h-index of this author since the distribution of the citations among the ten papers is not known. The exact value of h can range between 1 and 5. Adopting the measure

introduced in this section we immediately verify that while the author has enough papers to reach $h^{POT}(P) = 10$, the number of citations allows achieving only $h^{POT}(C) = 5$. Therefore, $h^{POT} = 5$. The opposite case can also occur. For example, if an author B has seven papers and 100 citations, the active restriction is given by the number of papers since the h-index cannot be higher than this value. When the distribution of citations is equally distributed among the papers of the author, $h = h^{POT}$. For example, if author C has ten papers with ten citations each, $h = h^{POT} = 10$.

Table 1

Number of papers, number of citations, and h^{Pot}-index

	Interval		Number of papers															
			0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Number of citations		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	3		1	1	1	1	1	1	1	1	1	1	1	1	1	1
	4	8		1	2	2	2	2	2	2	2	2	2	2	2	2	2
	9	15		1	2	3	3	3	3	3	3	3	3	3	3	3	3
	16	24		1	2	3	4	4	4	4	4	4	4	4	4	4	4
	25	35		1	2	3	4	5	5	5	5	5	5	5	5	5	5
	36	48		1	2	3	4	5	6	6	6	6	6	6	6	6	6
	49	63		1	2	3	4	5	6	7	7	7	7	7	7	7	7
	64	80		1	2	3	4	5	6	7	8	8	8	8	8	8	8
	81	99		1	2	3	4	5	6	7	8	9	9	9	9	9	9
	100	120		1	2	3	4	5	6	7	8	9	10	10	10	10	10
	121	143		1	2	3	4	5	6	7	8	9	10	11	11	11	11
	144	168		1	2	3	4	5	6	7	8	9	10	11	12	12	12
	169	195		1	2	3	4	5	6	7	8	9	10	11	12	13	13
	196	224		1	2	3	4	5	6	7	8	9	10	11	12	13	14
	225	255		1	2	3	4	5	6	7	8	9	10	11	12	13	14
	256	288		1	2	3	4	5	6	7	8	9	10	11	12	13	14
	289	323		1	2	3	4	5	6	7	8	9	10	11	12	13	14
	324	360		1	2	3	4	5	6	7	8	9	10	11	12	13	14
	361	399		1	2	3	4	5	6	7	8	9	10	11	12	13	14
	400	440		1	2	3	4	5	6	7	8	9	10	11	12	13	14

Note: columns 14 and 15 for the intervals from 225 onward read 14 and 15 respectively.

Empirical application

Aiming to apply the methodology introduced in the previous section, we consider a group of 4,055 economists. The sample includes members of the top 2019 Think Tanks and Research Institutes in Europe and the United States in economics, according to INOMICS (2019a, 2019b). Sixteen centers were considered (from the list of 20 institutions, one American and three European centers were not included due to their small size, i.e., less than 25 members). The data was drawn from the Scopus database in November 2019 by using a search query by affiliation. Our final sample includes 76,635 papers and 1,493,992 citations. Table 2 presents a characterization of the institutions included.

Table 2

Top 2019 Think Tanks and Research Institutes in Europe and in the United States

Institution	Location	No. of economists	No. of papers	No. of citations
European centers				
1. Centre for Economic Policy Research (CEPR)	United Kingdom	189	4,972	204,284
2. Institute of Labor Economics (IZA)	Germany	198	4,594	102,950
3. German Institute for Economic Research (DIW)	Germany	378	3,255	44,609
4. Information and Research Institute (IFO)	Germany	199	3,614	69,008
5. Research Institute of Industrial Economics (IFN)	Sweden	70	987	17,572
6. Economic and Social Research Institute (ESRI)	Ireland	125	816	9,668
7. Institute for Employment Research (IAB)	Germany	262	1,645	12,467
United States centers				
1. National Bureau of Economic Research (NBER)	Massachusetts	533	12,482	778,123
2. Peter G. Peterson Institute for International Economics (IIE)	District of Columbia	64	1,113	58,619
3. Brookings Institution	District of Columbia	483	3,431	75,181
4. Resources for the Future (RFF)	District of Columbia	183	2,486	58,504
5. Public Policy Institute of California (PPIC)	California	40	275	5365
6. American Enterprise Institute	District of Columbia	186	1,410	15,843
7. Levy Economics Institute, Bard College	New York	573	3,105	58,135
8. Urban Institute	District of Columbia	520	3,3391	70,618
9. Center for Economic and Policy Research (CEPR)	District of Columbia	52	328	5,701

It is important to stress the degree of heterogeneity of the sample, with differences emerging at several levels. The first aspect to highlight is the number of members. In the European sample, the members ranged between 70 (Research Institute of Industrial Economics) and 378 members (German Institute for Economic Research). And between 40 (Public Policy Institute of California) and 583 members (Levy Economics Institute, Bard College) in the American sample. Second, there are researchers with very different levels of publication performance. Table 3 presents the distribution of the h-index for each of the groups. In both cases, the proportion of economists with $h < 3$ is considerably high (51.02% for the European case and 64.20% for the American case). On the other hand, only a tiny fraction of

the researchers can achieve an h-index higher than 10 (13.72% for the European case and 14.20% for the American case). This distribution of publication performance is very close to what can be found in the evidence produced by the RePEc (Research Papers in Economics) ranking of authors for August 2020. Only 5% of registered authors have an $h \geq 14$. In our case, this percentage is 8.48% for the European centers and 10.21% for the American ones.

Table 3

Distribution of the h-index of the economists in the sample

h-index	European centers		American centers	
	N	%	N	%
0	168	11.82%	409	15.53%
1	381	26.81%	992	37.66%
2	176	12.39%	290	11.01%
3	134	9.43%	156	5.92%
4	79	5.56%	91	3.45%
5	44	3.10%	71	2.70%
6	60	4.22%	71	2.70%
7	53	3.73%	60	2.28%
8	44	3.10%	43	1.63%
9	50	3.52%	42	1.59%
10	37	2.60%	35	1.33%
11-15	103	7.25%	160	6.07%
16-20	40	2.81%	93	3.53%
21-30	34	2.39%	84	3.19%
>30	18	1.27%	37	1.40%

Table 4 shows critical information regarding publishing performance for each of the centers included in the analysis.

More specifically, we present the mean and the best author as defined by the h-index and the h^{POT}- index. An exciting conclusion emerging from this evidence is the fact that the author with the best score in both bibliometric measures is the same in the case of 14 out of 16 research centers considered, the exceptions being the Centre for Economic Policy Research (CEPR) and the Research Institute of Industrial Economics (IFN).

Table 4

Publishing performance in the top 2019 Think Tanks and Research Institutes in Europe and in the United States

Institution	h-index		h^{POT}-index	
	Mean	Max	Mean	Max
European centers				
1. Centre for Economic Policy Research (CEPR)	10.47	44 (J. Thisse)	9.70	95 (A. Venables)
2. Institute of Labor Economics (IZA)	8.22	49 (J. Currie)	15.05	97 (J. Currie)
3. German Institute for Economic Research (DIW)	2.94	32 (H. Lütkepohl)	5.03	74 (H. Lütkepohl)
4. Information and Research Institute (IFO)	5.14	63 (R. Tol)	9.25	114 (R. Tol)
5. Research Institute of Industrial Economics (IFN)	5.07	29 (C. Bjørnskov)	9.14	52 (M. Henrekson)
6. Economic and Social Research Institute (ESRI)	3.01	20 (B. Maître)	4.60	33 (B. Maître)
7. Institute for Employment Research (IAB)	2.44	14 (A. Kritikos; U. Blien)	4.54	28 (A. Kritikos)
United States centers				
1. National Bureau of Economic Research (NBER)	11.13	73 (J. Newhouse)	19.70	160 (J. Newhouse)
2. Peter G. Peterson Institute for International Economics (IIE)	6.73	43 (O. Blanchard)	13.16	103 (O. Blanchard)
3. Brookings Institution	2.56	28 (B. Bernanke)	4.73	67 (B. Bernanke)
4. Resources for the Future (RFF)	4.60	33 (R. Newell)	8.60	69 (R. Newell)
5. Public Policy Institute of California (PPIC)	3.40	28 (J. Mount)	5.45	47 (J. Mount)
6. American Enterprise Institute	2.29	22 (S. Satel)	4.20	41 (S. Satel)
7. Levy Economics Institute, Bard College	2.03	131 (A. Kontos)	3.48	131 (A. Kontos)
8. Urban Institute	3.00	34 (R. Berenson)	4.79	65 (R. Berenson)
9. Center for Economic and Policy Research (CEPR)	2.84	17 (S. Dasgupta)	4.88	34 (S. Dasgupta)

Figures 1 and 2 show the distribution of the h^{POT}- index and the h-index for the European and American cases, respectively.
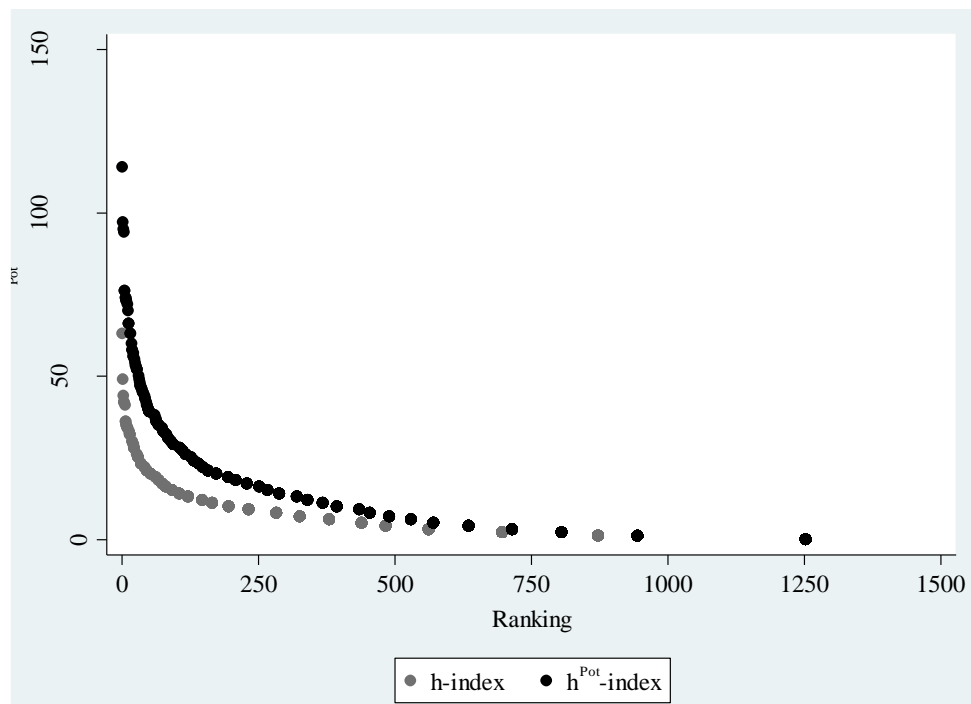
Figure 1: *Distribution of the h^{Pot}-index and the h-index - European centers*
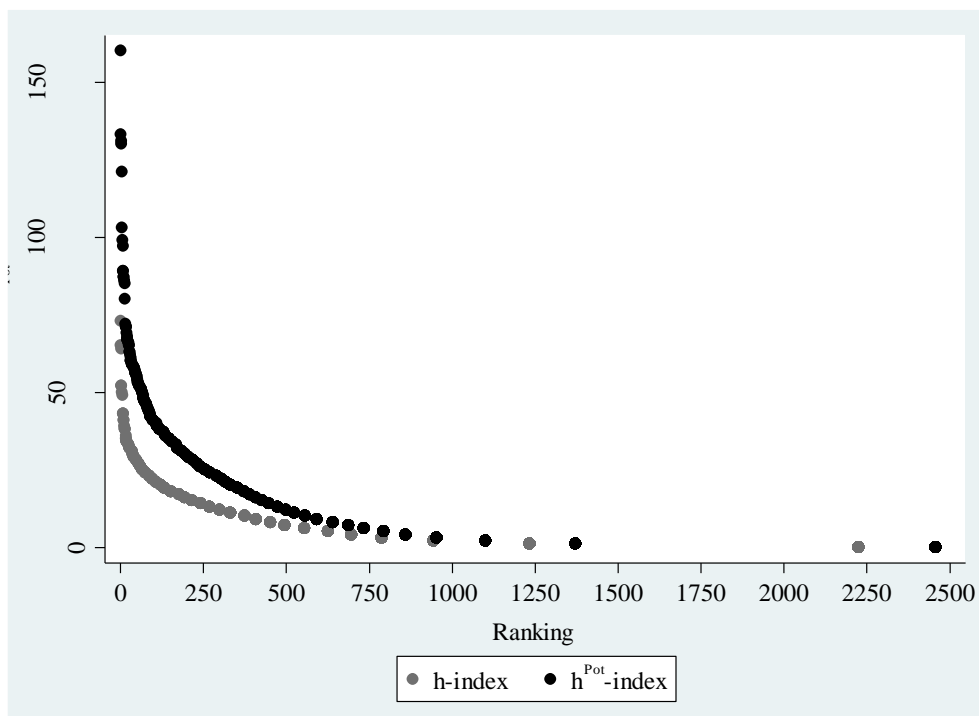

Figure 2: *Distribution of the h^{Pot}-index and the h-index - US centers*

The evidence presented clarifies a substantial similarity in what concerns the distribution of the scores according to both measures. Calculating the Spearman correlation coefficient between this new measure and the h-index we obtain, in both cases (American and European economists), is a very high correlation (0.98 in Europe and 0.95 in the American case). Summing up, the two measures produce the same general conclusions but the h^{POT}-index has the critical advantage of simplicity, requiring no calculations.

As a final remark, to shed some light on whether there were differences in the correlation between both measures for authors from different publication profiles, we divided the European and American samples into two groups: those with $h \leq 10$ and those with $h > 10$. For both groups, the correlations are high and very similar. For the European centers, the Spearman correlation coefficient is equal to 0.963 for those with $h \leq 10$ and 0.957 with $h > 10$. On the other hand, for the American centers, the correlation coefficient is 0.939 for those with $h \leq 10$ and 0.933 for those with $h > 10$.

Conclusion

Bibliometric measures are increasingly used to support significant decisions in universities and research units. However, despite the importance of these metrics, such a quantitative approach also suffers from important limitations as discussed, for example, by Hicks, Wouters, Waltman, De Rijcke, and Rafols (2015) and Hammarfelt and Rushforth (2017). Therefore, it is not surprising that many authors claim for combined use of quantitative and qualitative (peer review) approaches. Nevertheless, either as a sole indicator or an input for more informed peer review, bibliometric measures should be rigorous and straightforward to provide fairer decisions in various scientific dimensions.

The central contribution of the current study was to introduce a measure of publication performance that is highly correlated with the standard h-index but can be obtained from the table presented in this study without any calculations. This is a significant advantage and suggests that the additional effort associated with analyzing the distribution of citations does not pay off.

Acknowledgement

References

Abbas, A. M. (2011). Weighted indices for evaluating the quality of research with multiple authorship. *Scientometrics*, 88(1), 107-131. https://doi.org/10.1007/s11192-011-0389-7

Alonso, S., Cabrerizo, F., Herrera-Viedmac, E. & Herrera, F. (2009). h-index: a review focused in its variants, computation and standardization for different scientific fields. *Journal of Informetrics*, 3(4), 273-289. https://doi.org/10.1016/j.joi.2009.04.001

Bornmann, L., Mutz, R., Hug, S. E. & Daniel, H. D. (2011). A multilevel meta-analysis of studies reporting correlations between the h index and 37 different h index variants. *Journal of Informetrics*, 5(3), 346-359. https://doi.org/10.1016/j.joi.2011.01.006

Braun, T., Glänzel, W. & Schubert, A. (2006). A Hirsch-type index for journals. *Scientometrics*, 69(1), 169-173. https://doi.org/10.1007/s11192-006-0147-4

Crespo, N. & Simoes, N. (2019). Publication Performance Through the Lens of the h-index: How Can We Solve the Problem of the Ties?. *Social Science Quarterly*, 100(6), 2495-2506. https://doi.org/10.1111/ssqu.12696

Csajbók, E., Berhidi, A., Vasas, L. & Schubert, A. (2007). Hirsch-index for countries based on essential science indicators data. *Scientometrics*, 73(1), 91-117. https://doi.org/10.1007/s11192-007-1859-9

Egghe, L. (2008). Mathematical theory of the h-and g-index in case of fractional counting of authorship. *Journal of the American Society for Information Science and Technology*,

59(10), 1608-1616. https://doi.org/10.1002/asi.20845

Egghe, L. (2010). The Hirsch index and related impact measures. *Annual Review of Information Science and Technology*, 44(1), 65-114. https://doi.org/10.1002/aris.2010.1440440109

Fenner, T., Harris, M., Levene, M. & Bar-Ilan, J. (2018). A novel bibliometric index with a simple geometric interpretation. *PloS One*, 13, e0200098. https://doi.org/10.1371/journal.pone.0200098

Franceschini, F. & Maisano, D. (2010). The Hirsch spectrum: a novel tool for analyzing scientific journals. *Journal of Informetrics*, 4(1), 64-73. https://doi.org/10.1016/j.joi.2009.08.003

Hammarfelt, B. & Rushforth, A. (2017). Indicators as judgment devices: An empirical study of citizen bibliometrics in research evaluation. *Research Evaluation*, 26(3), 169-180. https://doi.org/10.1093/reseval/rvx018

Hicks, D., Wouters, P., Waltman, L., De Rijcke, S. & Rafols. I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature News*, 520(7548), 429. https://doi.org/10.1038/520429a

Hirsch, J. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*, 102(46), 16569-16572. https://doi.org/10.1073/pnas.0507655102

Hirsch, J. (2019). h_α: an index to quantify an individual's scientific leadership. *Scientometrics*, 118(2), 673-686. https://doi.org/10.1007/s11192-018-2994-1

INOMICS (2019a). https://inomics.com/insight/top-economics-think-tanks-and-research-institutes-in-europe-1025250

INOMICS (2019b). https://inomics.com/top-economics-think-tanks-research-institutes-us

Leydesdorff, L., Bornmann, L. & Opthof, T. (2019). h_α: the scientist as chimpanzee or bonobo. *Scientometrics*, 118(3), 1163–1166. https://doi.org/10.1007/s11192-019-03004-3

Marušić, A., Bošnjak, L. & Jerončić, A. (2011). A systematic review of research on the meaning, ethics and practices of authorship across scholarly disciplines. *Plos One*, 6(9), e23477. https://doi.org/10.1371/journal.pone.0023477

Schubert, A. & Schubert, G. (2019). All along the h-index-related literature: a guided tour. In *Handbook of science and technology indicators*; Glänzel, W.; Moed, H.; Schmoch, U.; Thelwall, M., Eds.; Springer, Cham. https://doi.org/10.1007/978-3-030-02511-3_12

Simoes, N. & Crespo, N. (2020a). A flexible approach for measuring author-level publishing performance. *Scientometrics*, 122(1), 331–355. https://doi.org/10.1007/s11192-019-03278-7

Simoes, N. & Crespo, N. (2020b). Self-Citations and scientific evaluation: Leadership, influence, and performance. *Journal of Informetrics*, 14(1), 100990. https://doi.org/10.1016/j.joi.2019.100990

Todeschini, R. & Baccini, A. (2016). *Handbook of bibliometric indicators: quantitative tools for studying and evaluating research*. John Wiley & Sons,Weinheim.

Tscharntke, T., Hochberg, M. E., Rand, T. A., Resh, V. H. & Krauss, J. (2007). Author sequence and credit for contributions in multiauthored publications. *PLoS Biol*, 5(1), e18. https://doi.org/10.1371/journal.pbio.0050018

Van Raan, A. (2006). Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups. *Scientometrics*, 67(3), 491-502. https://doi.org/10.1556/Scient.67.2006.3.10

Wan, J. K., Hua, P. H. & Rousseau, R. (2007). The pure h-index: calculating an author's h-index by taking co-authors into account. *COLLNET Journal of Scientometrics and Information Management,* 1(2), 1-5. https://doi.org/10.1080/09737766.2007.10700824

Wildgaard, L. (2019). An overview of author-level indicators of research performance. In: W. Glänzel, H. Moed, U. Schmoch, M. Thelwall (eds.) *Springer Handbook of Science and Technology Indicators*. Springer, Cham, 361-396.

Wildgaard, L., Schneider, J. & Larsen, B. (2014). A review of the characteristics of 108 author-level bibliometric indicators. *Scientometrics*, 101(1), 125-158. https://doi.org/10.1007/s11192-014-1423-3