International Journal of Scientific and Technical Research in Engineering (IJSTRE)
www.ijstre.com Volume 4 Issue 5 ǁ Sep-Oct 2019.
ISSN: 2581-9941

k-Index: A New Dynamic Research Performance Indicator

Reyad El-Khazali*
Reyad El-Khazali is with the ECSE Department of Khalifa University, Abu Dhabi, UAE.
reyad.elkhazali@ku.ac.ae

Abstract— *This paper introduces a new dynamic Research Performance Indicator (DRPI) as an attempt to effectively measure the productivity of a researcher by considering all his/her cited papers, the articles' age, the number of co-authors of each cited paper, and the order of each co-author. Unlike the h-index, or other ones, such as G, H(2), W, AR, E, K, Hw , and J indices, that do not differentiate between co-authors in the same field, and ignore articles below a certain threshold or of high number of citations, the proposed k-index strives to map the research outcome of each co-author onto a more practical measure based on his/her collective research contribution, article's age, number and order of co-authors, and the total number of citations, thus, highlighting the complete effort of each co-author separately. The k-index utilizes a recursive geometric sequence that distributes the merits of each article, as fairly as possible, among all co-authors. The effectiveness of the proposed k-index is demonstrated against the well-known h-index by investigating the research outcomes of researchers in the same field that almost have similar h-indices. As expected, the aging factor of the proposed measure implies that if the number of citations of a researcher does not increase in time, his/her research index should decrease, and that is successfully reflected by the proposed k-index.*
Keywords: *K–index, h-index, DRPI, Research Performance Index.*

I. INTRODUCTION

THE famous ***h***-index that was suggested by Jorge E. Hirsh [1] in 2005 is a measure of the scientific output of a scholar. It is based on the most cited papers and on the number of citations they have received. It considers the distribution of citations of a researcher's publications [1, 2, 3, 4, 5, 6]. It simply means that out of ***N*** publications for a given researcher, there are ***h*** papers that have received ***h*** citation each, regardless of the high volume of citation a paper might have received, or the number of co-authors. *This means that a scholar with an index of* ***h*** *has published* ***h*** *papers each of which has been cited* ***h****-times in other papers* as shown in Figure 1. This reflects the number of articles and the number of citations per publications [1]. It is a simple measure that can be used to compare between scientists working in the same field. It also helps to improve the research output of each scientist.

The ***h***-index is expected to increase with the age of an active and productive researcher. It does not account, however, for the number of co-authors per publication, nor does it for the age of
the paper. In calculating the ***h***-index, each co-author receives the same number of citations regardless whether he/she is the first, second, or an $\boldsymbol{n}^{th}$-author, even if he/she has a minor contribution in a particular publication. It does not also differentiate between two scholars who have the same ***h***-index but have two different volumes of citations, nor does it measure the impact of two articles that received the same number of citations in two different time span (say in half the time of one another). It also ignores large number of articles that are cited ***h*** or even (***h-1***) times. These factors represent a drawback in the ***h***-index that need to be rectified. However, it has few advantages such as it relies on the number of citation of the papers, and not on the journal or its impact factor (this may suites less active researchers), it is not influenced by a single well-cited paper, nor does it increase by a large number of poorly cited papers [7,8,9].

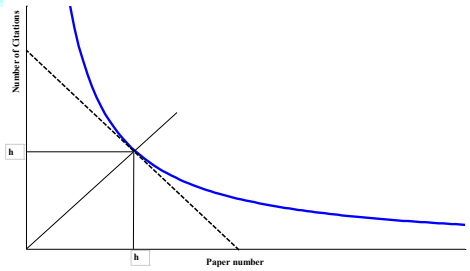

Fig. 1: A Schematic curve of number of citations versus paper number

Notice that the ***h***-index of a scholar means that his/her first ***h*** papers should receive an absolute minimum number of citations of h^2. However, if the number of citations of each first h papers sequentially increases by one, the expected minimum number of citations, $C_{h_{min}}$, can be calculated from the following arithmetic sum:

$$C_{h_{min}} = \sum_{k=h}^{2h-1}(k) = \sum_{k=0}^{h-1}(h+k) = \frac{h}{2}(3h-1) \quad (1)$$

However, if those ***h*** papers, for the same scholar, receive more than $h(3h\text{-}1)/2$ citations, his/her ***h***-index will not improve, in spite of his/her research being of interest to other scholars. This drawback in the ***h***-index will be corrected in the proposed ***k***-index.

In [5,6,7], a thorough comparative investigation was introduced to highlight the strengths and the differences between different types of indices such as the ***G***-, ***H***(2)-, ***HG***-, ***Q2*** -, ***AR***-, ***M***-quotient, ***M***-, ***W***-, ***Hw***-, ***E***-, ***A***-, ***R***- , and ***J***-indices. They all considered the ***h***-index as a base with some behavioral enhancement to overcome its limitation. When one considers averaging the ***h***-index by the number of years, or the square of the number of publications, as in the case of the ***g***-index, the tail publications become irrelevant in comparison to the core ones, similar to the case in the ***H***(2)-index. The total number of citations, the article's age, and the order of the co-authors, or even the total number of co-authors are not considered in a single index so that one can closely differentiate between researchers in order to reflect the true value of one's research contribution.

The ***h***-index does not also differentiate between different scholars of the same index but have different number of co-authors, e.g., the case of two scholars in the same field, where both have the same ***h***-index of 20, with one of them has a total of 25 co-authors, while the second one has a total of 75 co-authors, clearly shows the ineffectiveness of the ***h***-index as a measure to differentiate between researchers. This feature is also missing in the aforementioned indices for the same purpose.

II. *K*-INDEX CALCULATION

The ***k***-index is a dynamic measure that intends to effectively appraise both the productivity and the impact of a published work of a scholar. It accounts for all research outcomes of a scholar and considers the total number of citations of all research articles without placing any lower or upper thresholds. Whether there are high or low citations, the ***k***-index balances itself because it takes into account the research history of the scholar by considering the aging factor of each article.

It is difficult to quantify the contribution of co-authors of a single article because the process of ordering their names is not normalized throughout the globe. In order to come up with a fair distribution of the number of citations among co-authors, to see their individual effective research contributions, the distribution ought to be fair to give credits based on the order of co-authors. It is usually assumed that the first co-author receives more credit than the second one, and the second co-author is expected to earn more merit than the third one, and so on. Hence, the following scenario is proposed to generate a fair distribution of the number of citations of a research article, a book, or a book chapter among all co-authors. This scenario will be used in developing the proposed ***k***-index.

It is obvious that for a solo author, the whole number of citation of an article should fully be a credited to him/her. In the case of the ***h***-index, all co-authors share the whole merits regardless of their names' orders. Other scholars may argue to divide the merits of all co-authors equally likely among them, i.e., 50% each for two co-authors, and 33% each for three co-authors, and so on. Either method does not reflect reality since the first author is assumed to have put more effort in producing the research item than the second co-author, and the second co-author contributed more than the third one, and so on, otherwise, the order of the names will not have any meaning.

The proposed scenario assumes more merit to the first co-author than that of the later, while the second co-author should receive more credit than the third one, and so on. Hence, for an article with more than one co-author, the second co-author receives a merit 2/3 that of the first co-author and the third co-author receives a credit 2/3 that of the second co-author, and so on; i.e., each co-author receives a credit 2/3 of his/her predecessor.

Let the total number of citations of an article represents 100%. For an article with several co-authors, assume that the share of the first co-author be x of the total number of citations, then, according to the proposed scenario, the second co-author gets $(2/3)x$ of those citations, while the third co-author receives 2/3 that of the second co-author, *i.e.*, $(2/3)^2x$ of the first co-author, and so on. This will form a geometric sequence.

For n number of co-authors, the following relationship adds up to the total number of citations of each article as follows:

$$x + \frac{2}{3}x + \left(\frac{2}{3}\right)^2 x + \cdots \left(\frac{2}{3}\right)^{n-1} x = \sum_{i=1}^{n} r^{i-1} x = 100\% \quad (2)$$

where $r = 2/3$.

Evaluating the sum of the geometric sequence and solving for x yields:

$$x = \frac{1-r}{1-r^n} \times 100\ \% = \frac{1/3}{1-\left(\frac{2}{3}\right)^n} \times 100\ \% \qquad (3)$$

For a solo author, where $n = 1$, equation (3), as expected, yields x=100% of the total number of citations of an article. For two co-authors, however, it implies that the first and the second co-authors receive credits of 60% and 40% of the total number of citations, respectively, and for three co-authors, the credits is divided among the first, second, and third co-authors as 47.368 %, 31.579 %, and 21.0526 % of the total number of citations, respectively.

In general, let n_j be the number of co-authors of a j^{th}-article, then from equation (3), the merit of the i^{th}-co-author is denoted by $\rho_{ni,j}$, $1 \le i \le n$, and calculated from:

$$\rho_{n_j,i,j} = \left(\frac{2}{3}\right)^{i-1} \left(\frac{\frac{1}{3}}{1-\left(\frac{2}{3}\right)^n}\right) \times 100\ \%\ , \quad 1 \le j \le K \qquad (4)$$

where K is the total number of articles.

The percentage distribution of the number of citation of a j^{th}-article among several co-authors, up to seven co-authors, is defined by a merit matrix in (5) and summarized in Table 1. For example, the merit of a second co-author collaborated with two other researchers (i.e., a total of three co-authors) from a j^{th}-article; j, $1 \le j \le K$, is denoted by $\rho_{3,2,j}$ and equals to $\rho_{3,2,j} = 31.58\%$.

Notice that the size of M_j depends on the number of co-authors. If a j^{th}-article with three co-authors received 118 citations, then the citations distribution among all three authors will be $\rho_{31,j} \times 118 = 56$ citations for the first co-author, $\rho_{3,2,j} \times 118 = 37$ citations for the second one, and $\rho_{33,j} \times 118 = 25$ citations for the third co-author.

$$M_j = \begin{bmatrix} \rho_{11,j} & 0 & 0 & 0 & 0 & 0 & 0 \\ \rho_{21,j} & \rho_{22,j} & 0 & 0 & 0 & 0 & 0 \\ \rho_{31,j} & \rho_{32,j} & \rho_{33,j} & 0 & 0 & 0 & 0 \\ \rho_{41,j} & \rho_{42,j} & \rho_{43,j} & \rho_{44,j} & 0 & 0 & 0 \\ \rho_{51,j} & \rho_{52,j} & \rho_{53,} & \rho_{54,j} & \rho_{55,j} & 0 & 0 \\ \rho_{61,j} & \rho_{62,j} & \rho_{63,j} & \rho_{64,j} & \rho_{65,j} & \rho_{66,j} & 0 \\ \rho_{71,j} & \rho_{72,j} & \rho_{73,j} & \rho_{74,j} & \rho_{75,j} & \rho_{76,j} & \rho_{77,j} \end{bmatrix} \qquad (5)$$

Table 1: Citation distribution among co-authors

Merit / **No. of Authors**	$\rho_{n_j,1,j}$	$\rho_{n_j,2,j}$	$\rho_{n_j,3,j}$	$\rho_{n_j,4,j}$	$\rho_{n_j,5,j}$	$\rho_{n_j,6,j}$	$\rho_{n_j,7,j}$
1	100 %	0	0	0	0	0	0
2	60 %	40 %	0	0	0	0	0
3	47.37%	31.58 %	21.05 %	0	0	0	0
4	41.54%	27.69 %	18.46 %	12.31%	0	0	0
5	38.39%	25.59 %	17.06%	11.37 %	7.58 %	0	0
6	36.54%	24.36 %	16.24%	10.83 %	7.22 %	4.81%	0
7	35.41%	23.60 %	15.74 %	10.49 %	6.99 %	4.66 %	3.11%

Remark 1: *The number of citations for each co-author will be rounded to the nearest integer number. Therefore, as expected, the rounding process of the number of citations of an article with a few citations is always in favor of the first author.*

In order to calculate the ***k*-**index for an individual co-author, the distribution of all citations of each article among co-authors has to be executed first according to the matrix of merits, M_j**,** given by (5).

Supposed that an i^{th}-co-author has published K articles, each of c_j number of citations, and assume that the age of each article is denoted by T_j, where T_j represents the number of years since the article was published, then the

suggested ***k***-index is calculated from:

$$\boldsymbol{k} = \sum_{j=1}^{K} \frac{\rho_{n_j,i,j} \times c_j}{T_j} \qquad (6)$$

where $\rho_{n_j,i,j}$ is the merit of the i^{th}-co-author, $1 \leq i \leq n$, from the j^{th}-article; $1 \leq j \leq K$.

Remark 2: *If the age of a j^{th}-article, T_j; $1 \leq j \leq K$, is only few months since it was published, then T_j will be normalized by 12 months and will be taken as a fraction of a year in this case.*

The advantages of the proposed ***k***-index can be summarized as follows:

a) It accounts for all citations of a researcher, i.e., high or low volume of citations should not be discarded as in the case of the ***h***-index.
b) It accounts for the true individual research outcomes of a scholar by considering the total number of co-authors and their corresponding citation merits according to the order of their names order. Thus, in spite of encouraging group research, the research outcome of a solo-author should be appreciated as it happens.
c) It is a dynamic index since it introduces an aging factor for each article and consequently it reflects on the history of the scholar. Therefore, if the number of citations of a research item did not increase, the ***k***-index is expected to decrease. Consequently, if a scholar is keen to maintain at least the same ***k***-index, or higher, he/she must stay active in research and his/her research outcome ought to be of interest to other scholars.
d) It differentiates between individual scholars in the same field of research who have the same ***h***-index.
e) It minimizes (but not eliminates) the effect of a possible misconduct or unethical acts by certain individuals who may add their names to the list of co-authors without having a real contribution.
f) It accounts for both quantity and quality of research outcomes, especially for those "recent" articles that might get high volume of citation in very short time.

III. COMPARISON BETWEEN $\boldsymbol{k}$ AND $\boldsymbol{h}$ – INDEXES

The effectiveness of the ***k***-index is compared with that of the ***h***-index by investigating the research outcomes of two scholars. Tables 2 and 3 show a case study of calculating ***k***-indices for two researchers, A and B, in the same field of study. It is assumed that both scholars have published 20 papers each. The ***h***-index of scholar A is ***h*** =12 with a total number of 778 citations, while the ***h***-index of scholar B is ***h*** =14 with a total number of 537 citations. In these measures, the effect of the number of co-authors and the age of each article are not considered.

To calculate the ***k***-index, detailed information about each article is needed. Hence, the first column in the two tables shows the paper number, $1 \leq j \leq 20$, and the corresponding number of citations, c_j, for each paper is listed in descending order in the second column. The number of co-authors and the order of the concerned scholar are shown in the third and fourth columns, respectively. The merits of each paper for both scholars are then calculated using equation (4) and listed in the fifth columns, while the age of each paper is shown in the sixth columns. Applying equation (6) yields a cumulative ***k***-index as listed in the last columns, which grows with the increase of the number of citations. Clearly, the last number of the cumulative ***k***-indices is rounded to the nearest integer and represents the corresponding ***k***-index for both scholars.

Notice that the ***h***-index of scholar A, (***h*** = 12), will not change even if the first 12 papers receive hundreds of more citations, or even if all the papers from 13-20 papers receive up to 12 citations each. This represents a major drawback in the ***h***-index that is remedied by the proposed ***k***-index. Scholar B, however, has an ***h***-index of 14, which will not also improve unless the subsequent articles receive more citations.

Obviously, scholar B has higher ***h***-index than scholar A, which may leads to the conclusion that scholar B has better research outcome than that of scholar A. This is not actually the case. If one considers the number of co-authors, the order of each co-author with respect to their counterparts, and the age of each article, the ***k***-index of scholar A will be 31, while the ***k***-index of scholar B is 16.

As expected, the age factor of a research item will play an important role in indicating whether a scholar is active in research or not. If the number of citations does not increase with time, the ***k***-index is expected to decrease, which represents a more dynamic measure that differs from all other indices used for the same purposes.

Table 2: Calculation of ***k*-Index for scholar A**

No. of Paper J	No. of citation of co-author A c_j	No. of Co-Authors (n)	Order of co-author A (i)	Merit of co-author $\rho_{ni,j}$	Article's age (Years) T_j	k-Index
1	128	2	2	$\rho_{22,1} = 40\%$	10	5.12
2	118	3	2	$\rho_{32,2} = 31.58\%$	16	7.449
3	113	2	1	$\rho_{21,3} = 60\%$	22	10.53
4	68	1	1	$\rho_{11,4} = 100\%$	19	14.11
5	65	3	2	$\rho_{32,5} = 31.58\%$	13	15.69
6	59	4	2	$\rho_{42,6} = 27.69\%$	12	17.05
7	47	2	1	$\rho_{21,7} = 60\%$	25	18.18
8	27	3	2	$\rho_{32,8} = 31.58\%$	13	18.83
9	22	3	1	$\rho_{31,9} = 47.37\%$	13	19.64
10	21	3	1	$\rho_{31,10} = 47.37\%$	11	20.54
11	21	2	1	$\rho_{21,11} = 60\%$	24	21.1
12	19	1	1	$\rho_{11,12} = 100\%$	4	25.82
13	12	2	2	$\rho_{22,13} = 40\%$	14	26.16
14	12	2	1	$\rho_{21,14} = 60\%$	12	26.76
15	9	2	2	$\rho_{22,15} = 40\%$	26	26.90
16	8	1	1	$\rho_{11,16} = 100\%$	13	27.51
17	8	2	2	$\rho_{22,17} = 40\%$	13	27.76
18	8	2	1	$\rho_{21,18} = 60\%$	16	28.1
19	7	2	2	$\rho_{22,19} = 40\%$	17	28.22
20	6	1	1	$\rho_{11,20} = 100\%$	2	**31.22**
$J_T = 20$	$C_T = 778$	$n_{co} = 43$	$O_r = 1.45$	$\rho_{av} = 55.86$	$T_{av} = 11.346$	

k-index	**31**
h-index	**12**

Table 3: Calculations of ***k*-Index for scholar B**.

No. of Paper J	No. of citation of co-author B c_j	No. of Co-authors (n)	Order of co-author B (i)	Merit of co-author $\rho_{ni,j}$	Article's age T_j	k-Index
1	73	2	1	$\rho_{21,1} = 60\%$	24	1.83
2	65	4	2	$\rho_{42,2} = 27.69\%$	25	2.55
3	56	2	1	$\rho_{21,3} = 60\%$	24	3.95
4	52	3	1	$\rho_{31,4} = 47.37\%$	10	6.41
5	48	3	2	$\rho_{32,5} = 31.58\%$	13	7.57
6	39	4	2	$\rho_{42,6} = 27.69\%$	6	9.37
7	20	3	2	$\rho_{32,7} = 31.58\%$	12	9.90
8	19	3	2	$\rho_{32,8} = 31.58\%$	11	10.45
9	19	3	2	$\rho_{32,9} = 31.58\%$	14	10.87
10	18	3	2	$\rho_{32,10} = 31.58\%$	14	11.28
11	17	5	5	$\rho_{55,11} = 7.58\%$	6	11.50
12	15	2	1	$\rho_{21,12} = 60\%$	24	11.87
13	14	3	2	$\rho_{32,13} = 31.58\%$	12	12.2
14	14	2	1	$\rho_{21,14} = 60\%$	13	12.89
15	13	5	5	$\rho_{55,15} = 7.58\%$	4	13.13
16	13	1	1	$\rho_{11,16} = 7.58\%$	12	14.21
17	12	3	2	$\rho_{32,17} = 31.58\%$	11	14.56
18	11	3	2	$\rho_{32,18} = 31.58\%$	14	14.81
19	10	2	1	$\rho_{21,19} = 60\%$	12	15.31
20	9	3	3	$\rho_{33,20} = 21.05\%$	9	**15.52**
$J_T = 20$	$C_T = 537$	$n_{co} = 59$	$O_r = 2$	$\rho_{av} = 34.96$	$T_{av} = 10.8$	

k-index	**16**
h- index	**14**

Notice that even if scholar B has higher ***h***-index, scholar A has a better research impact than that of scholar B, as reflected by their corresponding ***k***-indices, which is further verified by the results shown in Table 4. Table 4 shows the effectiveness of their research outcomes by considering the average number of citations per paper, the average number of co-authors, and more importantly, the average order of the corresponding author with respect to their co-authors.

Obviously, with less number of co-authors, scholar A managed to have more citations than that of scholar B. In addition, the average order of scholar A is better than that of scholar B, which is reflected on the value of his (her) contribution. This distinctions between the two scholars could not be figured out by the ***h***-index. Hence, the proposed ***k***-index enjoys more parameters to consider that makes it a dynamic, adaptive, and more appealing measure than its counterparts.

Figure 2 summarizes the results of Tables 2 and 3 for both scholars. It shows the number of citations and the corresponding ***k***-indices of the two scholars versus the number of research items.

Table 4: Average research performance of scholars A and B

	Index Value	**No. of considered papers**	**Avg. no. of citation/paper**	**Avg. no. of co-authors/ paper**	**Avg. order of author**	**Article avg. age (year)**
A	$h = 12$	12	59	2.41	1.41	15.16
	$k = 31$	20	38.9	2.25	1.45	14.75
B	$h = 14$	14	33.5	3	1.85	14.85
	$k = 16$	20	26.85	2.95	2	13.5

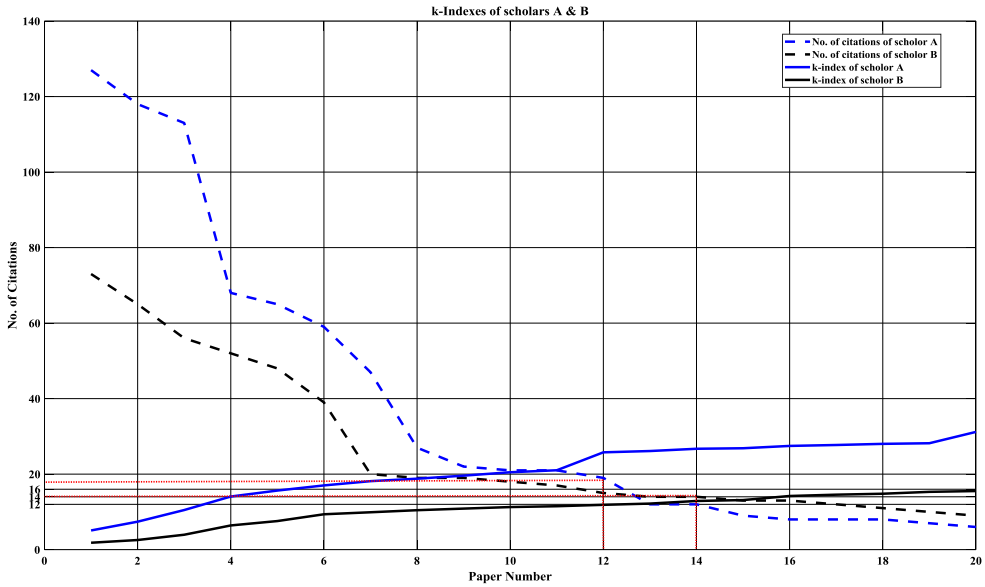

Fig. 2: Evolution of ***k***-Indexes vs number of papers of scholars A & B

IV. CONCLUSION

A new dynamic research indicator denoted by the ***k***-index is introduced to measure the research outcomes of scholars. Unlike the well-known ***h***-index, the ***k***-index considers all citations of all research articles of a scholar, the number of co-authors, the order of the corresponding scholar among his/her co-authors, and the age of each article. The proposed index involves four parameters, which makes it more dynamic than its counterparts. Having many co-authors is expected to yield more research outcomes that have to be fairly shared among scholars based on their name orders. This is not the case with the ***h***-index, where all co-authors equally receive the same merits. Whence a certain ***h***-index is achieved, it does not decrease with time. If there is no increase of new "selected" citations, the ***h***-index will remain constant and that eliminates an effective differentiating factors among scholars. Just like anything else, a research article ages with time, thus, the ***k***-index is introduced here to remedy this deficiency by considering an aging factor for each research article. The dynamic nature of the ***k***-index makes it more appealing to differentiate between scholars and defines a new measure to distinguish between them and indirectly minimizes ethical issues that might exist.

V. REFERENCES

[1] Hirsch JE (2005), "An index to quantify an individual's scientific research output," *Proc. National. Acad. Sci* , USA 102 (46), pp.16569–16572 (available online at: arXiv: physics/ 0508025, accessible via http://arxiv.org/abs/physics/ 0508025).

[2] Nils T. Hagen, "Harmonic publication and citation counting: sharing authorship credit equitably – not equally, geometrically or arithmetically," *Scientometrics* - Springer 2010 Sep; 84(3): 785–793.

[3] L. Egghe and R. Rousseau, "An informatics model for the Hirsch-index," *Scientometrics*, Vol. 69, No. 1 (2006), pp. 121-129.

[4] Kennedy D (2003), "Multiple authors, multiple problems," *Science,* 2003 Aug 8; 301 (5634):733. http://www.ncbi.nlm.nih.gov/pubmed/12907762

[5] Lutz Bornmann, and Hans-Dieter Daniel, "What do we know about the *h* index?" *J. of the Association for Inf. Sc. and Tech.*, vol. 58, No. 9, July, 2007.

[6] Lutz Bornmann, and Hans-Dieter Daniel, "Does the *h*-index for ranking of scientists really work?" *Scientometrics*, vol. 65, No. 3, pp 391-392, 2005.

[7] W. Glänzel, "On the *h*-index–A mathematical approach to a new measure of publication activity and citation impact. *Scientometrics*, 67 (2) 315–321, 2006.

[8] Ludo Waltman, and N. J. van Eck, "The inconsistency of the h-index," *J. of the American Society for Infor. Sc. and Tech*., Oct. 31, 2011, https://doi.org/10.1002/asi.21678.

[9] Alexander Yong, "Critique of Hirsch's Citation Index: A Combinatorial Fermi Problem," Notices of the American Mathematical Society, vol. 61 (2014), no. 11, pp. 1040–50.