RESEARCH ARTICLE

# The $h_a$-index: The average citation $h$-index

Yves Fassin

Department of Marketing, Innovation, and Organization, Faculty of Economics and Business Administration, Ghent University, Ghent, Belgium

**Keywords:** bibliometrics, citation, *h*-index, indicator, journal


## ABSTRACT

The ranking and categorizations of academic articles of a data set have traditionally been based on the distribution of their total citations. This ranking formed the basis for the definition of the *h*-index. As an alternative methodology, the ranking of articles of a data set can be performed according to the distribution of the average citations of the articles. Applying this same principle to the *h*-index itself leads to an average *h*-index, the $h_a$-index: the largest number of papers $h_a$ published by a researcher who has obtained at least $h_a$ citations per year on average. The new $h_a$-index offers more consistency, increased selectivity, and fairer treatment of younger scholars compared to the classic *h*-index. With its normalized time aspect, the method leads to better acknowledgment of progress. The evolution of the *h*-indices over time shows how the $h_a$-index reaches its full potential earlier and offers more stability over time. The average citation $h_a$-index partly solves the problem of the temporality of the *h*-index. The $h_a$-index can also be applied to academic journals. In particular, the application of the $h_a$-index to journals leads to more stability as they reach their limit sooner. The $h_a$-index offers a response to the inflation of *h*-index levels.




## 1. INTRODUCTION

The mathematical elegance and the arithmetic convenience of the *h*-index have been the major motivations for the widespread adoption of this indicator in the evaluation of science. It has become especially popular in applications concerning individual researchers. Defined as "the highest number of papers a scientist has that have each received at least that number of citations" (Ball, 2005; Hirsch, 2005), the *h*-index combines publication and citation counts, the most widely used proxies for productivity and impact. However, several bibliometric scholars have pointed to imperfections, and criticism has led to the proposal of alternatives that have not, however, succeeded in dethroning the *h*-index. In a recent ISSI newsletter, a new indicator was proposed (Fassin, 2020b) based on the ranking of articles of the data set according to the distribution of the average citations, rather than on the total citation distribution. The same principle as for the *h*-index is then applied, leading to a new $h_a$-index: the largest number of papers $h_a$ published by a researcher who have obtained at least $h_a$ citations per year on average. In this paper, a more in-depth analysis is realized that expands on the rationale behind this metric and positions it within existing *h*-index variants.

The structure of the paper is as follows. A literature study examines the criticism of the *h*-index and its proposed alternatives. From a systematization of the *h*-index variants, the new $h_a$-index is defined, followed by the practical application of the ranking of articles according to total citations and average citations. The $h_a$-index is then applied to individual

an open access journal

Check for updates

Citation: Fassin, Y. (2023). The $h_a$-index: The average citation $h$-index. *Quantitative Science Studies*. Advance publication. https://doi.org/10.1162/qss_a_00259

DOI: https://doi.org/10.1162/qss_a_00259

Peer Review: https://www.webofscience.com/api/gateway/wos/peer-review/10.1162/qss_a_00259

Supporting Information: https://doi.org/10.1162/qss_a_00259

Received: 17 July 2022
Accepted: 8 April 2023

Corresponding Author:
Yves Fassin
yves.fassin@ugent.be

Handling Editor:
Vincent Larivière



The MIT Press

researchers and to academic journals and compared to the existing indices. Its characteristics and its advantages are described. Particular attention is given to the evolution of the $h_a$- and *h*-indices over time, with applications.

## 2. LITERATURE STUDY OF THE *h*-INDEX

In a very short period, the *h*-index has been widely accepted as a measure of scientific performance and research achievement (Ball, 2005, 2007; Zhang, Thijs, & Glänzel, 2011). It is easily computable and provides a synthetic metric that combines the number of papers and their citations, the two traditional elements for evaluating researchers. Its simplicity allows comparison between scientists in a balanced way (Hirsch, 2005), even if other indicators are useful as complementary information and necessary for adding context (Wendl, 2007), as pointed out in the Leiden Manifesto for research metrics (Hicks, Wouters et al., 2015).

Strong criticisms have underlined the drawbacks and imperfections of the *h*-index (see e.g., Bouyssou & Marchant, 2011; Costas & Bordons, 2007; Waltman & van Eck, 2012; Wendl, 2007). Many bibliometric scholars advocate the use of multiple indicators and remain skeptical about the potential of any indicator to assess a scientist's work in one single metric. Even if it has become, in a very short time, the most popular of all bibliometric indicators, according to Bornmann and Leydesdorff (2018, p. 1,122), the *h*-index could be "the wrong type of summary statistics." Major imperfections concern problems of inadequate differentiation and selectivity and defective stability, due to their sharp and continuous increase over the years. A major drawback of the *h*-index is its time dependence. The *h*-index does not take into account any normalization of citation impact regarding the publication year or discipline.

As a consequence, the *h*-index engenders unfairness in assessment, as it has been privileging past achievements over recent contributions. In addition, different databases lead to different *h*-indices for the same researcher or group; *h*-indices should therefore be used with caution for formal academic purposes (Teixeira da Silva & Dobranszki, 2018).

### 2.1. Imperfections and Drawbacks of the *h*-Index

The *h*-index does not signal a contribution to seminal papers with extraordinarily high citation counts (Vinkler, 2010), nor does it fully reflect a scientist's accomplishment (Bras-Amorós, Domingo-Ferrer & Torra, 2011). An author who has published only one or several foundational papers is not rewarded by the *h*-index calculation, whereas modestly performing scientists with a larger number of moderately cited publications are unfairly favored (Dorogovtsev & Mendes, 2015).

The *h*-index values career achievement. It is subject to the principle of cumulative advantage, which in science means that scientists benefit from recognition from peers and prestige from published research (Allison, Long, & Krauze, 1982; Merton, 1968; Price, 1976; Wendl, 2007). By definition, the *h*-index cannot decrease (Hirsch, 2005). Older, more well-established researchers benefit from the lasting impact of publications from their early years and keep receiving new additional citations that can help to further increase their *h*-index (Rousseau, Egghe, & Guns, 2018). Moreover, papers that have earned sufficient citations in their first years but are not cited any more are still categorized in the researcher's *h*-index, even if they no longer have any impact. It is impossible for junior academics to reach that same level in a few years of academic research. In social sciences especially, it may take more than 5 years to accumulate a significant number of citations (Harzing, Alakangas, & Adams, 2014). Younger scholars have to build a portfolio of citable articles.

There has been a substantial increase in the number of academic publications and citations over recent decades. This growth has resulted in the exponential growth of database collections such as the Web of Science (Hu, Leydesdorff, & Rousseau, 2020). As a consequence, we have noticed inflation of the figures of the *h*-indices of all researchers and all journals. This phenomenon has diminished the significance and differentiation power of the *h*-index. Thresholds are rising, so it takes a few years for a new or recent paper to enter the existing *h*-core of a journal or a data set of a field. Once above a certain level of *h*-index, it is difficult for a junior scholar to achieve a further quick increase of the *h*-index; this requires different papers with substantial references, and the researcher's *h*-index is by definition limited to the number of published articles. Not only are citations important, but also the number of papers—so both elements that define the *h*-index constitute constraints for young researchers. The *h*-index is clearly a measure of a researcher's past accomplishments (Penner, Pan et al., 2013), not of their future achievements or potential. The *h*-index is, thus, not the most appropriate indicator for junior researchers (Harzing et al., 2014).

Redner (1998) observed that many papers obtain their citations over a limited period of popularity and then are no longer cited. The *h*-index tends to progress linearly during the most productive time of a researcher's career; it can then benefit from additional citations over a few years. Later, the increase in the *h*-index often results from citations to older publications close to the *h*-point (Schreiber, 2015). Then the *h*-index for authors tends to stabilize at the end of their productive careers. However, the *h*-indices of retired researchers do not decrease, even if they have little more impact. That inherent limitation of the size of the data set does not matter for academic journals or fields that can pursue a rise in their *h*-index.

This phenomenon of the exponential rise of the *h*-index of all researchers over the years (Hu, Leydesdorff, & Rousseau, 2020) makes the comparison of higher *h*-indices more difficult than at the time of its launch, where Hirsch suggested that an *h*-index of 20 after 20 years of research activity is a sign of success and exceptional *h*-indices of 40 outstanding (Ball, 2007). The same concerns for age arise for the *h*-index of journals. In the meantime, the *h*-index of the journals *Nature* and *Science* has exploded from 150 to 1,300 in less than 15 years (Braun, Glänzel, & Schubert, 2006). Econometric analysis tends to indicate that journal rankings reward older journals: The longer the existence of a journal, the more articles it publishes and the more likely a higher *h*-index becomes (Hudson, 2013; Ritzberger, 2008).

Another drawback of the *h*-index is the difficulty of comparing researchers from different disciplines (Batista, Campiteli et al., 2006). The *h*-index is "sensitive to differences in co-authorship patterns" between disciplines (Ryan, 2016, p. 578). The various fields adopt different publication practices with a varying number of publications and a different number of citations and, as a consequence, this results in a wide variation in the *h*-index.

### 2.2. Alternatives to the *h*-Index

The different critiques of the *h*-index have led to a huge number of alternatives. Bornmann, Mutz et al. (2011) performed a meta-analysis of studies on *h*-index variants; their comparative analysis, and especially the correlation between the *h*-index and 37 variants, concludes that most variants offer little added value over the *h*-index and are mostly redundant. A more recent comprehensive study on the *h*-index and its variants, realized by Bihari, Tripathi, and Deepak (2021), grouped the *h*-index variants into seven categories.

Some bibliometric scholars suggest counting the citations of the papers belonging to the *h*-core selection; others propose counting the top-cited papers in the corresponding fields, not only the *h*-core papers (Bornmann & Leydesdorff, 2018; Bras-Amorós et al., 2011). To better acknowledge exceptional contributions, Egghe suggested the *g*-index of a set of articles, defined as the highest rank *g* such that these *g* articles together received at least $g^2$ citations (Egghe, 2006a, 2006b). Jin, Liang et al. (2007) advance the *R*-index, as the square root of the total number of citations received by articles belonging to the *h*-core.

A number of variants of the *h*-index have been proposed, such as Kosmulski's $h^{(2)}$-index, equal to $h_2$ as the highest rank such that the first $h_2$ articles each received at least $(h_2)^2$ citations (Kosmulski, 2006). Analogously, an $h^3$-index has been defined (Fassin & Rousseau, 2019). The *w*-index indicates that a researcher has published *w* papers, with at least 10*w* citations each (Wu, 2010).



### 2.3. Normalization

Several proposals have been made to improve the fairness of the *h*-index. A number of variants of the *h*-index have suggested normalized alternatives that answer the distorting effects of the researcher's age or career length, multiple authorship, or scientific field.

#### *2.3.1. Career length*

To take into account the career length, Hirsch (2005) proposed an age-normalized variant of the *h*-index, namely the *m*-quotient, defined as $h/n$, where *n* defines the number of years since the researcher's first publication. As a complement to the *R*-index, Jin et al. (2007) put forward the *AR*-index, as the square root of the age-normalized total number of citations received by articles belonging to the *h*-core.

Schreiber (2015) presented the application of the *h*-index restricted to time windows. He described "the timed *h*-index $h_t$(5y) for publications from the year *y* and the previous *t* (5, or alternatively, 10) years in dependence to the length *t* of the utilized time window" (Schreiber, 2015, p. 150). However, in this time window approach, older major researchers who do not publish any more completely disappear from the classification.



#### *2.3.2. Multiple authorship and differences between fields*

A number of other proposals for alternatives or variants to the *h*-index have focused on the interdisciplinary character. To tackle these unequal characteristics between fields, bibliometricians have focused on the number of authors. Indeed, fields with a higher number of articles per researcher often work with larger (often international) teams and have more multiauthor papers (Batista et al., 2006). According to this reasoning, corrections for multiple authorship somewhat compensate for interdisciplinary differences. To reduce the effects of coauthorship, Batista et al. (2006) proposed the Individual *h*-index $h_I$ obtained by dividing the standard *h*-index by the average number of authors in the articles contributing to the *h*-index.

A more fine-grained alternative, the normalized individual *h*-index, $h_{I\text{norm}}$, was introduced by Harzing (2007) in the *Publish or Perish* program (PoP)[1]. The procedure of calculation for this $h_{I\text{norm}}$ is to first normalize the number of citations for each paper by dividing the number of citations by the number of authors for that paper, and only then to calculate the $h_{I\text{norm}}$ as the *h*-index of the *normalized* citation counts.

Multiple authorship has generated a huge debate between full-counting and fractional, reduced, or adapted counting formulas (Berker, 2018). Where full-counting privileges confirmed researchers who have published many papers with their network partners, fractional counting may be too penalizing for teams and especially for the leading authors and for the

[1] www.harzing.com.

corresponding authors. Several authors have, therefore, proposed different alternatives of adapted fractional counting with different methods of paper credit assignment (for a short overview, see Fassin (2020a)). The contribution of each author assigned with weighted factors can vary according to their role and position in the authorship list, but is not as diluted as with total fractional counting.

#### 2.3.3. Field normalization

Bornmann and Leydesdorff (2018) argue that *h*-index values are only comparable after proper field normalization. Several schemes for normalizing the citation counts have been proposed in recent years, based on percentiles (Leydesdorff & Bornmann, 2011), percentile rank classes (Leydesdorff, Bornmann et al., 2011), or variable *h*-type percentiles categories (Fassin, 2018). The HF-rating constitutes another attempt to cope with the interdisciplinary comparison (Fassin, 2020a).

To accommodate the issues of both disciplinary (multiple authorship) and career length, Harzing et al. (2014) proposed the individual, average annual increase of the *h*-index called $h_{I,\text{annual}}$, or $h_{I\text{a}}$. It is also calculated in the PoP program. It is defined as the $h_{In}$-index divided by the number of years of activity of the researcher. This criterion is also questionable; some bibliometric scholars propose the time that has elapsed between the researcher's first article, or since their PhD (Harzing et al., 2014).



## 3. SYSTEMATIZATION OF THE *h*-INDEX VARIANTS

In fact, to schematize the procedures for defining the *h*-index variants, we can distinguish three different operations: the ranking of the articles, the normalization or adjustments for age or career length, and the normalization of multiple authorship. The different *h*-index variants proposed depend on the order of calculation in the procedure.

The classic *h*-index executes the ranking only according to the article citations, without any normalization. The *m*-quotient corrects the *h*-index for age after ranking. Batista et al.'s (2006) $h_I$-index corrects for multiple authorship after ranking. Harzing et al.'s (2014) $h_{I\text{norm}}$-index first corrects for multiple authorship to normalize and then proceeds with the ranking (Ryan, 2016).

When drawing up a matrix of the variants on the basis of the order of calculation of the operations, in Table 1, one comes to a gap. Curiously, an alternative with a correction for age followed by the ranking and determination of an *h*-index variant has not yet been presented.

The ranking of a data set in the declining order of total citations formed the basis for the definition of the *h*-index. Now, the principle applied to define the *h*-index can also be applied to the average citation ranking. When ranking all papers of a researcher according to the average citation per paper, an average citation *h*-index can be defined: the $h_\text{a}$-index.

**Table 1.** Order of calculation

| | **Ranking first** | **Ranking second** |
|---|---|---|
| Multiple age correction | $m = {}^{h}/_{a}$ (Hirsch) | $\mathbf{h_a}$ |
| Multiple authorship correction | $h_I = {}^{h}/\text{avg } na$ (Batista et al.) | $h_{I\text{norm}}$ (Harzing et al.) |
| Double correction | $h_{Ia} = h_{In}/a$ (Harzing) | – |

Assuming $N$ publications $(p_k)_{k=1,...,N}$ with $c_k$ citations and age $a_k$ and $(na)_k$ coauthors for the $k$th publication:

*$h_a$ is the largest number of papers published by a researcher that have each obtained at least $h_a$ citations per year on average:* $\arg\max_{k\in\mathbb{N}}\left\{\frac{c_k}{a_k}\geq k\right\}$

$h$, $h_{\text{In}}$ and $h_{\text{a}}$ are obtained by ranking the citations or adapted ratios of citations. The units for ranking of the three $h$-indices differ. The $h$-index is obtained by ranking the citations of all publications, $h_{In}$ by ranking the number of citations divided by the number of authors for each publication taken individually, and $h_{\text{a}}$ by ranking the number of citations divided by the number of years since publication, also for each publication. This gives for $h$, $h_{In}$ and $h_{\text{a}}$ respectively:

$$c_k, \frac{c_k}{(na)_k}, \text{and } \frac{c_k}{a_k}$$

The order of the various publications will differ in the different orderings.

$m$ and $h_I$ are directly derived from the $h$-index.

$$m = {}^{h}/_{a} \qquad h_{Ia} = {}^{h_{In}}/_{a} \qquad hI = {}^{h_{In}}/_{\text{avg } na}$$

It is curious that this $h_{\text{a}}$ variant has not been proposed yet, probably because the $h$-index is an integer, but the average citations per paper are not integer numbers. In all honesty, we came across this finding not from the drawing up of the matrix, but indirectly from the study of the evolution of citations of publications over their lifetime. In the following section, we demonstrate how the average number of citations per year is an acceptable criterion for the normalization of the age effect.

We could round the $h_{\text{a}}$-index down to the integer below, as for the original $h$-index; however, the classic procedure of rounding up to the higher step if > 0.5 is preferable in this case[2]. Indeed, the Web of Science takes into account the current incomplete year in the calculation of the average citations per year, which penalizes papers during most of the year. The average can exceed the threshold by the end of the year, but then diminishes as the unit for the new year is added to the denominator.



## 4. THE TIME DEPENDENCE OF THE *h*-INDICES

Citations increase over time, and also the $h$-index and $h$-type indices rise over the years. The $h$-index is time dependent; it never declines and can at its maximum reach the number of articles published. In practice, for senior researchers, it rarely reaches more than 50% of the number of publications, as all authors have written articles that do not get citations. The proposed $h_{\text{a}}$-index rests upon the criterion of average citations per year as if the citations grow in a linear way. To what extent is this average citation per year representative of the citation pattern of academic publications? At an individual level too, citations of academic articles fluctuate significantly during the lifetime of an article and do not evolve in a linear pattern.

The evolution of a successful article's citations can be presented graphically in a simplified way, as a diachronous cumulative curve (Hu, Li, & Rousseau, 2020). The cumulative citation curve has the shape of an S-curve (Bejan & Lorente, 2012) with three successive periods of

[2] As an example, for five publications with average citations of 4.25, 3.5, 3.25, 3, and 2 the $h_{\text{a}}$-index will be 3 because $3 \leq 3.25 \leq 4$. With the rounding up, the $h_{\text{a}}$-index for five publications with average citations of 4.25, 3.75, 3.65, 3.5, and 2 the $h_{\text{a}}$ index will be 4 because $3 \leq 4$ (replacing 3.5) $\leq 4$.

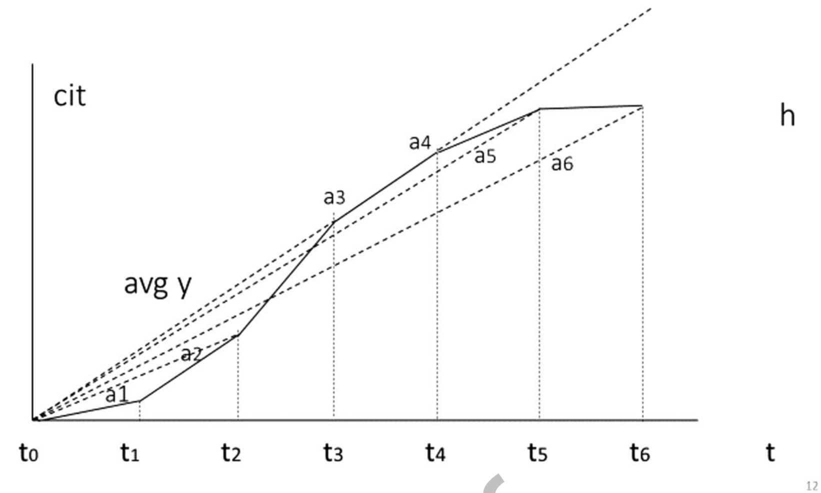


**Figure 1.** Cumulative citation curve and average citations.



increasing growth rate, to reach the average (avg *y*) after 15–20 years. The growth rate of citations $\frac{a\,(i+1)-a_i}{t\,(i+1)-t_i}$ will continue for an additional 5–10 years at a slower pace with less than half the average of the previous period, and will then gradually decrease further to remain stable or even decline to zero once the article is no longer cited.

In this simplified graph (Figure 1) of cumulative citations over time, the average is determined by the angle of the line drawn from the center of the axes to the point of the citation at a certain time. This average moves up from 0 to $a_1$ at $t_1$ and to $a_2$ at $t_2$; at $t_3$ it reaches its maximum average *ay*, where it stabilizes for a period $t_3$ until $t_4$ and then declines to $a_5$ at $t_5$ and to $a_6$ at $t_6$.

The average citation defined at a certain time varies; it is lower for the first years until the article reaches its peak. The assumption of the average number of citations leads to a temporary underestimation of the future potential of younger articles during the first years, but on an overall evaluation, as a moving average, it represents a good approximation, as the averages $a_3$, $a_4$, and $a_5$ stabilize and decline only slowly to $a_6$ after 25 years. Once that average gets below $h_a$, the article can fall out of the $h_a$-core, with a possible limited negative impact on the researcher's $h_a$-index.

This time evolution of the citations of a senior researcher's publications impacts the $h_a$-index, which will stabilize and ultimately slowly decline once the articles are no longer cited. In contrast with the *h*-index, the $h_a$-index is less time independent.

## 5. RESULTS: A PRACTICAL APPLICATION

To illustrate our proposal, we carry out two comparative analyses: the first at the level of citation distribution of a field, the second at the individual level.

### 5.1. Comparative Analysis at the Field Level

An example has been worked out on a homogeneous data set composed of all 67,052 articles in the field of entrepreneurship[3]. For this sample, the citation distribution tables and curves of

[3] Retrieved from the Web of Science (version Clarivate) by the query TS=entrepreneur*, field TOPIC (title, abstract, author key words, keywords plus), for all years (1956–2019).

academic articles are executed, first, on the basis of the total citations, and second, by applying the ranking on the basis of average citations per year. The two rankings show middling differences but higher variations for the most cited recent articles. Nearly 75% of the articles of the top 10% decile of the $h_a$-ranking also belong to the top 10% of the $h$-ranking. Some 70% of the top 1% percentile are common to both rankings, but 30% fall out of the 1% percentile in one of those rankings. The Spearman rank correlation of the 12,500 most cited entrepreneurship articles between their ranking according to the total citations count and their ranking according to the average citation per year reveals a moderate correlation of 0.634.

It should be noticed that the selection of the $h_a$-articles may differ from the selection according to rank in the $h$-index or $h^2$-index. The $h_a$-core can include articles that are not in the $h^2$-core and vice versa. The top of the ranking of articles also differs between both rankings (see Table 2). Although four of the five first-ranked articles in entrepreneurship research are the





**Table 2.** Comparison of the rankings of the $h$-core and $h_a$-core of entrepreneurship

| Authors | J | Year | TC | AVC | *r* avg | *r* cit |
|---|---|---|---|---|---|---|
| Teece | SMJ | 2007 | 3,439 | 245.6 | 1 | 3 |
| Shane & Venkataraman | AMR | 2000 | 4,621 | 220.1 | 2 | 1 |
| Uzzi | ASQ | 1997 | 4,461 | 185.9 | 3 | 2 |
| Mollick | JBV | 2014 | 945 | 135.0 | 4 | 52 |
| Lumpkin & Dess | AMR | 1996 | 3,315 | 132.6 | 5 | 4 |
| Zott, Amit, & Massa | JM | 2011 | 1,284 | 128.4 | 6 | 21 |
| Connelly et al. | JM | 2011 | 1,118 | 111.8 | 7 | 31 |
| Anderson et al. | JM | 2014 | 778 | 111.1 | 8 | 77 |
| Davidsson & Honig | JBV | 2003 | 1,741 | 96.7 | 9 | 12 |
| Greenwood et al. | AMA | 2011 | 951 | 95.1 | 10 | 51 |

| Authors | J | Year | TC | AVC | *r* avg | *r* cit |
|---|---|---|---|---|---|---|
| Shane & Venkataraman | AMR | 2000 | 4,621 | 220.1 | 2 | 1 |
| Uzzi | ASQ | 1997 | 4,461 | 185.9 | 3 | 2 |
| Teece | SMJ | 2007 | 3,439 | 245.6 | 1 | 3 |
| Lumpkin & Dess | AMR | 1996 | 3,315 | 132.6 | 5 | 4 |
| Harvey | GA | 1989 | 2,265 | 70.8 | 30 | 5 |
| Miller | MS | 1983 | 1,968 | 51.8 | 65 | 6 |
| Slater & Narver | JMK | 1995 | 1,878 | 72.2 | 27 | 7 |
| Shane | OS | 2000 | 1,868 | 89.0 | 13 | 8 |
| Desphandes et al. | JMK | 1993 | 1,815 | 64.8 | 39 | 9 |
| Sarasvathy | AMR | 2001 | 1,800 | 90.0 | 12 | 10 |

same in both rankings, only five other articles join the top 10 in both rankings. The six other most cited articles are in the top 65 of the best average citation, but they were all published before 2001. Four of the six remaining articles with the best average were published after 2010 and are placed 21, 31, 51, and 77 in the citation ranking. Lower ranked articles show more differences in ranking.

Well-cited articles climb gradually in the overall ranking, but it takes them 5–10 years to reach the top 10% decile or *h*-core, whereas they start in higher deciles in the average citation ranking. This average citation ranking allows a smoother transition, thanks to a unique overall temporal normalization. It compensates for the lower number of citations that newer publications can logically achieve.

In practice, the overall average citation distribution of any data set can be considered a good approximation; consequently, it can be put forward as a unique time-independent criterion. The overall average citation ranking can be proposed as a more accurate way to perform the ranking of a specific data set of articles.

### 5.2. A Comparative Analysis at the Individual Level

To illustrate the effect of the different variants of the *h*-index discussed in the scheme in Table 1, a comparative analysis has been performed.

A number of data are retrieved and their indices are calculated for some scholars in different scientific fields. Table 3 presents the WoS data of those scholars: number of papers in WoS − n, total citations in WoS − TC, the number of years the scholar has been active (number of years since their first publications) *y*, the average number of authors $a_a$; then their indices *h*, *m*, $h_I$, and $h_{I\text{norm}}$, $hI_a$, *AR*, *g*, *w*, and $h_a$ are computed on those WoS data. The asterisk * signals the presence of an article in $h_a$-core; the ratio $h_a/h$ is calculated in the last column. The sample consists of a dozen management scholars, a group of scholars in family firms and entrepreneurship research, members of the CYFE Center for Young and Family Enterprises at the University of Bergamo (with some of their coauthors), and the three leading scholars in entrepreneurship. They present a mix of established, well-known, midcareer, or younger scholars to offer a diversity of profiles. In addition, to study a comparison over different fields, the indices are calculated for a few bibliometric scholars, for a few top researchers in various fields such as medicine and plant sciences, and for physicist Ed Witten, the scientist with the highest *h*-index (Ball, 2005; Hirsch, 2005) and Jorge Hirsch for his impactful work in bibliometrics[4]. Just as the *h*-indices, so too do the $h_a$-indices depend on the size and citation patterns over the disciplines. The $h_a$-index of the field of entrepreneurship is 54; it is 30 for bibliometrics and exceeds 100 for medicine, physics, and plant sciences.

The last column in Table 3 presents the ratio between the $h_a$-index and the *h*-index, showing a great dispersion, even if there is a high correlation between both indices. Well-established authors have a lower ratio $h_a/h$ than midcareer or younger scholars, which underlines the selectivity of the $h_a$-index.

Table 4 displays the ranking of those 16 entrepreneurship scholars on the basis of various data and indices presented in Table 3. The following lines with the additional authors from

[4] With the increased multidisciplinarity of sciences, many scholars publish in different fields of research (for example Jorge Hirsch in physics and in bibliometrics). In such cases, three different *h*- and $h_a$-indices can be calculated: one in each discipline and one for their total oeuvre. This means that the general rule of cleaning the data first must be followed, depending on the objective of the evaluation.

**Table 3.** Indices of a selection of scholars (retrieved in September 2020)

| Author | F | $n$ | $TC$ | $y$ | $a_a$ | $h$ | $m$ | $h_I$ | $h_{Inorm}$ | $h_{Ia}$ | $AR$ | $g$ | $w$ | $h_a$ | | $h_a/h$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SZ | ENT | 125 | 24,147 | 29 | 2.42 | **66** | 2.28 | 27.2 | 52 | 1.8 | 37.0 | 125 | 26 | **19** | * | 0.29 |
| MW | ENT | 225 | 12,438 | 29 | 3.35 | **62** | 2.14 | 18.5 | 38 | 1.3 | 32.4 | 106 | 18 | **19** | | 0.31 |
| ADM | ENT | 74 | 2,812 | 13 | 3.69 | **29** | 2.23 | 7.9 | 16 | 1.2 | 19.5 | 52 | 8 | **13** | | 0.45 |
| SS | ENT | 82 | 15,626 | 41 | 1.85 | **46** | 1.12 | 24.9 | 37 | 1.1 | 28.3 | 82 | 17 | **12** | * | 0.26 |
| TZ | ENT | 33 | 2,497 | 14 | 3.00 | **22** | 1.57 | 7.3 | 15 | 1.1 | 17.3 | 33 | 11 | **11** | * | 0.50 |
| HA | ENT | 85 | 7,915 | 45 | 2.39 | **38** | 0.84 | 15.9 | 24 | 0.5 | 18.4 | 85 | 12 | **10** | | 0.26 |
| MN | ENT | 39 | 2,410 | 16 | 3.16 | **25** | 1.56 | 7.9 | 16 | 1.0 | 15.4 | 39 | 9 | **9** | | 0.36 |
| AVG | ENT | 16 | 823 | 15 | 3.38 | **13** | 0.87 | 3.8 | 10 | 0.7 | 9.4 | 16 | 6 | **7** | | 0.54 |
| MK | ENT | 38 | 867 | 8 | 2.82 | **17** | 2.13 | 6.0 | 12 | 1.5 | 9.9 | 29 | 4 | **6** | | 0.35 |
| TM | ENT | 25 | 347 | 13 | 3.75 | **12** | 0.92 | 3.2 | 5 | 0.4 | 8.3 | 18 | 3 | **5** | | 0.42 |
| GC | ENT | 19 | 406 | 8 | 3.82 | **11** | 1.38 | 2.9 | 5 | 0.6 | 8.8 | 19 | 4 | **5** | | 0.45 |
| LC | ENT | 13 | 408 | 16 | 3.44 | **9** | 0.56 | 2.6 | 6 | 0.4 | 7.9 | 13 | 4 | **5** | | 0.56 |
| FH | ENT | 27 | 1,074 | 39 | 2.77 | **13** | 0.33 | 4.7 | 10 | 0.3 | 7.0 | 27 | 5 | **4** | | 0.31 |
| DD | ENT | 8 | 96 | 6 | 3.17 | **6** | 1.00 | 1.9 | 4 | 0.7 | 4.5 | 8 | 2 | **3** | | 0.50 |
| DH | ENT | 7 | 33 | 4 | 3.67 | **3** | 0.75 | 0.8 | 2 | 0.5 | 3.1 | 5 | 1 | **2** | | 0.67 |
| MB | ENT | 3 | 35 | 5 | 4.33 | **3** | 0.60 | 0.7 | 2 | 0.4 | 3.9 | 3 | 1 | **2** | | 0.67 |
| LB | BIBL | 334 | 8,157 | 17 | 2.57 | **44** | 2.59 | 17.1 | 28 | 1.6 | 22.9 | 77 | 11 | **12** | * | 0.27 |
| JH | BIBL | 5 | 5,026 | 16 | 1.40 | **5** | 0.25 | 2.9 | 4 | 0.3 | 18.1 | 5 | 3 | **4** | * | 0.80 |
| ER | BIBL | 8 | 190 | 28 | 4.20 | **5** | 0.18 | 1.2 | 5 | 0.2 | 3.0 | 8 | 3 | **2** | | 0.40 |
| FP | BIBL | 24 | 84 | 18 | 3.33 | **6** | 0.33 | 1.8 | 3 | 0.2 | 2.2 | 8 | 1 | **1** | | 0.17 |
| EW | PHY | 361 | 102,280 | 43 | 1.88 | **144** | 3.35 | 76.8 | 120 | 2.8 | 57.9 | 315 | 57 | **32** | * | 0.22 |
| LM | MED | 420 | 31,566 | 54 | 7.09 | **79** | 1.46 | 11.1 | 30 | 0.6 | 27.6 | 172 | 23 | **11** | * | 0.14 |
| DI | PLS | 489 | 46,818 | 36 | 6.94 | **119** | 3.31 | 17.1 | 39 | 1.1 | 44.8 | 201 | 28 | **21** | * | 0.18 |
| HN | PLS | 37 | 1,201 | 18 | 11.10 | **21** | 1.17 | 1.9 | 7 | 0.4 | 12.5 | 34 | 6 | **7** | | 0.33 |

Auth: author's initials; F: field; $n$: number of papers in WoS; $TC$: total citations in WoS; $y$: the number of years the scholar has been active (since their first publications); $a_a$ average number of authors; $h$-index; $m$-quotient ($h/y$); $h_I$ and $h_{Inorm}$-indices; individual $h_{Ia}$-index; $AR$-index; $g$-index; w-index; the new $h_a$ (* for an article in $h_a$-core); the ratio $h_a/h$.

other disciplines give their respective rankings through (individual) interpolation in this short entrepreneurship list to emphasize comparison. Table 4 is completed with the HF-ratings of the scholars, divided into tiers of authors of comparable categories.

The differences between the absolute values of the $h$-indices change significantly over the various indices. Even if there is a high correlation between the $h$-index and the $h_a$-index, changes in the order can be noted when comparing the rankings, especially for the middle

**Table 4.** Ranking of a selection of scholars following various indices

| Author | F | *n* | *TC* | *h* | *m* | $h_I$ | $h_{Inorm}$ | $h_{Ia}$ | *AR* | *g* | *w* | $h_a$ | HF | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SZ | ENT | 2 | 1 | **1** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | **1** | AAA | * |
| MW | ENT | 1 | 3 | **2** | 3 | 3 | 2 | 3 | 2 | 2 | 2 | **1** | AA | |
| ADM | ENT | 5 | 5 | **5** | 2 | 6 | 6 | 4 | 4 | 5 | 7 | **3** | A | |
| SS | ENT | 4 | 2 | **3** | 8 | 2 | 3 | 5 | 3 | 4 | 3 | **4** | AAA | * |
| TZ | ENT | 8 | 6 | **7** | 5 | 7 | 7 | 6 | 6 | 7 | 5 | **5** | A | |
| HA | ENT | 3 | 4 | **4** | 12 | 4 | 4 | 11 | 5 | 3 | 4 | **6** | AAA | |
| MN | ENT | 6 | 7 | **6** | 6 | 5 | 5 | 7 | 7 | 6 | 6 | **7** | A | |
| AVG | ENT | 12 | 10 | **9** | 11 | 10 | 10 | 8 | 9 | 12 | 8 | **8** | BA | |
| MK | ENT | 7 | 9 | **8** | 4 | 8 | 8 | 2 | 8 | 8 | 10 | **9** | BBB | |
| TM | ENT | 10 | 13 | **11** | 10 | 11 | 12 | 14 | 11 | 11 | 13 | **10** | CCC | |
| GC | ENT | 11 | 12 | **12** | 7 | 12 | 13 | 10 | 10 | 10 | 11 | **11** | BBC | |
| LC | ENT | 13 | 11 | **13** | 15 | 13 | 11 | 15 | 12 | 13 | 12 | **12** | BBB | |
| FH | ENT | 9 | 8 | **10** | 16 | 9 | 9 | 16 | 13 | 9 | 9 | **13** | A | |
| DD | ENT | 14 | 14 | **14** | 9 | 14 | 14 | 9 | 14 | 14 | 14 | **14** | C | |
| DH | ENT | 15 | 16 | **16** | 13 | 15 | 15 | 12 | 16 | 15 | 15 | **15** | D | |
| MB | ENT | 16 | 15 | **15** | 14 | 16 | 16 | 13 | 15 | 16 | 16 | **16** | D | |
| | | | | | | | | | | | | | | |
| LB | BIBL | 1 | 4 | **4** | 1 | 4 | 4 | 2 | 4 | 5 | 5 | **5** | AAA | |
| JH | BIBL | 16 | 5 | **15** | 17 | 13 | 14 | 17 | 6 | 15 | 13 | **13** | AAA | * |
| ER | BIBL | 14 | 14 | **15** | 17 | 15 | 12 | 17 | 16 | 14 | 13 | **15** | BBB | |
| FB | BIBL | 11 | 15 | **14** | 16 | 15 | 15 | 17 | 16 | 14 | 15 | **17** | C | |
| EW | PHY | 1 | 1 | **1** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | **1** | AAA | * |
| LM | MED | 1 | 1 | **1** | 7 | 5 | 4 | 10 | 4 | 1 | 2 | **5** | AAA | * |
| DI | PLS | 1 | 1 | **1** | 1 | 4 | 3 | 5 | 1 | 1 | 1 | **1** | AAA | * |
| HN | PLS | 8 | 8 | **8** | 8 | 14 | 11 | 13 | 8 | 7 | 8 | **8** | BBB | |

Legend: see Table 3, and in addition: HF-rating based on full citation count.



category but also for some top-ranked authors. However, some of the absolute differences are smaller than their sole ranking may suggest. Although it is difficult to draw significant results from such a small sample, the analysis of the comparative rankings presents some indications.

The *h*-index ranking of most scholars lies in between the citation ranking and the productivity ranking. The *m*-quotient and the $h_{Ia}$-index offer the widest divergence from the *h*-index ranking. The *AR* ranking lies in between the *h*- and $h_a$-indices, except for authors with few papers but one exceptional highly cited paper, such as Hirsch in bibliometrics.

The $h_a$-index improves the position of midcareer researchers with 15–20 years of activity; they benefit from the increase in the number of publications that have accumulated sufficient citations, and from the increase in their average citations. Older scholars still benefit from the citations but not from continuous growth, which makes their $h_a$-indices stagnate.

In the *m*-quotient, the number of years of activity ($y$) plays an overarching role; it can seriously distort the index, especially if the most important articles are not published at the beginning of the author's career. A suggestion for refinement is to take a different time-space into consideration: the (average) time that has elapsed since the most cited publication(s) as single, first, or corresponding author. With increased precision, $h_a$ is calculated with the year average citation per publication and takes this asymmetry in time into account.

### 5.3. Multiple Authorship and Field Normalization

Likewise, the same objection of imprecision applies to the multiple authorship correction, where the $h_r$-index considers the average number of publications in the *h*-core. In some cases, a publication with many authors (even up to 20) ranked just within the *h*-core considerably reduces the $h_r$-index, whereas the more fine-grained method of the $h_{Inorm}$ would have replaced this publication with a single-authored article.

The application of corrections for both age and multiple authorship in the $h_{Ia}$-index also takes into account different career stages but suffers from the same problem related to the correct choice of the number of years.

The $h_{Inorm}$ variant and the $h_r$-indices have been presented as indices that allow interdisciplinary comparison, where the average number of authors is used as a proxy for different scientific fields. This assumption is based upon the observation that scholars in life sciences publish many more publications with large teams than social scientists who publish mostly as single authors or, nowadays, in small teams of two to four researchers.

By comparing entrepreneurship scholars with scholars of other fields, it has been shown that this generalization is not as correct as previously imagined. Although the top authors in physics and in plant sciences largely exceed the entrepreneurship scholars in the *h*-index ranking, the difference is attenuated in other rankings, especially in the $h_{Inorm}$- and $h_a$-indices. However, a top scholar in bibliometrics would fall into the subtop of entrepreneurship for several indices yet subtop researchers in life sciences may surpass the top entrepreneurship scholars. This comparison shows the importance of the differences in size, age, collaboration, and publication practices in different scientific fields and their evolution over time.

Normalization for multiple authorship remains a difficult exercise. For the interdisciplinary comparison, a bibliometric normalization performed directly on the total distribution of the field is to be preferred. The HF-rating is based on such a normalization (Fassin, 2020a). Where this method does not provide an exact ranking, it supplies tiers of authors of comparable categories. In this categorization, all top authors of the various disciplines achieve the same AAA rating, including Hirsch in bibliometrics for his impactful article, even though he has written a limited number of articles.

In the next section, we examine the evolution of the *h*- and $h_a$-indices over the years, presenting evidence of greater stability, one of the main advantages of the $h_a$-index.

## 6. THE EVOLUTION OF RESEARCHERS' $h_a$- AND *h*-INDEX OVER THE YEARS

Table 5 displays a few examples of the evolution of the $h_a$-index (left part of the table) and the *h*-index (right part) over the years 1960–2020 for researchers of different fields

**Table 5.** Examples of evolution of $h_a$-indices (left) and $h$-indices (right) over time

| Author | 1960 | 70 | 80 | 85 | 90 | 95 | 2000 | 5 | 10 | 15 | 2020 | 1960 | 70 | 80 | 85 | 90 | 95 | 2000 | 5 | 10 | 15 | 2020 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SZ | | | | | | 2 | 3 | 5 | 11 | 15 | 19 | | | | | | 6 | 15 | 18 | 35 | 53 | 66 |
| MW | | | | | | 1 | 1 | 5 | 10 | 14 | 19 | | | | | 2 | 3 | 7 | 14 | 30 | 55 | 62 |
| SS | | | | | | 1 | 3 | 7 | 11 | 13 | 14 | | | | | | 3 | 8 | 20 | 31 | 43 | 46 |
| TZ | | | | | | | | | 2 | 7 | 11 | | | | | | | | | 2 | 12 | 22 |
| TM | | | | | | | | | 1 | 2 | 5 | | | | | | | | | 1 | 4 | 12 |
| GC | | | | | | | | | | 2 | 5 | | | | | | | | | | 3 | 11 |
| LB | | | | | | | | 2 | 7 | 11 | 12 | | | | | | | | 2 | 13 | 31 | 44 |
| EW | | | 5 | 17 | 26 | 27 | 31 | 32 | 33 | 31 | 32 | | | 8 | 33 | 59 | 77 | 98 | 115 | 125 | 132 | 144 |
| PDG | 2 | 7 | 9 | 11 | 13 | 14 | 15 | 16 | 16 | 16 | 16 | 4 | 22 | 35 | 45 | 57 | 66 | 76 | 83 | 89 | 99 | 105 |
| PA | 3 | 9 | 13 | 14 | 17 | 21 | 20 | 19 | 20 | 19 | 19 | 10 | 27 | 45 | 54 | 64 | 75 | 80 | 82 | 83 | 86 | 87 |
| JH | | | 3 | 6 | 12 | 11 | 10 | 10 | 10 | 10 | 10 | | | 5 | 15 | 29 | 38 | 44 | 50 | 54 | 57 | 58 |

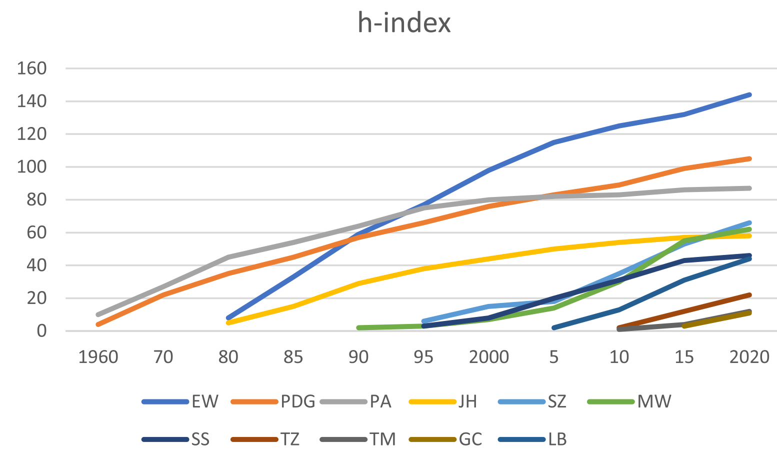


**Figure 2.** Evolution of the *h*-index over time.

(entrepreneurship, bibliometrics, and physics) and with different experience and age (most of the authors are those mentioned in Fassin (2020a)). Figures 2 and 3 present the graphs of those evolutions over time.

The first authors are the three major scholars in entrepreneurship research: Mike Wright (MW), Scott Shane (SS), and Shaker Zahra (SZ). The rise of their $h_a$-index from 2005 to 2020 amounts to between 50 and 70% of the rise of their *h*-index, after more than 20 year careers. A researcher who has been active for 15 years, Thomas Zellweger (TZ), has doubled his *h*-index in the last 5 years, and his $h_a$-index has increased by 50%. Younger authors Minola (TM) and Campopiano (GC), with about 10 years of research experience, are still expanding their *h*- and $h_a$-indices. In bibliometrics, Lutz Bornmann (LB) also stabilizes his $h_a$-index around 12 compared to his *h*-index of 44, reflecting the smaller size of the field. It is interesting to notice the evolution of the top physics researchers who were selected in *Nature*'s article when launching the *h*-index: Witten (EW), Anderson (PA), and DeGennes (PDG) (Ball, 2005, 2007). The peaks of their $h_a$-indices rose to respectively 33, 20, and 16 after 30–40 years of academic research. Since 2005, their $h_a$-indices have stabilized and can even decline, whereas their *h*-index has still risen by 20%.

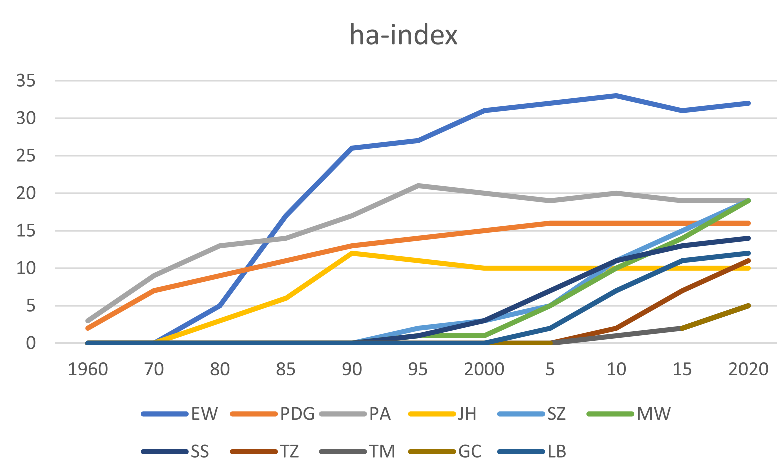


**Figure 3.** Evolution of the $h_a$-index over time.

The graphs perfectly illustrate the time dependence of the *h*-index and the stabilization that the $h_a$-index provides over time during the maturity phase. They also elucidate the differences between fields and between ages. For senior and retired researchers (the physicists), the *h*-curve is concavely increasing in function of time towards an asymptotic value, and the $h_a$-index shows a stabilization followed by a slight decline. Compared to the classic *h*-index, the $h_a$-index has a time limit and reaches its full potential at an earlier stage.

## 7. ADVANTAGES OF THE AVERAGE CITATION $h_a$-INDEX

As noted, the selection of the $h_a$-articles differs from the selection according to rank in the *h*-index or $h^2$-index. This $h_a$-index offers advantages. Compared to the classic *h*-index, the average $h_a$-index offers better selectivity and allows better differentiation, more stability, and a quicker acknowledgment of potential.

The $h_a$-index will help to better distinguish those articles that have an impact or that sustain a certain interest. Consider two scholars with equal *h*-indices of 10: One has accumulated 10 or more citations for his 10 publications 10 years ago, with the following citations distribution (50, 30, 20, 20, 15) and five articles with 10 citations; the second, a younger researcher, has a citations distribution (25, 20, 16, 12, 10) and five other articles with 10 citations published in the last three years. The first scholar's $h_a$-index will select only two articles; the second scholar's $h_a$-index will be 4 or 5 if two or three of the articles with 10 citations were published only 2 years ago.

The average $h_a$-index is lower in number and more selective than the high levels attained today by the classic *h*-index. The number of articles in the $h_a$-core lies more in the same range as the $h^2$-index, although, in general, they rank somewhat higher. Only those articles that sustain a higher growth rate over a longer period will remain in the $h_a$-core. An increase in the *h*-index demands two conditions: All papers in the *h*-core have at least one additional citation, which is usually the case, but also that one of the papers outside the *h*-core reaches the level of $h + 1$ citations, usually the next paper closest to the *h*-threshold or a more recent paper with high growth rate. An increase in the $h_a$-index requires more than one additional citation for the existing papers, and especially for the potential papers: If the difference is 1, it will need as many citations as the number of years since its publication; if the difference is only a fraction, it will necessitate the number of years, multiplied by the fraction of the difference. In this sense, the selectivity of the $h_a$-index is much wider, although not as wide as for the $h^2$-index, where the number of extra citations required for an increase of the $h^2$-index increases exponentially. An additional advantage relates to another critique of the *h*-index; this stricter criterion prevents manipulation of the $h_a$-index through self-citations.

In fact, the $h_a$-index allows the differentiation of articles according to their citation evolution pattern. Foundational papers continue to be quoted extensively and survive in the $h_a$-core and in the $h^2$-core. Older papers gradually decline over the years, but a basis of older papers persists: Those highly cited papers that have accumulated a sufficient reserve of citations for the next years. Older publications that continue to receive a reasonable number of citations maintain their place in the $h_a$-core, which means they still have an impact. Older papers with reasonably high cumulative citations retain their place in the *h*-core, but will slowly disappear from the $h_a$-core once their contributions lose impact. Older, lower cited articles cannot enter the $h_a$-core, even if they belong to the *h*-core. Recently published younger papers with potential can quickly enter the $h_a$-core, a few years before gathering sufficient citations to reach the threshold of the *h*-core and certainly many years before the $h^2$-core. The earlier detection of potential papers helps to acknowledge the visibility of younger researchers. It can also highlight important recent contributions of established researchers.

The $h_a$-index is more stable, as it does not increase as much as the *h*-index. Contrary to the *h*-index, which cannot decline, the $h_a$-index can stagnate and decline, for older scholars, when older publications around the $h_a$-index no longer receive many citations, whereas those articles around the *h*-core can enter the *h*-index in the long run with a few additional citations. Those articles around the $h_a$-core have to arouse sustainable interest. This new indicator combines career achievement with potential, recognizing experience and recent contributions. In fact, as some scholars contend, a combination of metrics can give more guidance. With respect to the $h_a$-index, the ratio between the $h_a$- and *h*-indices also provides valuable information.

The $h_a$-index does not solve all the drawbacks of the *h*-index; the major shortcoming of the *h*-index is that it does not signal the existence of an exceptional contribution. This condition is not fulfilled yet with the $h_a$-index. The *g*-index or the *AR*-index adds this information on excess citations to a certain extent. The alternative solution is to mark the presence of a paper in the $h_a$-core of the data set of the field under study with an asterisk (*) added to the researcher's $h_a$-index (see Table 3), following the suggestion for the HF-rating (Fassin, 2020a).

A characteristic of the $h_a$-index is its ease of calculation. As it is lower than the *h*-index, a preselection can be restricted to the *h*-core (or even, in many cases, the upper half of the *h*-core). The ranking of the *h*-core articles according to the average citation per paper provides a temporary provisional figure; the final $h_a$-core can be determined after an additional check with a limited number of already highly cited recent papers of the *h*-core that can be interpolated in that provisional ranking. Harzing's PoP program proposes an automatic selection choice for classification following the average citation per paper based on Google Scholar citations. Scopus and Web of Science, which are more selective in their selection of citations, do not provide this choice option yet.

In the following section, we extend the application of the $h_a$-index to academic journals.

## 8. THE AVERAGE CITATION $h_a$-INDEX FOR ACADEMIC JOURNALS

The same principle for the *h*-index can be applied to academic journals to define the journal's $h_a$-index. Data sets of journals group much larger numbers of articles than citation distributions of individual scientists. They are also more homogeneous, with smooth transitions, whereas data sets of individual researchers may present huge differences between their most cited papers. As a result, journals present a wider citation distribution with a smoother citation distribution curve. This makes the application and comparison of $h_a$-indices even more useful.

Applied to *Scientometrics* (SCIM), the $h_a$-index advances towards 22 by the end of 2020, whereas its *h*-index reaches 118 and the $h^2$-index 18.

In Table 6, the $h_a$-index of some journals in bibliometrics and in management are compared with their *h*-, $h^2$-, and $h^3$-indices, and with the average number of citations per paper. The journals display a variety of profiles and different levels of *h*-indices. The selection of management journals contains three of the top five management journals. The table also includes the data for the two leading journals *Science* and *Nature*, which exceed the level of 200 for their $h_a$-indices, compared to 1,300 for their *h*-indices.

The asymmetry of the composition of the $h^2$-core and $h_a$-core for journals is illustrated in the Supplementary material for an example of a journal data set. The $h_a$-core includes more recent articles than the $h^2$-core.

Table 7 exhibits the evolution of various *h*- and $h_a$-indices over time, from 1985 to 2020, and also yearly from 2015 to 2020, for three different journals with different profiles and

**Table 6.** Comparison of $h$- and $h_a$-indices of bibliometrics and management journals (Retrieved on January 13, 2021)

| | *n* | avg cit | *h* | $h^2$ | $h^3$ | $h_a$ |
|---|---|---|---|---|---|---|
| SCIM | 6,367 | 18.4 | 118 | 18 | 7 | 22 |
| JOI | 1,038 | 22.6 | 67 | 14 | 6 | 18 |
| JDOC | 3,555 | 8.9 | 73 | 16 | 7 | 12 |
| MJLIS | 282 | 4.4 | 15 | 5 | 3 | 3 |
| JASIS | 3,376 | 21.8 | 110 | 18 | 8 | 18 |
| JASST | 3,870 | 22.2 | 114 | 19 | 7 | 24 |
| JIS | 2,349 | 11.4 | 62 | 12 | 6 | 10 |
| JDIS | 71 | 3.7 | 8 | 3 | 2 | 3 |

| | *n* | avg cit | *h* | $h^2$ | $h^3$ | $h_a$ |
|---|---|---|---|---|---|---|
| AMJ | 3,249 | 156.2 | 363 | 36 | 12 | 54 |
| AMR | 2,458 | 182.4 | 330 | 39 | 15 | 59 |
| SMJ | 3,013 | 153.2 | 316 | 36 | 13 | 51 |
| ASQ | 2,949 | 105 | 284 | 36 | 13 | 45 |
| RP | 3,727 | 69.5 | 235 | 27 | 10 | 43 |
| JMS | 2,950 | 48.3 | 183 | 23 | 9 | 30 |
| JBE | 8,557 | 30.8 | 180 | 20 | 8 | 31 |
| | | | | | | |
| Science | > 150,000 | n/a | 1,311 | 75 | 20 | 200 |
| Nature | > 200,000 | n/a | 1,336 | 75 | 21 | 210 |

SCIM: *Scientometrics*; JOI: *Journal of Informetrics*; JDOC: *Journal of Documentation*; MJLIS: *Malaysian Journal of Library and Information Science*; JASIS: *Journal of the Association for Information Systems*; JASST: *Journal of the Association for Information Science and Technology*; JIS: *Journal of Information Science*: JDIS: *Journal of Data and Information Science Management*; AMJ: *Academy of Management Journal*; AMR: *Academy of Management Review*; SMJ: *Strategic Management Journal*; ASQ: *Administrative Science Quarterly*; RP: *Research Policy*; JMS: *Journal of Management Studies*; JBE: *Journal of Business Ethics*.





different levels of $h$-indices: the *Journal of Business Ethics*, the *Academy of Management Review* (the most cited journal in management research), and a leading journal in bibliometrics, *Scientometrics*, to analyze an example in a smaller discipline.

The same observation applies to the journals' $h_a$-indices, which reveal more stability than the $h$-index as their increase is slower; especially for higher indices, the increase by one unit per year of existence is a severe criterion. Although the $h_a$-index of journals will not decrease, they approach the asymptotic limit of the parabolic form of their evolution sooner. Younger journals, such as JBE and SCIM, have not yet reached this limit. JBE's $h$-index continues to rise by 14 units per year, somewhat less than 10%; its $h_a$-index rises by one unit per year on

**Table 7.** The evolution of the journals' $h_a$-, $h$-, $h^2$, and $h^3$-indices

| $h_a$ | 1985 | 1990 | 1995 | 2000 | 2005 | 2010 | 2015 | 2020 | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| JBE | | 2 | 3 | 4 | 4 | 10 | 16 | 31 | 16 | 19 | 21 | 25 | 28 | 31 |
| AMR | | 7 | 10 | 14 | 20 | 34 | 46 | 59 | 46 | 49 | 51 | 54 | 56 | 59 |
| SCIM | | | | 4 | 6 | 15 | 18 | 22 | 18 | 18 | 19 | 19 | 20 | 22 |

| $h$ | 1985 | 1990 | 1995 | 2000 | 2005 | 2010 | 2015 | 2020 | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| JBE | 2 | 8 | 17 | 28 | 38 | 66 | 105 | 180 | 105 | 118 | 133 | 145 | 159 | 180 |
| AMR | 7 | 25 | 49 | 77 | 118 | 185 | 262 | 330 | 261 | 274 | 289 | 300 | 317 | 330 |
| SCIM | | | | 28 | 36 | 60 | 88 | 118 | 88 | 92 | 96 | 102 | 108 | 118 |

| $h^2$ | 1985 | 1990 | 1995 | 2000 | 2005 | 2010 | 2015 | 2020 | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| JBE | 1 | 3 | 4 | 6 | 8 | 10 | 14 | 20 | 14 | 15 | 16 | 18 | 19 | 20 |
| AMR | 3 | 6 | 9 | 13 | 17 | 24 | 31 | 39 | 31 | 33 | 35 | 36 | 37 | 39 |
| SCIM | | | | 7 | 8 | 10 | 14 | 18 | 14 | 15 | 16 | 17 | 17 | 18 |

| $h^3$ | 1985 | 1990 | 1995 | 2000 | 2005 | 2010 | 2015 | 2020 | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| JBE | 1 | 2 | 3 | 3 | 4 | 5 | 6 | 8 | 6 | 7 | 7 | 7 | 7 | 8 |
| AMR | 2 | 3 | 4 | 5 | 7 | 10 | 13 | 15 | 13 | 13 | 13 | 14 | 14 | 15 |
| SCIM | | | | 3 | 4 | 5 | 6 | 7 | 6 | 6 | 6 | 7 | 7 | 7 |

JBE: *Journal of Business Ethics*; AMR: *Academy of Management Review*; SCIM: *Scientometrics*.



average. SCIM's $h$-index continues to rise by 5% a year, which is around five units a year. The rise of the $h_a$-index is somewhat slighter, but in absolute values, it increases only once every 2 years. AMR has reached a higher level of 37 for its $h_a$-index and, due also to its smaller size, it is closer to its natural limits, but its $h$-index is still progressing at a rate of around 5%.

## 9. LIMITATIONS AND FURTHER RESEARCH

First of all, it is worth repeating that metrics and indicators are not substitutes for a more comprehensive assessment of scientific contributions (Teixeira da Silva & Dobranszki, 2018); $h$-related indices depend on citations, which are reliant on the theme and the content of the article, and also on the size of the specific topic in the field. Metrics should be complemented with qualitative information and context as recommended by the Leiden Manifesto for research metrics (Hicks et al., 2015) and by the San Francisco Declaration on Research Assessment[5] (DORA).

[5] See sfdora.org.

The $h_a$-index has been presented as an integer number in the same way as the index values of the *h*-index are also restricted to integers. However, the numbers taken into account, namely average citations per year, are not integers. An alternative is to adopt a kind of interpolated version of the *h*-index (Rousseau, 2006). However, as the $h_a$-index is, just as the *h*-index, a Probably Approximately Correct (PAC) indicator (Rousseau, 2016), the $h_a$-index does not need the precision suggested by the interpolated version. Also, the choice of an average citation is an approximation, as the linear growth of citations does not hold exactly.

The application of the integer number can have consequences for young scholars in less popular research areas: When all their articles have average citations lower than 1, they would obtain an $h_a$ of 0, and an $h_a$ of 1 when their actual average citations are higher than 1 but lower than 2. This makes comparison and differentiation rather difficult. The suggestion is to use the figure with one decile for $h_a$-indices of lower than 3. For $h_a$-indices lower than 1, the $h_a$ will be defined by the best average citation per year of the researchers' articles limited to 1; for $h_a$-indices between 1 and 2, the $h_a$ will be defined with the second best average citation per year; or by the third best average citation per year when three articles have more than two citations per year.

Time remains a critical factor in bibliometrics. As the time elapsed is still a major determinant of the publication record and the number of citations, by analogy with the *h*-index (Schreiber, 2015; Schubert & Glänzel, 2007) it might be interesting to introduce a recent $h_a'$-index for the last five or 10 years, especially to compare the work of younger scholars.

The recent growth rate of an article during its first 5–15 years is higher than its average citation per year. Such a criterion increases selectivity. This means that articles selected to join the higher averages group certainly deserve their place in the $h_a$-core. Mention has also been made of another observation that can slightly affect the count: the imprecision in the calculation due to the inclusion of the current year[6].

Future work on the $h_a$-index should explore the theoretical foundations to support the practical advantages of the new indicator, following the theoretical work on the *h*-index and its alternatives by Bertoli-Barsotti and Lando (2015, 2017), Egghe and Rousseau (2006), and Glänzel (2006) and mathematical approaches on stochastics (Burrell, 2007).

Insights into the pattern of citation distribution merit further research. It would be advisable to conduct more empirical research in other sciences, such as in health and life sciences, with different citation patterns and higher immediacy in referencing than in management or social sciences. If the average citation in the first year is highest, recent papers could obtain a disproportionate advantage, although this would rapidly decrease. The use of the WoS average, which also counts the incomplete ongoing year in the denominator, could be a mitigating factor that could compensate for that advantage. Similar empirical research could be performed based on Google Scholar data with their higher levels of *h*-indices. Also, other *h*-related indices could be integrated into the comparative empirical analysis. Another avenue for further research could investigate in depth the evolution of the $h_a$- and *h*-indices over time in more and more varied scientific disciplines.





[6] The analysis of JBE would award an $h_a$-index of 28 by the end of 2019, and when using the average citations of the WoS in May 2020, with one (incomplete) additional year it drops to 26. This approximative nature is an additional argument to round the figure when using the data as extracted from the Web of Science. In this case it would award JBE an $h_a$-index of 27.

## 10. CONCLUSION

The present analysis of the comparison of the categorizations of academic articles based on total citations or average citations offers new insights into bibliometrics. The overall average citations categorization based on the complete data set seems to propose a better method, as it offers more consistency over time and provides a fairer evaluation than the overall total citations categorization. It gives a more precise view of the researcher's sustained impact and, as it provides greater acknowledgment of progress through the inclusion of more recent contributions, it is in favor of younger scholars. With its normalized time aspect, it mitigates other *h*-index variants and allows a better comparison—to a certain degree—between articles from different periods of publication. It can partially alleviate some of the unfair outcomes of the *h*-index as it distinguishes progress in an earlier phase. The average citation $h_a$-index partly solves the problem of the temporality of the *h*-index[7].

The comparison between variants of the *h*-index also questions the validity of the correction for multiple authorship as a proxy for interdisciplinary comparison and suggests that normalization be performed at the level of the total citation distribution of the field. It also suggests better alternatives of adapted fractional counting for multiauthorship.

With its increased selectivity and better differentiation, the proposed new $h_a$-index offers a valuable informative alternative to the *h*-index. It responds to the inflation of levels of *h*-indices. Furthermore, it brings us closer to the more comparable figures of the original *h*-index at the time of its launch. The $h_a$-index reaches its full potential earlier than the *h*-index. It offers stability over time, especially in the application of the $h_a$-index to larger data sets, such as those for academic journals, where the $h_a$-index reaches its limit sooner. It has the same ease of calculation as the *h*-index. This new indicator combines achievement with potential and recognizes experience as well as recent contributions.



## ACKNOWLEDGMENTS

The author thanks Professor Ronald Rousseau for his fruitful feedback and suggestions in the development of this paper, especially for the mathematical formulation of the $h_a$-index and for the observation for cases of lower $h_a$-index. He also thanks the editor and the anonymous reviewers for their suggestions.

## COMPETING INTERESTS

The author has no competing interests.

## FUNDING INFORMATION

The author received no funding.

## DATA AVAILABILITY

The indices are calculated from data of the Web of Science. For legal reasons, data from Clarivate's Web of Science cannot be made openly available.

[7] I owe this nice formulation to an anonymous reviewer, whom I wish to thank.

**REFERENCES**


Allison, P. D., Long, J. S., & Krauze, T. K. (1982). Cumulative advantage and inequality in science. *American Sociological Review*, *47*, 615–625. https://doi.org/10.2307/2095162

Ball, P. (2005). The index aims for fair ranking of scientists. *Nature*, *436*(7053), 900. https://doi.org/10.1038/436900a, PubMed: 16107806

Ball, P. (2007). Achievement index climbs the ranks. *Nature*, *448*(7155), 737. https://doi.org/10.1038/448737a, PubMed: 17700666

Batista, P. D., Campiteli, M. G., Kinouchi, O., & Martinez, A. S. (2006). Is it possible to compare researchers with different scientific interests? *Scientometrics*, *68*(1), 179–189. https://doi.org/10.1007/s11192-006-0090-4

Bejan, A., & Lorente, S. (2012). The physics of spreading ideas. *International Journal of Heat and Mass Transfer*, *55*(4), 802–807. https://doi.org/10.1016/j.ijheatmasstransfer.2011.10.029

Berker, Y. (2018). Golden ratio as a substitute for geometric and harmonic counting to determine multi-author publication credit. *Scientometrics*, *114*(3), 839–857. https://doi.org/10.1007/s11192-017-2632-3

Bertoli-Barsotti, L., & Lando, T. (2015). On a formula for the *h*-index. *Journal of Informetrics*, *9*(4), 762–776. https://doi.org/10.1016/j.joi.2015.07.004

Bertoli-Barsotti, L., & Lando, T. (2017). A theoretical model of the relationship between the *h*-index and other simple citation indicators. *Scientometrics*, *111*, 1415–1448. https://doi.org/10.1007/s11192-017-2351-9, PubMed: 28596626

Bihari, A., Tripathi, S., & Deepak, A. (2021). A review on *h*-index and its alternative indices. *Journal of Information Science*. https://doi.org/10.1177/01655515211014478

Bornmann, L., & Leydesdorff, L. (2018). Count highly-cited papers instead of papers with *h* citations: Use normalized citation counts and compare "like with like"! *Scientometrics*, *115*(2), 1119–1123. https://doi.org/10.1007/s11192-018-2682-1, PubMed: 29628536

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

Bouyssou, D., & Marchant, T. (2011). Ranking scientists and departments in a consistent manner. *Journal of the American Society for Information Science and Technology*, *62*(9), 1761–1769. https://doi.org/10.1002/asi.21544

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

Bras-Amorós, M., Domingo-Ferrer, J., & Torra, V. (2011). A bibliometric index based on the collaboration distance between cited and citing authors. *Journal of Informetrics*, *5*(2), 248–264. https://doi.org/10.1016/j.joi.2010.11.001

Burrell, Q. L. (2007). Hirsch's *h*-index: A stochastic model. *Journal of Informetrics*, *1*(1), 16–25. https://doi.org/10.1016/j.joi.2006.07.001

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*(3), 193–203. https://doi.org/10.1016/j.joi.2007.02.001

Dorogovtsev, S. N., & Mendes, J. F. (2015). Ranking scientists. *Nature Physics*, *11*(11), 882–883. https://doi.org/10.1038/nphys3533

Egghe, L. (2006a). An improvement of the *h*-index: The *g*-index. *ISSI Newsletter*, *2*(1), 8–9.

Egghe, L. (2006b). Theory and practice of the *g*-index. *Scientometrics*, *69*(1), 131–152. https://doi.org/10.1007/s11192-006-0144-7

Egghe, L., & Rousseau, R. (2006). An informetric model for the Hirsch-index. *Scientometrics*, *69*(1), 121–129. https://doi.org/10.1007/s11192-006-0143-8

Fassin, Y. (2018). A new qualitative rating system for scientific publications and a fame index for academics. *Journal of the Association for Information Science and Technology*, *69*(11), 1396–1399. https://doi.org/10.1002/asi.24059

Fassin, Y. (2020a). The HF-rating as a universal complement to the *h*-index. *Scientometrics*, *125*, 965–990. https://doi.org/10.1007/s11192-020-03611-5

Fassin, Y. (2020b). The average citation $h_a$-index. *ISSI Newsletter*, *16*(4), 64–67.

Fassin, Y., & Rousseau, R. (2019). The *h*(3)-index of academic journals. *Malaysian Journal of Library & Information Science*, *24*(2), 41–53. https://doi.org/10.22452/mjlis.vol24no2.3

Glänzel, W. (2006). On the *h*-index—A mathematical approach to a new measure of publication activity and citation impact. *Scientometrics*, *67*, 315–321. https://doi.org/10.1007/s11192-006-0102-4

Harzing, A. W. (2007). *Publish or perish*. Retrieved from https://www.harzing.com/.

Harzing, A., Alakangas, S., & Adams, D. (2014). $h_{Ia}$: An individual annual *h*-index to accommodate disciplinary and career length differences, *Scientometrics*, *99*(3), 811–821. https://doi.org/10.1007/s11192-013-1208-0

Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature*, *520*(7548), 429–431. https://doi.org/10.1038/520429a, PubMed: 25903611

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences USA*, *102*, 16569–16572. https://doi.org/10.1073/pnas.0507655102, PubMed: 16275915

Hudson, J. (2013). Ranking journals. *Economic Journal*, *123*(570), 202–222. https://doi.org/10.1111/ecoj.12064

Hu, X., Li, X., & Rousseau, R. (2020). Describing citations as a function of time. *Journal of Data and Information Science*, *5*(2), 1–12. https://doi.org/10.2478/jdis-2020-0011

Hu, X., Leydesdorff, L., & Rousseau, R. (2020). Exponential growth in the number of items in the WoS. *ISSI Newsletter*, *16*(2), 32–38.

Jin, B., Liang, L., Rousseau, R., & Egghe, L. (2007). The *R*- and *AR*-indices: Complementing the *h*-index. *Chinese Science Bulletin*, *52*(6), 855–863. https://doi.org/10.1007/s11434-007-0145-9

Kosmulski, M. (2006). A new Hirsch-type index saves time and works equally well as the original *h*-index. *ISSI Newsletter*, *2*(3), 4–6.

Leydesdorff, L., & Bornmann, L. (2011). Integrated impact indicators compared with impact factors: An alternative research design with policy implications. *Journal of the American Society for Information Science and Technology*, *62*(11), 2133–2146. https://doi.org/10.1002/asi.21609

Leydesdorff, L., Bornmann, L., Mutz, R., & Opthof, T. (2011). Turning the tables on citation analysis one more time: Principles for comparing sets of documents. *Journal of the American Society for Information Science and Technology*, *62*(7), 1370–1381. https://doi.org/10.1002/asi.21534

Merton, R. K. (1968). The Matthew effect in science: The reward and communication systems of science are considered. *Science*, *159*(3810), 56–63. https://doi.org/10.1126/science.159.3810.56, PubMed: 5634379

Price, D. de Solla. (1976). A general theory of bibliometric and other cumulative advantage processes. *Journal of the American Society for Information Science*, *27*(5), 292–306. https://doi.org/10.1002/asi.4630270505

Penner, O., Pan, R. K., Petersen, A. M., Kaski, K., & Fortunato, S. (2013). On the predictability of future impact in science. *Scientific Reports*, *3*(1), 3052. https://doi.org/10.1038/srep03052, PubMed: 24165898

Redner, S. (1998). How popular is your paper? An empirical study of the citation distribution. *European Physical Journal B—Condensed Matter and Complex Systems*, *4*(2), 131–134. https://doi.org/10.1007/s100510050359

Ritzberger, K. (2008). A ranking of journals in economics and related fields. *German Economic Review*, *9*(4), 402–430. https://doi.org/10.1111/j.1468-0475.2008.00446.x

Rousseau, R. (2006). Simple models and the corresponding *h*- and *g*-index. E-LIS: ID 6153. Available at https://hdl.handle.net/10760/7501.

Rousseau, R. (2016). Citation data as a proxy for quality or scientific influence are at best PAC (probably approximately correct). *Journal of the Association for Information Science and Technology*, *67*(12), 3092–3094. https://doi.org/10.1002/asi.23525

Rousseau, R., Egghe, L., & Guns, R. (2018). *Becoming metric-wise: A bibliometric guide for researchers*. Chandos Publishing.

Ryan, J. C. (2016). A validation of the individual annual *h*-index ($h_{Ia}$): Application of the $h_{Ia}$ to a qualitatively and quantitatively different sample. *Scientometrics*, *109*(1), 577–590. https://doi.org/10.1007/s11192-016-1972-8

Schreiber, M. (2015). Restricting the *h*-index to a publication and citation time window: A case study of a timed Hirsch index. *Journal of Informetrics*, *9*(1), 150–155. https://doi.org/10.1016/j.joi.2014.12.005

Schubert, A., & Glänzel, W. (2007). A systematic analysis of Hirsch-type indices for journals. *Journal of Informetrics*, *1*(3), 179–184. https://doi.org/10.1016/j.joi.2006.12.002

Teixeira da Silva, J. A., & Dobranszki, J. (2018). Multiple versions of the *h*-index: Cautionary use for formal academic purposes. *Scientometrics*, *115*, 1107–1113. https://doi.org/10.1007/s11192-018-2680-3

Vinkler, P. (2010). Indicators are the essence of scientometrics and bibliometrics. *Scientometrics*, *85*(3), 861–866. https://doi.org/10.1007/s11192-010-0159-y

Waltman, L., & van Eck, N. J. (2012). The inconsistency of the *h*-index. *Journal of the American Society for Information Science and Technology*, *63*(2), 406–415. https://doi.org/10.1002/asi.21678

Wendl, M. C. (2007). *H*-index: However ranked, citations need context. *Nature*, *449*, 403. https://doi.org/10.1038/449403b, PubMed: 17898746

Wu, Q. (2010). The *w*-index: A measure to assess scientific impact by focusing on widely cited papers. *Journal of the American Society for Information Science and Technology*, *61*(3), 609–614. https://doi.org/10.1002/asi.21276

Zhang, L., Thijs, B., & Glänzel, W. (2011). The diffusion of *H*-related literature. *Journal of Informetrics*, *5*(4), 583–593. https://doi.org/10.1016/j.joi.2011.05.004