# Preprints.org

Essay

# Goodbye H-index, Welcome C-score

Valenti Rull [*]

*Essay*

# Goodbye H-Index, Welcome C-Score

**Valentí Rull** [1,2]

[1]  Botanic Institute of Barcelona, Spanish National Research Council (CSIC), Pg. del Migdia s/n, 08038, Barcelona, Spain; vrull@csic.es

[2]  Institut Català de Paleontologia Miquel Crusafont (ICP-CERCA), Universitat Autònoma de Barcelona, c/ Columnes s/n, 08193 Cerdanyola del Vallès, Barcelona, Spain

**Abstract:** Using the h-index to assess individual researchers is not only unethical and unfair but also inaccurate and misleading. This index fails to offer a reliable measure of a researcher's impact based on their citation scores. Beyond statistical and conceptual considerations, authorship practices, especially in the case of multi-authored publications, give rise to significant problems that often go unnoticed. While some modifications of the h-index have been proposed to mitigate these weaknesses, the fundamental deficiencies persist. Most of these flaws have been effectively addressed by the c-score, a composite citation index. The c-score excludes self-citations, normalizes the number of citations by considering the number of authors in each paper, and takes into account first, single, and last authorship. This approach provides a more realistic measure of the impact of each individual researcher based on raw citations.

**Keywords:** h-index; c-score; individual evaluation; citation scores; first authorship; multiauthorship; scientific misconduct

## Introduction

How to utilize available bibliometric indices in the evaluation of individual researchers has been a long-standing topic of discussion in both general and specialized scientific literature. This goes beyond mere statistical analysis, as individual assessments have far-reaching implications for funding, hiring, promotion, recognition, and prizes, significantly influencing the career trajectory of scientists (Shen & Barabasi, 2014). The most commonly used indicators for evaluation are based on the impact of publications, gauged by the number of citations received. Despite its inherent limitations, the h-index remains the most widely used, likely due to its attempt to encapsulate all relevant information into a single numerical value.

As is well known, the h-index represents the intersection between the number of papers published by an author and the number of citations each paper receives, arranged in descending order (Hirsh, 2005). For instance, an h-index of 50 indicates that a particular author has 50 papers with 50 or more citations. It is important to note that the h-index can vary depending on the bibliographic database used. The lowest values are typically based on journals with an Impact Factor (IF) listed in Clarivate's Journal Citation Records (JCR), while the highest values are provided by Google Scholar, which considers the entire available literature. For example, the h-index of a randomly chosen researcher is 50 in the JCR and 65 in Google Scholar.

This commentary briefly summarizes the main shortcomings of the h-index and synthetizes how a recently introduced composite index, the c-score, addresses these drawbacks thus paving the way toward a fairer evaluation of individual researchers on the basis of their citation scores. Emphasis is placed on multiauthorship, a practice that has significantly grown in the last decades and has created unexpected measurement problems favoring biased strategies, and eventually, scientific misconduct (Koltun & Hafner, 2021).

**Weaknesses of the h-index**

Alonso et al. (2009) conducted an analysis of the h-index, scrutinizing its pros and cons from a statistical perspective. One notable drawback is its inapplicability for intra-discipline comparisons, owing to significant variations in productivity and citation practices. The h-index heavily relies on highly cited papers, but once these are consolidated at the top of the list, the number of citations they receive ceases to impact the index. Consequently, a researcher with a few extremely highly cited papers might share a similar index with a researcher boasting numerous moderately cited papers. Furthermore, the h-index fails to identify self-citations, allowing researchers and research teams to artificially inflate their scores through this practice (Bartneck & Kokkelmans. 2010). The index is also sensitive to the duration of a researcher's career, making it unsuitable for comparing scientists at different career stages (Stupnanova, 2024).

According to Gaster & Gaster (2012), the h-index measures quantity, rather than quality. Another drawback is that researchers having a lower h-index may have much more citations than others with higher h-index (Zhang. 2009). Additionally, the number of citations for a given paper is inflated as a multiple of the number of authors, as the same number of citations is counted on the web pages of all co-authors (Bi, 2023). Recognizing these weaknesses, there has been a proliferation of h-index modifications aimed at enhancing its performance as a measure of scientific impact (e.g., Egghe, 2008; Schreiber, 2008; Alonso, 2009; Zhang 2009; Negahdary et al. 2018; Koltun & Hafner 2021). A useful and thorough review of the h-index and the debate around it in simple non-technical terms was performed by Barnes (2017).

Another significant drawback is associated with authorship practices, which constitutes the focus of this paper. The recent surge in multiauthored papers (Wuchty et al., 2007) has posed challenges to the utility of the h-index, particularly concerning the fair allocation of credit among authors. The h-index, by not considering the number of authors in each paper and their respective contributions, gives rise to substantial issues (Bi, 2023). The primary flaw lies in the fact that every author of a multiauthored paper receives full credit for the publication, regardless of their actual contribution. This practice is not only ethically questionable but also unfair. Each author of a multiauthored paper is scored equally, akin to an author of a single-authored paper, where full credit is indisputable. Additionally, the h-index fails to account for gift authorship and, more broadly, fake authorship, making it impossible to discern whether authors contributed to the paper or if other contributors have been omitted (ghost authorship).

While some general rules have been recommended to enhance transparency in authorship procedures (McNutt et al., 2018), the responsibility ultimately lies with the authors. Inappropriate authorship practices have prompted several modifications to the h-index. For instance, Egghe (2008) and Schreiber (2008) introduced the fractional h-index, where the credit to each author in a multiauthored paper is determined by the number of citations divided by the number of authors. According to Kolun & Hafner (2021) such fractionation has been discussed in the literature but not implemented at scale. Negahdary et al. (2018) assigned progressively decreasing coefficients to the first/corresponding author (1), second 0.5), third (0.2) and fourth (0.1) authors. Alternatively, authors may explicitly state their percentage contribution to a given paper (Bi, 2023), but this approach remains highly subjective and susceptible to misconduct.

**H-Based Strategies**

Two types of authors whose contributions are indisputable are FAs and single authors (SAs). In recent decades, the practice of reserving the last position for project/group leaders and/or fund-raisers has become commonplace (although this was different in the past, creating inequities, especially for long careers). However, the criteria for identifying these LAs are subjective and may involve fake authorship practices, particularly gift authorship and coercive authorship, which are two forms of the so-called honorary authorship.

In multiauthored publications, it is not uncommon for the FA to be a young Ph.D. or postdoc researcher, the LA to be the team leader, and the co-authors in between (the IAs) to be members of the research team and other colleagues whose contributions are difficult to ascertain. In this scenario,

the FA is often called the junior author, and the LA is the senior author, typically serving as the advisor to the FA. However, this is a specific case of multiauthored papers that cannot be generalized. In this paper, the FA is referred to as the author, the IA as collaborators, and the LA as the manager.

It is also common for the FA and IAs from the same team, or a subset of them, to rotate in a collection of papers based on their contributions in each specific case, while the LA remains constant. This phenomenon has been referred to as a 'publication gang,' and it often leads to high rates of individual h-index inflation through cross self-citation and fake authorship. Another noteworthy scenario involves researchers who are supposedly engaged in a larger number of projects than they can genuinely handle. Consequently, they are listed as co-authors (rarely as FA or LA) on an unrealistically high number of publications, sometimes reaching one or more papers per week on average.

Beyond statistical and conceptual considerations, a more detailed empirical analysis of h-based citation patterns enables us to discern various approaches. To illustrate this, we present a real example based on the Google Scholar h-index of three researchers referred to as A, B, and C (Table 1). Although all three share a similar h-index range (64-69), an examination of authorship patterns within this range reveals three distinct strategies, whether intentional or not.

**Table 1.** Values of *h*-index (Google Scholar) for three real cases indicating the position of each author in the papers that define their *h*-value. FA, first author; SA, single author; LA, last author; IA, intermediate author. Retrieved 16 January 2024.

|   | Total | FA | SA | LA | IA | Strategy |
|---|---|---|---|---|---|---|
| **A** | 64 | 35 (55%) | 10 (16%) | 7 (11%) | 12 (19%) | Author |
| **B** | 69 | 15 (22%) | 1 (1%) | 5 (7%) | 48 (70%) | Collaborator |
| **C** | 65 | 3 (5%) | 0 (0%) | 42 (65%) | 20 (31%) | Manager |

Researcher A is the FA/SA in 71% of the papers that contribute to their index, signifying a heavy reliance on their own contributions. Conversely, researcher B's h-index is predominantly associated with an IA role, constituting 70% of their papers. In the case of researcher C, 65% of their h-index is closely tied to LA practices. Therefore, for researcher A, the h-index is substantially dependent on their individual contribution (author strategy), while for researchers B and C, the same index is primarily a result of others' contributions. In the case of IAs (collaborator strategy), this dependency is evident, but concerning LAs (manager strategy), it is important to note that significant contributions by the researcher cannot be disregarded, although such contributions are difficult to assess. However, the h-index assigns the same score to all three, potentially overlooking these nuanced distinctions.

Encouraging authors to disclose their contributions in multiauthored papers has not proven to be a solution for two main reasons. Firstly, journal requirements in this regard vary widely, making objective comparisons problematic. Secondly, there is no means of verifying what authors declare about their contributions, and there is a clear tendency to overestimate their own input (Herz et al., 2020). However, in publications where authors are arranged alphabetically, this becomes the only viable option. Fair credit allocation demands high integrity, which is difficult to gauge in many publications. This challenge is particularly pronounced in the case of hyperauthorship (more than 50 authors), a practice not uncommon today, with papers featuring over 1000 authors significantly increasing in the last decade (Sing, 2019).

Credit allocation in such papers is intricate and may lead to considerable unfairness, given that contributions such as providing reagents, is considered by the Contributor Role Taxonomy (CRediT; https://credit.niso.org/) as qualifying for authorship (Allen et al., 2018). Therefore, the h-index assigns the same score to an IA who provides a reagent to a research team as it does to a FA/SA who makes a substantial discovery. It should be stressed that CRediT guidelines are followed or recommended by the most influential journals and publishers.

**The c-score and the Stanford ranking**

In recent years, a new index addressing the limitations of the h-index has been developed. This composite citation index, known as the c-score, utilizes Elsevier's Scopus database and holds particular significance for single, first, and last authorship (Ioannidis et al., 2016). The c-score is also notable for excluding self-citations and employing the fractional modification of the h-index (Schreiber, 2008).

The c-score for the top 1000,000 scientists across disciplines is publicly released annually in two versions: the single-year version, reflecting the past year, and the career-long version, encompassing an entire researcher's career (Ioannidis et al., 2020). All this information is stored in the Mendeley database from 2019 onward (https://elsevier.digitalcommonsdata.com/datasets/btchxktzyw/6), and it includes partial rankings by specialty, thereby avoiding problematic interdisciplinary comparisons. The ranking based on the c-score is commonly known as the Stanford ranking, as it was developed and is maintained by researchers from that university.

The Stanford ranking presents a significantly different perspective on citation scores. For instance, Ioannidis et al. (2016) highlight that only 322 out of the top 1000 scientists based on total citations are included in the top 1000 using the c-score. This discrepancy mainly arises from the fact that many of the top-1000 authors based on total citations are not FAs or single authors SAs of their papers.

Referring to the example in Table 1, only researcher A is featured in the Stanford ranking. On the other hand, the c-scores of researchers B and C do not meet the criteria for inclusion, neither in the top 100,000 scientists in the general version nor in the top 2% of their respective specialties, which is an additional criterion for the Stanford ranking. It is noteworthy that a fourth researcher from the same specialty, referred to as D, with a significantly lower h-index (48) compared to A, B, and C, secures a higher position in the Stanford ranking than researcher A. Researcher D has a high percentage of FA/SA papers (75% of those included in the h-index), a few LA papers (20%), and very few IA papers (5%).

In summary, the c-score tends to favor the author and manager strategies while minimizing the highly opportunistic collaborator strategy (Table 1). Although the c-score is not a perfect indicator – achieving such perfection with a single number is challenging – it addresses most of the weaknesses associated with the flawed h-index. One potential critique could argue that the author strategy, the only one ensuring full involvement and a major contribution, should be prioritized over the manager strategy, where the real implication is difficult to ascertain. Another limitation is that this score cannot be applied in cases of alphabetically arranged authors.

**Summary and final remarks**

Using the h-index to assess individual researchers is not only unethical and unfair but also inaccurate and misleading. This index fails to offer a reliable measure of a researcher's impact based on their citation scores. Beyond statistical and conceptual considerations, authorship practices, especially in the case of multiauthored publications, give rise to significant problems that often go unnoticed. While some modifications of the h-index have been proposed to mitigate these weaknesses, the fundamental deficiencies persist.

Most of these flaws have been effectively addressed by the c-score, a composite citation index. The c-score excludes self-citations, normalizes the number of citations by considering the number of authors in each paper, and takes into account first, single, and last authorship. This approach provides a more realistic measure of the impact of each individual researcher based on raw citations.

As is well-known, citation-based impact is not a measure of scientific quality, a concept that is challenging to capture with a single numerical value. In the words of the creators of the c-score, 'Multiple indicators and their composite may give a more comprehensive picture of impact, although no citation indicator, whether single or composite, can be expected to select all the best scientists' (Ioannidis et al., 2016). However, the nature of impact measures significantly influences the strategies adopted by researchers to maximize their scores.

The use of the h-index, for instance, promotes self-citation and fake authorship, distorting the allocation of contributions by favoring opportunistic collaborator strategies and underrating first, single, and last authorship. In other words, the h-index may inadvertently stimulate undesirable authorship practices (Koltun & Hafner, 2021). While the c-score is not a panacea, it serves as a more objective measure of the citation-based impact of individual researchers and is less susceptible to opportunism and scientific misconduct.

Some issues that still need resolution include how to obtain a reliable measure for the contribution of each author and how to deal with nonsensical hyperproductivity and hyperauthorship. It is expected that finding a suitable solution for the former will also impact the latter by discouraging opportunistic authorship. While integrity and transparency are considered the best options, an objective measure independent from the author's criteria would be highly welcomed. It is noteworthy that most of these problems are linked to the increasing and fashionable trend of multiauthorship, which, while beneficial for scientific progress, poses significant challenges for measuring the citation-based impact of individual researchers. Other issues, such as incorrect or debunking citations, require a specialized reading of papers, making them unfeasible to address comprehensively.

Hopefully, these issues will be successfully tackled in the future. In the meantime, the c-score stands out as the best tool we have, while the h-index remains the worst and should be avoided. Regrettably, it is now beyond remedy to rectify the injustices brought about by this index.

## References

Allen L, O'Connel A, Kiermer V (2019) How can we ensure visibility and diversity in research contributions? How the Contributor Role Taxonomy (CRediT) is helping the shift from authorship to contributorship. *Learn Publ* 32: 71-74.

Alonso S, Cabrerizo FJ, Herrera-Viedma E, Herrera F (2009) h-index: A review focused in its variants, computation and standardization for different scientific fields. *J Informetr* 3: 273-289

Barnes C (2017) The h-index debate: an introduction for librarians. *J Acad Librariansh* 43: 487-494

Bartneck C, Kokkelmans S (2010) Detecting h-index manipulation through self-citation analysis. *Scientometrics* 87: 85-98

Bi HH (2023) Four problems of the *h*-index for assessing the research productivity and impact of individual authors. *Scientometrics* 128: 2677-2691

Chawla DS (2019) Hyperauthorship: global projects spark surge in thousan-author papers. *Nature*: doi 10.1038/d41586-019-03862-0

Egghe L (2008) Mathematical theory of the h- and g-index in case of fractional counting and authorship. *J Am Soc Info Sci Tech* 59: 1608-1616

Gaster N, Gaster M (2012) A critical assessment of the h-index. *Bioessays* 34: 830832

Herz N, Dan O, Censor N, Bar-Haim Y (2020) Authors overestimate their contribution to scientific work, demonstrating a strong bias. *Proc Natl Acad Sci USA* 117: 6282-6285

Hirsh FE (2005) An index to quantify an individual's scientific output. *Proc Natl Acad Sci USA* 102: 16569-16572

Ioannidis JPA, Klavans R, Boyak KW (2016) Multiple citation indicators and their composite across scientific disciplines. *PLoS Biol* 14: e1002501

Ioannidis JPA, Baas J, Klavans R, Boyack KW (2019) A standardized citation metrics author database annotated for scientific field. *PLoS Biol* 17: e3000384

Koltun V, Hafner D (2021) The h-index is no longer an effective correlate of scientific reputation. *PLoS ONE* 16: e0253397

McNutt MK, Bradford M, Drazen JM, Hanson B, Howard B, Jamieson KH, Kiermer V, Marcus E, Pope BK, Schekman R, Swaminathan S, Stang PJ, Verma IM (2018) Transparency in authors' contributions and responsibilities to promote integrity in scientific publication. *Proc Natl Acad Sci USA* 115: 2557-2560

Negahdary M, Jafarzadeh M, Rahimi G, Naziri M, Negahdary A (2018) The modified h-index of Scopus: a new way in fair scientometrics. *Pub Res Q* 34: 430-455

Schreiber MA (2008) A modification of the h-index: The hm-index account for multiauthored manuscripts. *J Informatics* 2: 211-216

Shen H-W, Barabasi A-L (2014) Collective credit allocation in science. *Proc Natl Acad Sci USA* 111: 12325-12330

Stupnanova A (2024) Author-level metrics dependent on time. *Fuzzy Sets Syst* 477: 108795

Wuchty S, Jones BF, Uzzi B (2007) The increasing dominance of teams in knowledge production. *Science* 316: 1036-1039

Zhang C-T (2009) The e-index, complementing the h-index for excess citations. *PLoS ONE* 4: e5429