Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

An index to quantify the disparity between coauthorship and individual's scientific impact

Shaibu Mohammed

shaibu.mohammed@uenr.edu.gh

University of Energy and Natural Resources https://orcid.org/0000-0001-7858-0103

Research Article

Keywords: ciations, h index, coauthorship disparity index, scientometrics,

Posted Date: July 8th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10283177/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

An index to quantify the disparity between coauthorship and individual's scientific impact

Abstract

Research collaboration advances the frontier of knowledge at a relatively fast pace. Regrettably, coauthorship—a consequence of research collaboration—has been abused. Coauthorship abuse stems from at least three episodes. First, the criteria for promotion in academia include the number of papers published. Second, as part of the requirements for positions in a university or research institutions, applicants should have a "decent" number of publications. Third, grant committees consider publications and other bibliometric indices. These have led to gift and guest coauthorships. In this paper, I propose an author-level metric, namely coauthorship disparity index (CDI), that compares an author's scientific impact and the number of coauthors on his/her h-core papers. The number of distinct coauthors of individual researchers were correlated with their number of publications, number of citations, and h indices. A positive trend was observed in all cases, with no apparent outliers. However, when the number of coauthors of the h-core papers was normalized by the h-index, outliers were detected. The CDI is able to decipher authors with and without outliers, thus quantifying the disparity between coauthorship and researcher's citation impact. Authors with high CDI's may want to reflect on their research output relative to the number of their coauthors.

Introduction

It has been argued that multiauthorship breeds, or is expected to breed, highly impactful publications compared to single-authorship (Fan et al., 2022; Glänzel and Thijs, 2004; Lawal 1986). A researcher needs not contribute to all sections, but a section or a few sections of a paper (data gathering and analysis, experimental setup, model development, statistical analysis, etc.) to qualify as an author. This has been reported to have a positive correlation with quality and quantity of research output (Larivieve et al., 2015). Therefore, research collaboration should be encouraged provided that every collaborator contributes significantly to an aspect of the research (Lin and Huang, 2012).

Truth be told, authors need publications for job applications, promotions, grant applications, and even for altruism (Hicks 2012). Therefore, they need their names not on one, but on several papers. This has, unfortunately, led to coauthorship misconduct. As an attempt to "fact check" guest and gift authors, some journals require that individual authors' contributions are stated during the submission stage. Regrettably, journal editors and reviewers cannot detect gift and guest authors from the stated contributions, and so these contributions are accepted anyway. The burden of integrity rests on the authors, but if they do otherwise, humanity suffers because grants and research positions would be given to underserving individuals. Inflating an author's list as a consequence of a publication cartel only increases one's publications; such individuals rise in the academic ranks faster than those who do not.

Several attempts have been made to award credits to coauthors based on the order in which their names appear on their publications (Schreiber, 2009, 2008a, b; Sidiropoulos et al., 2007). Other methods are based on authors' contributions to the research (Vavryčuk, 2018; Ioannidis, 2015; Stallings, et al., 2013; Tscharntke et al., 2007). On the one hand, coauthorship position is dependent

on the discipline; on the other hand, assigning credits based on authors' contributions requires that members of the committee read the contributions of every coauthor. This is a time-consuming endeavor. Sanches-Jimenez et al. (2026) conducted a comprehensive evaluation of researchers' output of over 300 categories. Their results suggest that while multiauthorship is on the rise, single authorship is on the decline. They argue that excessive numbers of coauthors will only inflate authors' lists while compromising on the accountability of the research content. The authors recommended the use of a discipline-specific threshold instead of a global one. In regard to using hyperauthorship as a proxy for research collaboration, Dinh et al. (2024) warned against the inclusion of hyperauthored papers to study coauthorship network.

Conventional metrics (h index and total citations) measure the productivity (publications) and impact (citations) of authors' output (Tzitzikas and Dovas, 2024; Tahamtan et al., 2016; Hirsch, 2005). Unfortunately, these metrics are unconcerned about single authorship and multiauthorship, and so all authors take full credit. As stated earlier, although several metrics have been proposed to address the abuse of coauthorship, the application of these metrics is time-consuming. In this paper, a new author-level metric is presented to measure the disparity between coauthorship and individual's scientific impact.

Coauthorship index

The h index and the number of citations of the h-core papers measure different dimensions of an author's output. Consequently, coupling these metrics seems a plausible idea compared to either of the two. On the one hand, the h index discards excess citations, which are clearly impactful output. On the other hand, the number of citations of all of an author's publications considers not only high-cited papers, but many low-cited papers too. A drawback of using total number of publications is that not all publications, and hence their citations, are captured owing to the coverage of the existing databases. Therefore, the actual numbers of publications and citations of most authors are not known. On the contrary, the total number of citations of h-core papers considers high-cited papers, which are almost certain to be captured in any given database; the low-cited papers are discarded. Thus, I coin a terminology, called impact number, which couples the h index and the total number of citations of the h-core papers. As will been shown in the results section, an author's impact number (I_h) can be described by:

$$I_h = \alpha h {C_h}^{\beta} \qquad (1)$$

where h is the h index, C_h is the total number of citations of the h-core papers, α and β are constants determined by a regression analysis. Then, an author-level metric, namely coauthorship disparity index (CDI), defined in Eq. 2, is proposed:

$$\text{CDI} = \frac{n_{k,h}}{I_h} \qquad (2)$$

where $n_{k,h}$ is the number of distinct coauthors of the h-core papers and I_h (computed from Eq. 1) is the impact number of the h-core papers of an author. Suppose that an author has a single authorship for all of his/her h-core papers, then $n_{k,h} = 0$. The coauthorship disparity index establishes a comparison between the number of coauthors and the citation impact of the h-core papers of an author.

Results

Here, the effect of the number of coauthors of the h-core papers ($n_{k,h}$) on h index and number of citations of the h-core papers is investigated. Authors' citation data (which we treat anonymously for an obvious reason) were taken from a published paper and others from Google scholar. The data covers diverse disciplines (Chemistry, Physics, Mathematics, Biological science, Biomedical engineering, and Petroleum engineering). The h indices considered in this study span 6 to 35. In all, the research output of 48 authors in diverse disciplines was used in this study.

The h indices of the authors were correlated with the number of citations (Fig 1a) and number of coauthors (Fig. 1b) of the h-core papers. Both give a positive correlation with no clear outliers. In Fig. 2, the number of coauthors of h-core papers was correlated with the number of citations of the authors' h core papers and number of publications. Although the correlation is not so good, a positive trend is observed with, again, no clear outliers. The trends in Figs 1 and 2 suggest that research collaborations generally translate to higher performance output. To reveal any outliers, the number of coauthors was normalized by the publication counts (Fig. 3a) and h-core papers (Fig. 3b) and plotted against the number of citations of the h-core papers. While a positive trend was exhibited in each case, clear outliers were observed (those in the circles in Fig. 3). This suggests that authors with varying numbers of coauthorship can have a comparable number of citations, and vice versa.

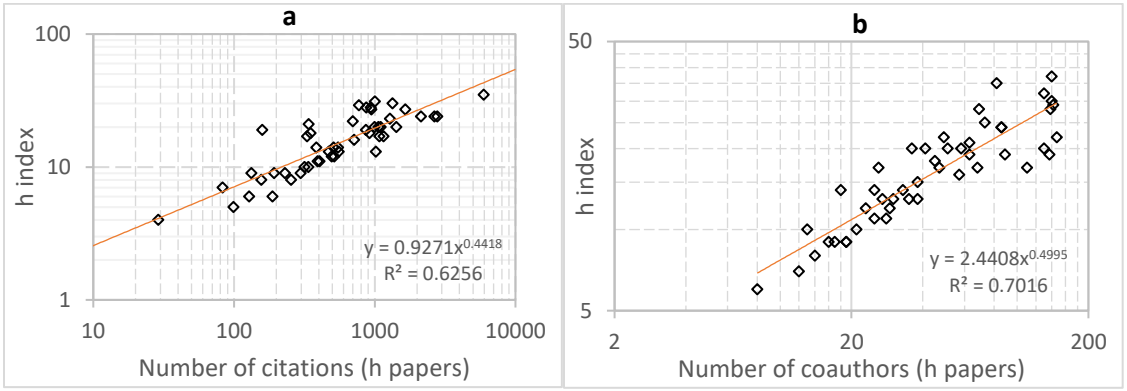


Fig. 1. Relation of h index with **a** Number of citations of h core papers. **b** Number of coauthors of h-core papers. In both cases, a positive correlation is observed with no clear outliers.

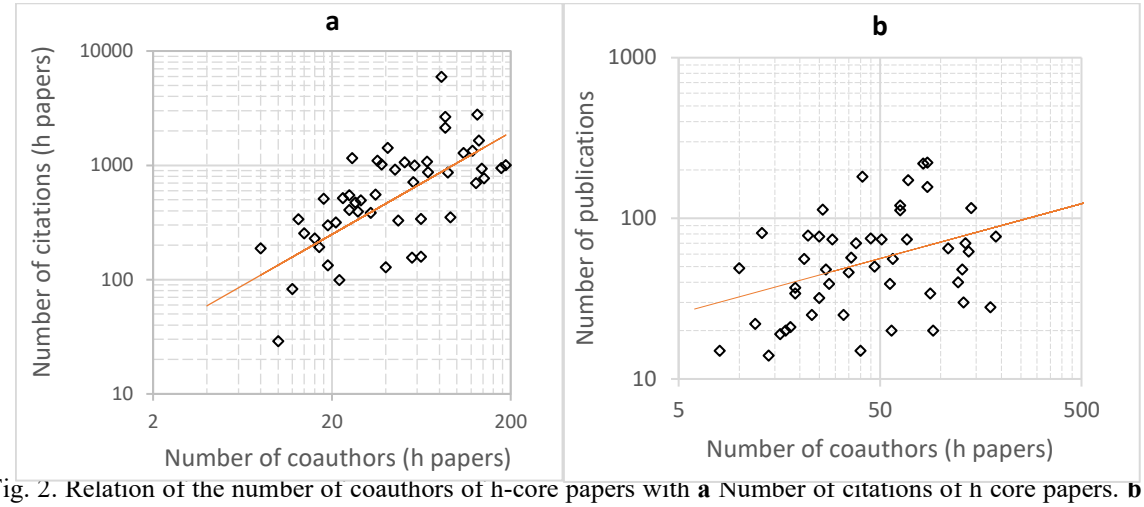


Fig. 2. Relation of the number of coauthors of h-core papers with **a** Number of citations of h core papers. **b** Number of publications. In both cases, a positive trend is observed with no clear outliers.

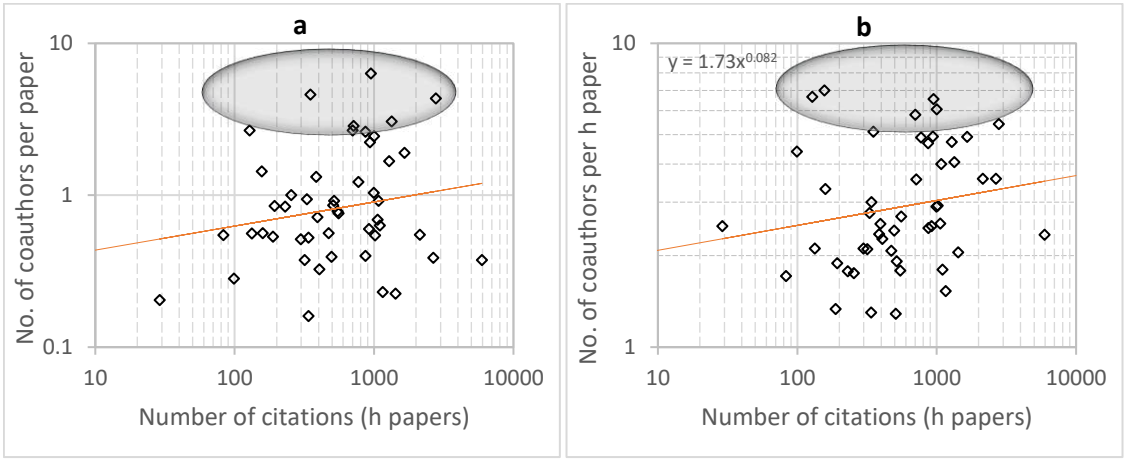


Fig. 3. Number of citations of h-core papers with **a** average number of coauthors per publications and **b** average number of coauthors per h-core papers. A positive trend is observed with clear outliers (in circles).

Let $n_{k,h}$ denote the number of coauthors of an author's h-core papers; then, the average number of coauthors per h-core paper is $n_{k,h}/h$. Relying on the trend in Fig. 3b, we can relate $n_{k,h}$ to h and C_h as:

$$n'_{k,h} = 1.73 h {C_h}^{0.082} \qquad (3)$$

where $n'_{k,h}$ is the number of coauthors of an author's h-core papers, given the author's h-index (h) and citations of the h-core papers (C_h). Thus, an author's impact number (see Eqs 1 and 3: i.e., $i_h = n'_{k,h}$) constitutes the number of coauthors an author should have. Then, the coauthorship disparity index (CDI), given previously in Eq. 2, can be re-expressed as:

$$\text{CDI} = \frac{n_{k,h}}{n'_{k,h}} \qquad (4)$$

where $n_{k,h}$ is the number of distinct coauthors of the h-core papers an author actually has and $n'_{k,h}$ (computed from Eq. 3) is the number of distinct coauthors of the h-core papers an author should have, given his/her citation impact.

I calculated the CDI's of over 500 authors in diverse disciplines. The results seem to suggest that authors with CDI's below 2 have had a good balance between the number of coauthors and citation impact. A high h index and CDI of less than 1 is truly an impressive score (Fig. 4). An author with a CDI greater than 2 may want to take a look at his/her coauthorship pattern relative to his/her citation impact.

When I derived Eq. 4, it dawned on me that the CDI is analogous to the Kardashian index (Hall 2014), which was developed to measure the discrepancy of scientists' research impact and twitter followers.

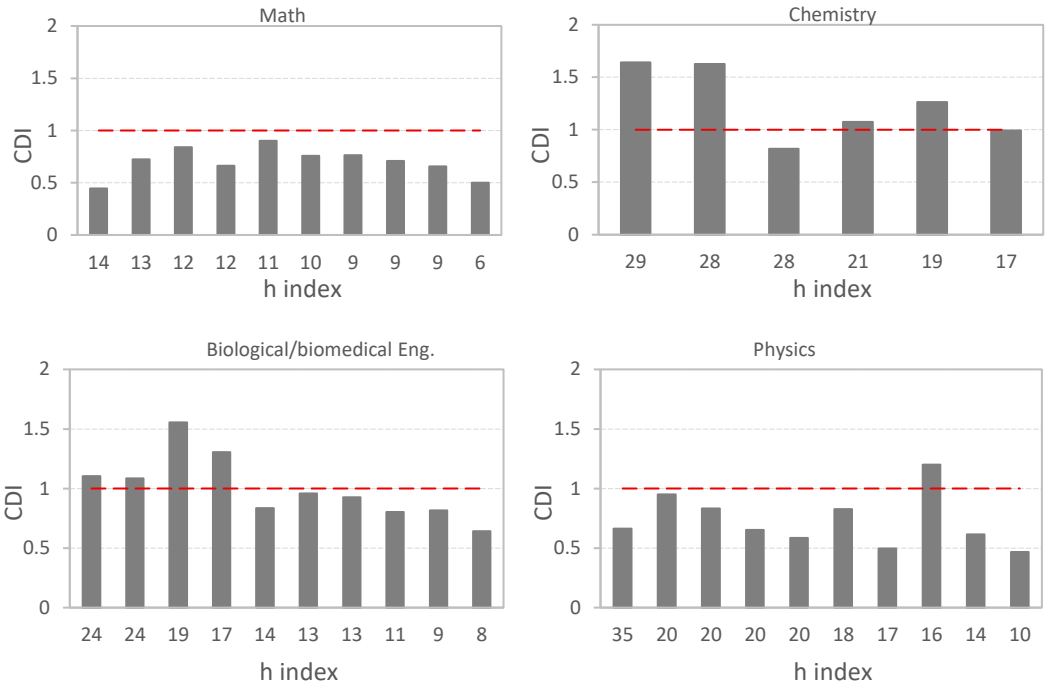


Fig. 4. Coauthorship disparity indices (CDI's) of authors in different disciplines.

Some Caveats

Although the results of this study suggest that authors with CDI greater than 2 have many coauthors relative to their impact, I need to point out that the metric, like all other metrics, has caveats. I do recognize the existence of sleeping beauties; I am also aware of early and mid-career researchers who are in a large group of senior collaborators; also, authors in different disciplines may have a marked difference in their CDI scores, and so should not be compared (Fig. 4). Once these caveats are borne in mind, the proposed metric would be useful to ascertain the influence of coauthorship on a researcher's citation counts and number of highly cited papers.

Conclusion

Authors who do the real work but gift or invite guest authors on their papers are not doing humanity any good. The coauthorship disparity index, presented in this paper, requires that for any given number of coauthors of the h-core papers, a minimum number of high quality papers and citations must be achieved. It's an attempt to "justify" the number of coauthors of an author's h core papers. Currently, the h indices of authors are provided in different databases; if $n_{k,h}$ and C_h are provided too, the calculation of CDI will be a lot easier. The findings of this study are summarized below.

There is a positive correlation between h index and number of citations; a positive correlation also exists between the number of coauthors and h index, number of citations, and impact number. These correlations suggest that research collaborations do, generally, produce high-impact output and should be encouraged. Despite these correlations, outliers were detected, suggesting that not all research collaborations result in high-impact output. If the limitations of the proposed metric are borne in mind, the coauthorship disparity index can serve as a useful measure to decipher the impact of the number of coauthors on research collaboration output.

Acknowledgment
The author expresses his gratitude to Seth Dwumah for reading and providing constructive feedback on the first draft.

Declaration
The author did not receive any funding for this study.

Reference

Bi, H. H. (2023). Four problems of the *h*-index for assessing the research productivity and impact of individual authors. *Scientometrics*, 128, 2677–2691. https://doi.org/10.1007/s11192-022-04323-8

Dinh, L., Barley,W. C., Johnson, L., & Allan, B. F. (2024). Hyperauthored papers disproportionately amplify important egocentric network metrics. *Quantitative Science Studies*, 5(3), 613–636. https://doi.org/10.1162/qss_a_00307

Fan, L. Guo, L., Wang, X., Xu, L., et al. (2022). Does the author's collaboration mode lead to papers' different citation impacts? An empirical analysis based on propensity score matching. *Journal of Informetrics*, 16(4), 1751–1577. https://doi.org/10.1016/j.joi.2022.101350

Glänzel, W., & Thijs, B. (2004). Does co-authorship inflate the share of self-citations? *Scientometrics*, *61*(3), 395–404. https://doi.org/10.1023/B:SCIE.0000045117.13348.bl.

Hall, N. (2014). The Kardashian index: a measure of discrepant social media profile for scientists. *Hall Genome Biology*, 15(424), 1–3. https://doi.org/10.1186/s13059–014–0424–0

Hicks, D. (2012). Performance-based university research funding system. *Research Policy*, *41*(2), 251–261.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences, 102*(46), 16569–16572. https://doi.org/10.1073/pnas.05076 55102.

Ioannidis, J. P. A. (2015). A generalized view of selfcitations, coauthor, collaborative, and coercive induced selfcitation. *Journal of Psychosomatic Research*, *78*(1), 7–11. article number e3000384. https://doi.org/10.1371/journal.pbio.3000384

Lin, W. Y. C. & Huang, M. H. (2012). The relation between co-authorship, currency of references and author self-citations. *Scientometrics*, *90*(2), 343–360. http://doi.org/10.1007/s11192-011-0523.

Larivieve, V., Gingras, Y., Sugimoto, C. R., & Tsou, A. (2015). Team size matters: Collaboration and scientific impact since 1900. Journal for the Association for Information Science and Technology, *66*(7), 1323–1332. http://doi.org/10.1002/asi.23266.

Lawal, S. M. (1986). Some bibliometric correlates of quality in scientific research. *Scientometrics*, 9, 13–25. https://doi.org/10.1007/BF02016604

Sanchez-Jimenez, R. Rodriguez, J. T. & Guerrero-Bote, V. P. (2026). Identification of hyperauthorship in scientific publications. An approach based on outlier detection methods. *Journal of Informetrics*. 20. hhtps://doi.org/10.1016/joi.202.101803.

Schreiber, M. (2009). The influence of self-citation corrections and the fractionalised

counting of multi-authored manuscripts on the Hirsch index. *Annalen der Physik, 521*(9), 607–621. https://doi. org/10.1002/andp.20095 210903

Schreiber, M. (2008a). A modification of the h-index: The hm-index accounts for multi-authored manuscripts. *Journal of Informetrics, 2*(3), 211–216. https:// doi. org/10. 1016/j. joi. 2008.05.001

Schreiber, M. (2008b). To share the fame in a fair way, hm modifies h for multi-authored manuscripts. *New Journal of Physics, 10*(4), 040201. https:// doi. org/10.1088/1367- 2630/10/4/040201

Sidiropoulos, A., Katsaros, D., & Manolopoulos, Y. (2007). Generalized Hirsch h-index for disclosing latent facts in citation networks. *Scientometrics,* 72, 253–280. https://doi. org/10.1007/s11192- 007- 1722-z.

Stallings, J., Vance, E., Yang, J., Vannier, M., W., Liang, J., Pang, L., Dai, L., Ye, I. & Wang, G. (2013). Determining scientific impact using a collaboration index, *Proc. Natl. Acad. Sci.* U.S.A. 110 (24) 9680-9685, https://doi.org/10.1073/pnas.1220184110 (2013*).*

Tahamtan, I., SafipourAfshar, A., & Ahamdzadeh, K. (2016). Factors affecting number of citations: A comprehensive review of the literature. *Scientometrics, 107*(3), 1195–1225. https://doi. org/10. 1007/s11192- 016- 1889-2.

Tscharntke, T., Hochberg, M. E., Rand, T. A., Resh, V. H., & Krauss, J. (2007). Author sequence and credit for contributions in multiauthored publications. *PLoS Biology, 5*(1), e18. https:// doi. org/ 10. 1371/ journal. pbio. 00500 18

Tzitzikas, Y., Dovas, G. (2024). How co-authorship affects the H-index? *Scientometrics* **129**, 4437–4469. https://doi.org/10.1007/s11192-024-05088-y

Vavryčuk, V. (2018). Fair ranking of researchers and research teams. *PLoS ONE, 13*(4), e0195509. https://doi. org/ 10. 1371/ journal. pone. 01955 09