TransInformação

DATA AND INFORMATION IN ONLINE ENVIRONMENTS

Editor
Carlos Luis González-Valiente

Support
Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) - CNPq/MCTI, nº26/2021. Process 402042/2022-0 and 200937/2022-7. Post-Doctorate Abroad, Protocol: 7763808095369450.

Conflict of interest
The authors declare they have no conflict of interests.

Received
Jan. 26, 2023

Final Version
Apr. 20, 2023

Approved
June 12, 2023

Journal self-citation on the h5-index of Ibero-American journals

Autocitação de periódico no índice-h5 de periódicos Ibero-Americanos

Fabio Lorensi do Canto[1], Adilson Luiz Pinto[1], Marcos Talau[2], Thiago Magela Rodrigues Dias[3]

[1] Universidade Federal de Santa Catarina, Centro de Ciências da Educação, Departamento de Ciência da Informação. Florianópolis, SC, Brasil. Correspondence to/*Correspondência para*: F. L. CANTO. *E-mail*: fabio.lc@ufsc.br

[2] Universidade Tecnológica Federal do Paraná, Centro de Educação. Curitiba, PR, Brasil.

[3] Centro Federal de Educação Tecnológica de Minas Gerais, Programa de Pós-Graduação em Modelagem Matemática e Computacional. Belo Horizonte, MG, Brasil.

Article based on the thesis of F. L Canto, entitled "*Avaliação de impacto de periódicos com base no índice h5 do Google Scholar Metrics*". Universidade Federal de Santa Catarina, 2022.

How to cite this article/*Como citar este artigo*: Canto, F. L. *et al*. Journal self-citation on h5-index of Google Scholar Metrics. *Transinformação*, v. 35, e237318, 2023. Doi: https://doi.org/10.1590/2318-0889202335e237318

Abstract

It aims to analyze journal self-citation in Ibero-American journals from the h5-index of the Google Scholar Metrics. The bibliometric tool *Gsm_hdata* was used to identify 4049 Ibero-American journals indexed simultaneously in Latindex and Google Scholar Metrics. Self-citations were identified, self-citation rates by country and research area were calculated, and the h5-index was recalculated without self-citations (hs5-index). No self-citations were identified in almost 40% of the journals, especially those with an h5-index lower than 5. The overall average self-citation rate was 3.6%. Among the 1859 most cited journals with at least one self-citation, the rate was 4.8%, lower than that of research based on the Impact Factor. Journals of Engineering, Exact and Natural Sciences, and Agricultural Sciences had the highest self-citation rates, while Social Sciences and Humanities journals presented the lowest. Journals with excessive rates (outliers) were identified in all areas. These results suggest that the prior exclusion of journal self-citations in the calculation of the h5-index is not necessary. However, monitoring journals with excessive self-citation rates is recommended to avoid distortions in impact assessment procedures based on h5-index of Google Scholar Metrics.

Keywords: Bibliometrics. Google Scholar Metrics. Ibero-American Journals. Journal self-citation.

Resumo

Este estudo analisa a autocitação de periódicos a partir do índice-h5. Foi utilizada a ferramenta bibliométrica Gsm_hdata para identificar 4049 periódicos ibérico-americanos indexados simultaneamente no Latindex e no Google Scholar Metrics. As autocitações foram identificadas, as taxas de autocitação por país e área de pesquisa foram calculadas e o índice-h5 foi recalculado sem autocitações (índice-hs5). Quase 40% dos periódicos não registraram autocitações, especialmente aqueles com índice-h5 menor que 5. A taxa média geral de autocitação foi de 3,6%. Entre os 1859 periódicos mais citados e com pelo menos uma autocitação, a taxa foi de 4,8%, taxa inferior à de pesquisas baseadas no Fator de Impacto. Periódicos de Engenharia, de Ciências Exatas e Naturais e de Ciências Agrárias apresentaram as maiores taxas de autocitação, enquanto

periódicos de Ciências Sociais e Humanas as mais baixas. Periódicos com taxas excessivas (outliers) foram identificados em todas as áreas. Estes resultados sugerem que a exclusão prévia de autocitação de periódicos no cálculo do h5-índice não é necessária. No entanto, é recomendado monitorar periódicos com taxas excessivas para evitar distorções nos procedimentos de avaliação de impacto baseados no h5-índice do Google Scholar Metrics.

Palavras-chave: *Bibliometria.* Google Scholar Metrics. *Periódicos Ibero-americanos. Autocitação de periódico.*

Introduction

Journal self-citation has received significant attention recently (Chorus; Waltman, 2016; Gazni; Didegah, 2021; Heneberg, 2016; Taşkin *et al.*, 2021; Wilhite; Fong, 2012). This interest may have originated from the perception that self-citations, and not the growth of impact, may be the main reason for the increase in indicators of some journals (Campanario, 2018; Heneberg, 2016). This alerted the need to identify self-citation patterns in databases and analyze journals with excessive self-citation rates (Gazni; Didegah, 2021; Taşkin *et al.*, 2021; Yu; Yu; Wang, 2014).

Identifying the average and growth rates of self-citations (Chorus; Waltman, 2016) makes it possible to draw a parameter in impact assessment processes. Normal rates reflect the source identity of the citing work and the work cited, which is natural when there is a thematic relationship. Normal rates do not produce significant changes to indicators or sudden changes in position in rankings, strata, or quartiles of databases (Flatt; Blasimme; Vayena, 2017; Yu; Yu; Wang, 2014). Journal self-citations are accepted in science when they perform their function, *i.e.*, provide information about works previously published in the same journal.

Journal self-citation rates considered abnormal should be analyzed and, if possible, justified. They may be more frequent in journals with low visibility or from specialized areas (Sanfilippo *et al.*, 2021). They may also originate from a specific document (an editorial or review article) that contains many self-citations. However, abnormal self-citation rates may also indicate editorial practices to boost impact indicators (Wilhite; Fong, 2012; Yu; Yu; Wang, 2014).

The Impact Factor (IF) and other impact indicators based on the average number of citations per article is sensitive to the increase in citations received (Gazni; Didegah, 2021; Liu; Fang, 2020; Waltman, 2016). Thus, journal self-citations have become a path to increase impact (Taşkin *et al.*, 2021). The IF version that excludes self-citations facilitates large-scale analyses, particularly identifying rates by research areas. Due to this ease of data extraction and analysis, most self-citation studies use data from the Journal Citation Reports (JCR) (Huang; Cathy Lin, 2012; Yu; Yu; Wang, 2014).

Although the h-index is not so sensitive to excess citations – including self-citations (Hirsch, 2005; Waltman, 2016) – the influence of journal self-citations on the h5-index result is not well known. Most studies based on the h-index have focused on author self-citations (Bartneck; Kokkelmans, 2011; Schreiber, 2009; Vîiu, 2016), the original unit of analysis of this indicator (Hirsch, 2005).

Google Scholar (GS) ignores the existence of self-citations in its two extensions dedicated to citation analysis: Google Scholar Citations (GSC) and Google Scholar Metrics (GSM). There is no resource for identifying self-citations, and the bibliometric indicators of these sources (i10-index, h5-index, h5-median) record citations and self-citations as a single element (López-Cózar; Cabezas-Clavijo, 2013; Jacsó, 2012).

In this sense, an analysis of self-citations of Ibero-American journals from GSM data. The self-citation rates were calculated by the country of origin of the journals and by area of knowledge.

The h5-index was recalculated after excluding self-citations. The journals were classified with normal and abnormal (outliers) self-citation rates.

The proposal is justified by the fact that GSM has been considered an alternative source of impact analysis, especially for journals not indexed in JCR and Scopus, which are the majority in the Ibero-American context (López-Cózar; Cabezas-Clavijo, 2013; Jacsó, 2012).

The results can help to answer the following questions: what are the average self-citation rates of Ibero-American journals from h5-index data? Do journals with high self-citation rates perform better on the h5-index than journals with normal rates? Is excluding self-citations in impact assessment processes necessary based on the h5-index?

Additionally, the findings can also help to enhance the understanding the influence of journal self-citations on the h-index, an alternative indicator to the traditional IF and CiteScore (CS) that has been gaining space in the scenario of impact assessment of Ibero-American scientific production.

Journal self-citation

Self-citation is defined as the practice of an author citing a previous work of their authorship or co-authorship (Ioannidis, 2015; Heneberg, 2016; Waltman, 2016; Szomszor; Pendlebury; Adams, 2020; Taşkin *et al.*, 2021). However, self-citation may be configured in other ways, such as at the collaboration level by citing research colleagues, at the journal level by citing articles from the same journal (Rousseau, 1999; Hartley, 2009), at the institutional level by citing a publication from the same institution/publisher or scientific peers (Zhou, 2021), at the country level by citing a publication from the same country (Huang; Cathy Lin, 2012; Frandsen, 2007; Waltman, 2016), and so on.

There is no unanimity regarding excluding self-citations in impact indicators (Campanario, 2018; Huang; Cathy Lin, 2012). There is a consensus that at the micro (author, journal) and meso (institution) levels, the influence of self-citation on the impact analysis is significant (Yu; Yu; Wang, 2014), and exclusion may be necessary. In turn, at the macro level (country), the influence is not significant, and the exclusion is unnecessary (Waltman, 2016). However, it is unanimous that the existence of self-citations in scientific communication cannot be ignored (Frandsen, 2007).

Unlike author self-citation, in which there is an unequivocal intention to self-cite (whether justifiable or not), journal self-citation may occur involuntarily (Hartley, 2009). It is not uncommon for several studies on the same topic to be published in the same journal. Thus, the citation of part of these studies in new work in the same publication is scientifically valid (Frandsen, 2007; Yu; Yu; Wang, 2014; Gazni; Didegah, 2021).

The problem begins when self-citation is unnecessary and there is no thematic or methodological relationship between the citing work and the work cited. Unnecessary or excessive self-citations are indicators of abnormality that may affect the credibility of the research and the publication itself. In excess, journal self-citations may interfere with the impact assessment, artificially increasing citation-based indicators (Bartneck; Kokkelmans, 2011; Yu; Yu; Wang, 2014).

In 2007, the IF started to have a version excluding self-citations. In 2009, the JCR began to identify and monitor self-citations of indexed publications. Several publications suffered sanctions of Clarivate Analytics due to irregularities involving self-citations in the IF (Huang; Cathy Lin, 2012; Liu; Fang, 2020).

At the same time, new impact indicators were created based on the exclusion of self-citations, such as the Eigenfactor Score (ES) and the Article Influence Scores (AIS) (Bergstrom; West; Wiseman, 2008) or based on the limitation of self-citations, such as the Scientific Journal

Rankings (SJR) (González-Pereira; Guerrero-Bote; Moya-Anegón, 2010). These indicators assume that self-citations interfere with the impact or influence of the publication in its scientific field.

The differentiation between citations and self-citations facilitated large-scale studies, especially identifying self-citation patterns or rates. Journal self-citation rates are defined in two ways. The first, self-citing, is defined by the percentage of self-citations among the total number of references in a journal. The second, self-cited, by the percentage of self-citations in the total number of citations received by a journal (Rousseau, 1999; Frandsen, 2007).

Self-citation rates are established at different levels. They may be calculated by journal, discipline or research area, publisher or institution, language, and country (Taşkin *et al.*, 2021).

The self-citation rates are higher soon after the publication of the work, decreasing their occurrence over the years. They tend to be higher in journals of lower impact (Heneberg, 2016), regional languages (Taşkin *et al.*, 2021), specialized subjects (Livas; Delli, 2018; Sanfilippo *et al.*, 2021), and areas of lower visibility (Rousseau, 1999; Frandsen, 2007).

On the other hand, the average self-citation rates do not seem to show relevant differences among subject categories, remaining at close percentages in studies of different disciplines (Taşkin *et al.*, 2021).

A study of 5876 journals indexed in the 2002 JCR found that 82% of the publications had rates lower than 20%, with a general self-citation average of 5.87% and a median of 9%. These JCR patterns do not appear to have varied in the following years.

A new study conducted in 2018 with 11866 journals indexed in the JCR found that approximately 10% had 25% or more self-citations, and 5% did not have any self-citations. The average self-citation rates, from 5% to 10%, increase as the quartile of publications decreases, i.e., they are higher in publications of lower impact (Taşkin *et al.*, 2021).

The self-citation rates of 35 intensive care medicine journals range from 0% to 35.4%, with a median of 8.8% (Sanfilippo *et al.*, 2021). The rates of 85 dentistry journals decreased from 13.725% in 2014 to 10.667% in 2016 (Livas; Delli, 2018). In both studies, the journals of specific subjects in each discipline recorded higher rates than those of broad scopes (Livas; Delli, 2018; Sanfilippo *et al.*, 2021).

The research aimed to identify excessive journal self-citation patterns. Usually, rates higher than 20% or 25% are considered excessive, regardless of the area of knowledge. Even more extreme rates may indicate poor practices (Taşkin *et al.*, 2021). However, self-citation rates above the discipline average, for example, are not necessarily due to manipulation Yu; Yu; Wang, 2014). The performance of large-scale analyses is recommended for a more accurate assessment (Frandsen, 2007).

Poor journal self-citation practices may be voluntary (anticipated) or coercively induced (Chorus; Waltman, 2016). In the first case, the authors voluntarily insert self-citations imagining that this may facilitate the acceptance of the manuscript by a given journal. In the second case, the editors or reviewers require or strongly recommend that the authors cite one or more works from the journal (Chorus; Waltman, 2016; Wilhite; Fong, 2012).

In these two forms, voluntary or coercive, journal self-citations have, as a direct consequence, the artificial (abnormal) increase in impact indicators based on the frequency of citations received (Ioannidis, 2015).

Hs5-index

Unlike with the IF, journal self-citation studies from h-index data are rare. A possible reason for this gap is that only the most cited articles are considered in calculating this indicator (Barnes, 2017). It is estimated that only 5% of the total articles published by an author or a journal integrate the *h* result. Thus, in theory, self-citations could interfere with the h-index to a lesser extent than with indicators based on the average citations, such as the IF and CS (Hirsch, 2005; Waltman, 2016; Barnes, 2017).

A version of the h-index without self-citations called the hs5-index was proposed by Schreiber (2009). This proposal was based on the perception that, contrary to what Hirsch (2005) imagined, self-citations can influence the outcome, especially in the case of extreme self-citation rates (Vîiu, 2016). However, this proposal, as most studies on self-citation based on the h-index, refers only to author self-citations (Vîiu, 2016; Teixeira da Silva; Dobránszki, 2018).

The influence of journal self-citations on the h-Index, on the other hand, is not yet well known. In the case of Google Scholar Metrics (GSM), the inexistence of data analysis and extraction resources makes it challenging to carry out broader studies (López-Cózar; Cabezas-Clavijo, 2013). The lack of standardization of titles (Harzing, 2014) may also interfere with the analysis of the correspondence of the titles of citing and cited sources.

The lack of self-citation parameters may be an obstacle to using the h5-index in evaluation systems. Journals with excess self-citations may gain an advantage in defining impact levels (rankings, quartiles, or strata) over publications with normal self-citation rates. It is recommended to monitor self-citations in the h5-index to avoid manipulations.

Methodological procedures

This work presents partial results of the doctoral research by Canto (2022), carried out from 4049 Ibero-American journals indexed in the 2021 edition of the GSM, which has the h5-index calculated from articles published from 2016 to 2020 and citations registered until July 2021.

The procedures described refer to the processes of identification and analysis of journal self-citations. The *Gsm_hscite* resource was used to identify self-citations from the h5-index data of all journals in the research universe.

Gsm_hscite is the self-citation analysis feature of the *Gsm_hdata* tool (Canto, 2022). It performs the functions of identifying self-citations from GSM citation data and recalculating impact indicators. This feature calculates four indicators from the identified self-citations: the hs5-index (Schreiber, 2009), the hs5-median, the self-citation rate per article of the h5-index, and the self-citation rate per journal.

Gsm_hscite is configured to identify the main title and any equivalent titles, considering the non-standardization of titles in GSM. This configuration increases its accuracy because GSM does not standardize entries by journal titles. Equivalent titles may be indexed in the list of articles of the h5-index.

Figure 1 describes the operation of *Gsm_hscite* in a journal with title variations.

The (1) main title, the (2) (3) equivalent titles, and the (4) h5-index and h5-median are identified.

Figure 1 – Identification of equivalent titles in Google Scholar Metrics.
Source: Google Scholar Metrics (2022).

In a second moment, the list of articles that cite each article in the h5-core is accessed (5). A journal's titles (main and equivalent) are compared with the titles of the citing sources, recording the occurrence of journal self-citations (6), as shown in Figure 2.

Figure 2 – Identification of journal self-citation in Google Scholar Metrics.
Source: Google Scholar Metrics (2022).

This process is repeated for all citations received by the articles of the analyzed journals. Finally, the h_s5-index and h_s5-median, the percentage rate of self-citations per article (*article self-cited rate* – ASR), and the percentage rate of self-citations per journal (*self-cited rate* – SCR) (Rousseau, 1999) are calculated.

The self-citation rate per journal was chosen as the primary indicator to represent the self-citation phenomenon, considering its recurrent use in related studies.

The self-citation rate per journal (*self-cited rate*) represents the percentage of self-citations relative to the total citations received by a journal, including self-citations (Rousseau, 1999).

The self-citation analysis was divided into two parts: a general analysis and a specific analysis. The general analysis included all journals in the research universe. The specific analysis was conducted only with journals of more significant impact with identified self-citations.

The selection of the journals analyzed in the second part was based on two criteria, one of exclusion and the other of inclusion: (1) exclusion of journals with a zero self-citation rate; (2) inclusion of publications with at least 50 citations, approximately, identified by the formula h5-index × h5-median = n, with $n > 49$.

This definition was intended to achieve greater statistical accuracy. It ensures that a self-citation implies a maximum percentage of 2% in the SCR, avoiding distortions in analyses by percentage values.

It also aims to avoid interference caused by publications that received few citations, in which case a self-citation may assume a high percentage. For example, in a publication with an h5-index = 2 that received five citations, only two self-citations would represent an SCR of 40%.

Results and discussion

Figure 3 shows the SCR distribution according to the h5-index of all journals in the research universe regarding the general analysis of self-citations. No self-citations were identified in 1586 journals, corresponding to 39.17%. Of the publications with SCRs equal to zero, it was identified that 1158 (73%) had an h5-index not higher than 5, i.e., they had a low level of impact.

On the other hand, SCRs greater than 20% were not detected in journals with an h5-index greater than 25.

These data suggest that SCRs higher than the general limit of 20% set in previous studies (Taşkin *et al.*, 2021) are more frequent in journals of low and intermediate impact levels, with an h5-index ranging from 5 to 25.

The median and mean SCRs of the h5-index for the whole set were 1.37% and 3.61%, respectively. These values are lower than those obtained in IF-based self-citation studies (Livas; Delli, 2018; Sanfilippo *et al.*, 2021; Taşkin *et al.*, 2021; Gazni; Didegah, 2021), in which the mean and median rates were around 5% and 8%, approximately.

This indicates that self-citations less influence the h5-index than the IF, a characteristic that may be considered positive for using this indicator in evaluation processes.

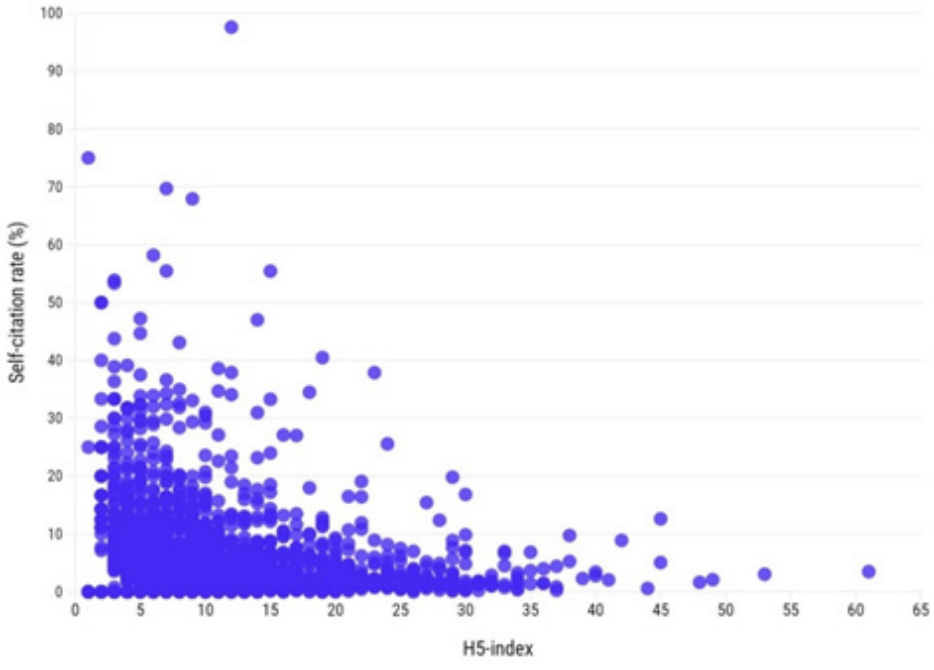

Figure 3 – Distribution of self-citation rates according to the h5-index.
Source: Research data.

Self-citation rate by country

The following analysis was conducted from a subset of journals that received at least 50 citations, approximately, with at least one self-citation. Application of these criteria resulted in identifying a subset of 1859 journals.

Upon analyzing self-citations by the country of origin of the journals, Figure 4 shows that Cuba had the highest SCRs, with a mean of 9.8% and a median of 5.6%. Brazil placed second, with a mean of 5.1% and a median of 3.9%. Both had journals with excessive SCRs, especially those above 30%. Brazil also published a journal with almost 100% self-citations. Colombia and Argentina presented a similar distribution, with median SCRs close to 3%. Colombia, however, had more outliers, especially with SCRs greater than 20%.

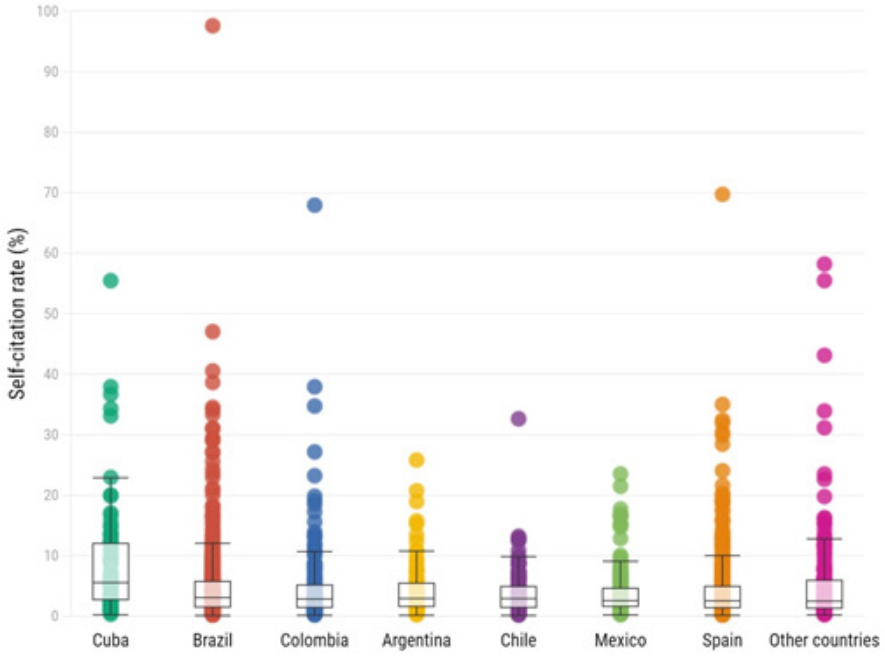

Figure 4 – Distribution of self-citation rates by country.
Source: Research data.

Chile, Mexico, and Spain had the lowest median SCRs. However, Spain and the other countries had more outliers, including at rates exceeding 40%.

The data suggest that self-citation occurs more frequently in Cuban journals than in other countries in the region. On the other hand, journals with excessive SCRs were observed in virtually all countries. This finding requires attention to the practice of self-citation in journals in the region, especially those not indexed in the JCR and Scopus.

Indexed journals tend to control their SCR, given that both databases have citation monitoring mechanisms.

Self-citation rate by area

Figure 5 shows the SCR distribution according to the area of knowledge. The analysis is supplemented by calculating centrality measures (mean and median) by area.

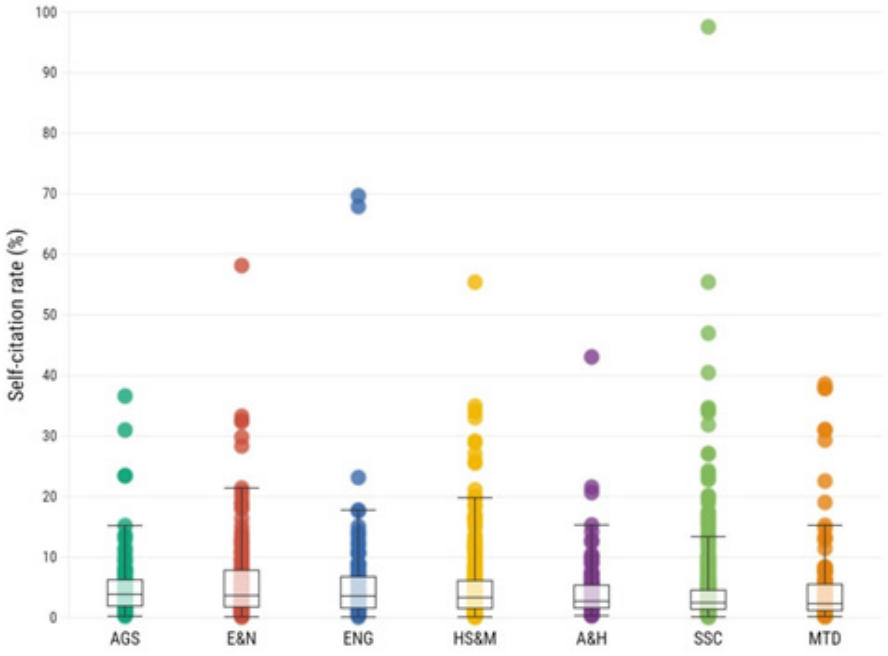

Figure 5 – Distribution of self-citation rates by area.
Source: Research data.

The mean SCRs by area were 6.28% for Engineering (ENG), 6.24% for the Exact and Natural Sciences (E&N), 5.36% for the Agricultural Sciences (AGS), 5% for Multidisciplinary (MTD), 5.12% for Health Sciences and Medicine (HS&M), 4.31% for Arts and Humanities (A&H), and 4.17% for Social Sciences (SSC). The overall mean was 4.83%.

The highest median SCRs were those of the AGS (3.89%), E&N (3.68%), ENG (3.62%), and HS&M (3.33%) journals. The areas of MTD (2.33%), SSC (2.48%), and A&H (2.74%) had the lowest median SCRs.

This difference between the two centrality measures in all areas may have been caused by outliers, i.e., publications with discrepant rates. However, discrepant SCRs were a minority, as 88.31% of journals had less than 10% of self-citations.

These results confirm that the SCRs calculated based on the h5-index tend to be lower than those calculated from the IF (Taşkin *et al.*, 2021). This reinforces the thesis of the lower susceptibility of the index to the influence of self-citations (Hirsch, 2005; Waltman, 2016), considering its sample formula based on the most cited articles.

This may support arguments against the exclusion of self-citations in impact assessment. On the other hand, it would be difficult to defend the maintenance of the performance of journals that, notably, had a significant increase in performance due to the excess of self-citations.

Based on the premise that excessive self-citation should be monitored (Frandsen, 2007), the analysis compares journals with normal and excessive SCRs. Journals with normal SCRs are those within mean distribution levels. Journals with excessive SCRs are those statistically identified as outliers in journals with common characteristics, such as the same research area of knowledge (Figure 5) or the same country of origin (Figure 4).

The differentiation between normal and outlier SCRs thus stems from statistical analysis. No value judgment is made in the scope of this research about the reasons for any extravagant SCRs in the analyzed journals.

Variation between the h5 and hs5 indices

Figure 6 shows the average variation between the h5 and hs5 indices (VrH) in journals with normal and outlier SCRs of the entire subset of the specific self-citation analysis.

The outliers had a mean h5-index lower than the other journals, as reported in studies based on the IF (Heneberg, 2016; Taşkin *et al.*, 2021). The difference is accentuated in the comparison using the hs5-index, with a mean decrease of 33.5% in the average impact of *outliers*, compared to a drop of less than 5% in other journals.

This result indicates that the excess of self-citations may benefit journals and distort the results of evaluation processes based on the h5-index of the GSM. It is advisable to monitor self-citations and, based on this, set boundaries between normal and excessive SCRs.

Figure 7 shows that the outliers obtained an h5-index growth at a higher pace than the others in the last three editions of GSM (2019–2021). The cumulative advantage was 10% when the two measurements were combined.

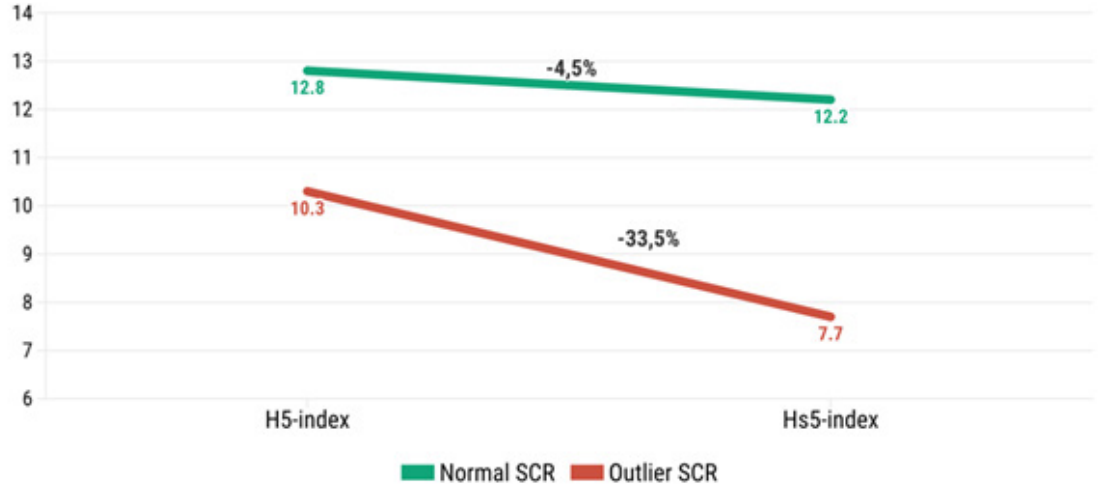

Figure 6 – Variation between the h5-index and h_s5-index.
Source: Research data.

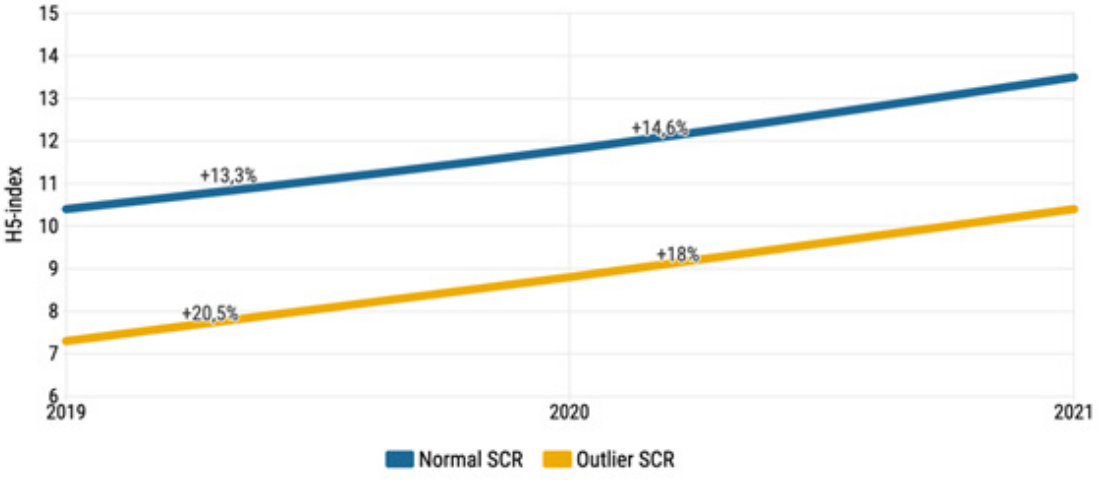

Figure 7 – Growth rate of the h5-index.
Source: Research data.

However, it is impossible to attribute all the additional growth of outlier journals to self-citations alone. However, based on the defined indicators, this justified the conduction of a more in-depth analysis, especially by comparing the average performance of journals with normal and outlier SCRs.

Conclusions

The results suggest that the h5-index is less susceptible to self-citations' influence than the IF. The calculation of the h-index results in disregarding a significant portion of the citations received by a journal, including self-citations.

Some findings corroborated findings from previous studies. The areas with the highest SCRs were those with the fewest journals, i.e., the most specialized areas. This trend may be because more specialized areas form citation networks with lower density, which increases the probability of identity between the citing source and cited source, thus configuring journal self-citations.

Journals with higher SCRs had lower average impact levels. This indicates that self-citations may be a self-promotion strategy adopted by more recent publications that still lack recognition. A hypothesis to be tested in future studies is if, after an initial period with more self-citations, reaching an intermediate impact level can increase the proportion of external citations and decrease the SCR.

In general, the data supported the lack of a need for prior exclusion of journal self-citations in impact analyses based on the h5-index of the GSM, as advocated by part of the literature.

From another perspective, it has been shown that not excluding self-citations beforehand cannot lead to the total disregard of this parameter. It is necessary to monitor the SCR to prevent distortions, given that journals with statistically discrepant percentages were identified in all areas.

The resistance of the h5-index to self-citations decreases with discrepant rates. At the outlier level, self-citations can decisively interfere with impact levels, conferring an advantage compared to journals with moderate SCRs.

This indicates the need for monitoring self-citations in the h5-index and, if necessary, correcting the results of evaluation indicators in some cases. Corrections may involve lowering the position in rankings, strata, or quartiles, adjusting the value of the h5-Index, or classifying journals as non-scientific, given the preponderance of self-citations relative to other citations.

References

Barnes, C. The h-index Debate: An Introduction for Librarians. *The Journal of Academic Librarianship*, v. 43, n. 6, p. 487-494, 2017. Doi: https://doi.org/10.1016/j.acalib.2017.08.013

Bartneck, C.; Kokkelmans, S. Detecting h-index manipulation through self-citation analysis. *Scientometrics*, v. 87, n. 1, p. 85–98, 2011. Doi: https://doi.org/10.1007/s11192-010-0306-5

Bergstrom, C. T.; West, J. D.; Wiseman, M. A. The EigenfactorTM metrics. *Journal of Neuroscience*, v. 28, n. 45, p. 11433–11434, 2008. Doi: https://doi.org/10.1523/JNEUROSCI.0003-08.2008

Campanario, J. Journals that Rise from the Fourth Quartile to the First Quartile in Six Years or Less: Mechanisms of Change and the Role of Journal Self-Citations. *Publications*, v. 6, n. 4, p. 47, 26 nov. 2018. Doi: https://doi.org/10.3390/publications6040047

Canto, F. L. *Avaliação de impacto de periódicos ibero-americanos com base no índice h5 do Google Scholar Metrics*. 2022. 189 f. Tese (Doutorado em Ciência da Informação) – Universidade Federal de Santa Catarina, Florianópolis, 2022. Disponível em: https://repositorio.ufsc.br/handle/123456789/237341. Acesso em: 17 nov. 2022.

Chorus, C.; Waltman, L. A large-scale analysis of impact factor biased journal self-citations. *PLoS ONE*, v. 11, n. 8, p. 1-11, 2016. Doi: https://doi.org/10.1371/journal.pone.0161021

López-Cózar, E.; Cabezas-Clavijo, Á. Ranking journals: could Google Scholar Metrics be an alternative to Journal Citation Reports and Scimago Journal Rank? *Learned Publishing*, v. 26, n. 2, p. 101-113, 2013. Doi: https://doi.org/10.1087/20130206

Flatt, J.; Blasimme, A.; Vayena, E. Improving the Measurement of Scientific Success by Reporting a Self-Citation Index. *Publications*, v. 5, n. 3, p. 20, 2017. Doi: https://doi.org/10.3390/publications5030020

Frandsen,T. F. Journal self-citations-Analysing the JIF mechanism. *Journal of Informetrics*, v. 1, n. 1, p. 47-58, 2007. Doi: https://doi.org/10.1016/j.joi.2006.09.002

Gazni, A.; Didegah, F. Journal self-citation trends in 1975–2017 and the effect on journal impact and article citations. *Learned Publishing*, v. 34, n. 2, p. 233-240, 2021. Doi: https://doi.org/10.1002/leap.1348

González-Pereira, B.; Guerrero-Bote, V. P.; Moya-Anegón, F. A new approach to the metric of journals' scientific prestige: The SJR indicator. *Journal of Informetrics*, v. 4, n. 3, p. 379-391, 2010. Doi: https://doi.org/10.1016/j.joi.2010.03.002

Hartley, J. On the need to distinguish between author and journal self-citations. *Scientometrics*, v. 81, n. 3, p. 787-788, 2009. Doi: https://doi.org/10.1007/s11192-008-2241-2

Harzing, A.-W. A longitudinal study of Google Scholar coverage between 2012 and 2013. *Scientometrics*, v. 98, n. 1, p. 565-575, 2014. Doi: https://doi.org/10.1007/s11192-013-0975-y

Heneberg, P. From excessive journal self-cites to citation stacking: analysis of journal self-citation kinetics in search for journals, which boost their scientometric indicators. *PLoS ONE*, v. 11, n. 4, p. 1-20, 2016. Doi: https://doi.org/10.1371/journal.pone.0153730

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

Huang, M.; Cathy Lin, W. The influence of journal self-citations on journal impact factor and immediacy index. *Online Information Review*, v. 36, n. 5, p. 639-654, 2012. Doi: https://doi.org/10.1108/14684521211275957

Ioannidis, J. A generalized view of self-citation: direct, co-author, collaborative, and coercive induced self-citation. *Journal of Psychosomatic Research*, v. 78, n. 1, p. 7-11, 2015. Doi: https://doi.org/10.1016/j.jpsychores.2014.11.008

Jacsó, P. Google Scholar Metrics for Publications: The software and content features of a new open access bibliometric service. *Online Information Review*, v. 36, n. 4, p. 604-619, 2012. Doi: https://doi.org/10.1108/14684521211254121

Liu, X. Z.; Fang, H. A comparison among citation-based journal indicators and their relative changes with time. *Journal of Informetrics*, v. 14, n. 1, p. 101007, 2020. Doi: https://doi.org/10.1016/j.joi.2020.101007

Livas, C.; Delli, K. Journal Self-Citation Rates and Impact Factors in Dentistry, Oral Surgery, and Medicine: A 3-year Bibliometric Analysis. *Journal of Evidence-Based Dental Practice*, v. 18, n. 4, p. 269-274, 2018. Doi: https://doi.org/10.1016/j.jebdp.2017.09.001

Rousseau, R. Temporal differences in self-citation rates of scientific journals. *Scientometrics*, v. 44, n. 3, p. 521-531, 1999. Doi: https://doi.org/10.1007/bf02458493

Sanfilippo, F. *et al.* Self-citation policies and journal self-citation rate among Critical Care Medicine journals. *Journal of Intensive Care*, v. 9, n. 1, p. 1-5, 2021. Doi: https://doi.org/10.1186/s40560-021-00530-2

Schreiber, M. The influence of self-citation corrections and the fractionalised counting of multi-authored manuscripts on the Hirsch index. *Annalen der Physik (Leipzig)*, v. 18, n. 9, p. 607-621, 17 ago. 2009. Doi: https://doi.org/10.1002/andp.200910360

Szomszor, M.; Pendlebury, D. A.; Adams, J. How much is too much? The difference between research influence and self-citation excess. *Scientometrics*, v. 123, n. 2, p. 1119-1147, 2020. Doi: https://doi.org/10.1007/s11192-020-03417-5

Taşkin, Z. *et al.* Self-Citation Patterns of Journals Indexed in the Journal Citation Reports. *Journal of Informetrics*, v. 15, n. 4, 2021. Doi: https://doi.org/10.1016/j.joi.2021.101221

Teixeira da Silva, J. A.; Dobránszki, J. Multiple versions of the h-index: cautionary use for formal academic purposes. *Scientometrics*, v. 115, n. 2, p. 1107-1113, 2018. Doi: https://doi.org/10.1007/s11192-018-2680-3

Vîiu, G. A. A theoretical evaluation of Hirsch-type bibliometric indicators confronted with extreme self--citation. *Journal of Informetrics*, v. 10, n. 2, p. 552-566, 2016. Doi: https://doi.org/10.1016/j.joi.2016.04.010

Waltman, L. A review of the literature on citation impact indicators. *Journal of Informetrics*, v. 10, n. 2, p. 365-391, 2016. Doi: https://doi.org/10.1016/j.joi.2016.02.007

Wilhite, A. W.; Fong, E. A. Coercive citation in academic publishing. *Science*, v. 335, n. 6068, p. 542-543, 2012. Doi: https://doi.org/10.1126/science.1212540

Yu, T.; Yu, G.; Wang, M. Y. Classification method for detecting coercive self-citation in journals. *Journal of Informetrics*, v. 8, n. 1, p. 123-135, 2014. Doi: https://doi.org/10.1016/j.joi.2013.11.001

Zhou, Y. Self-citation and citation of top journal publishers and their interpretation in the journal-discipline context. *Scientometrics*, v. 126, n. 7, p. 6013-6040, 2021. Doi: https://doi.org/10.1007/s11192-021-03995-y

Contributors

F. L. Canto and M. Talau contributed to the study conception, design and interpretation. A. L. Pinto and T. M. R. Dias contributed to the review and approval of the final version of the manuscript.

Acknowledgment

We would like to thank the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) – CNPq/MCTI, Notice nº26/2021. Process 402042/2022-0 and 200937/2022-7. Post-Doctorate Abroad. Protocol 7763808095369450.