Revista Española de Documentación Científica
47(1), enero-marzo 2024, e375
https://doi.org/10.3989/redc.2024.1.1454

ESTUDIOS / *RESEARCH STUDIES*

Invisible Publications: A Study of Academic Productivity in the Web of Science Database

Eliandro Schvirck*, Caroline Lievore**, Priscila Rubbo***, José Roberto Herrera Cantorani****, Luiz Alberto Pilatti*****

Federal University of Technology – Paraná (UTFPR). Brasil
*Correo-e: eliandro@utfpr.edu.br | ORCID iD: https://orcid.org/0000-0003-1957-9115
**Correo-e: carolievore1@gmail.com | ORCID iD: https://orcid.org/0000-0003-2448-089X
***Correo-e: priscilarubbo@utfpr.edu.br | ORCID iD: https://orcid.org/0000-0002-2488-4627
*****Correo-e: lapilatti@utfpr.edu.br | ORCID iD: https://orcid.org/00000-0003-2679-9191
****Federal Institute of São Paulo (IFSP)
Department of Program in Production Engineering. Brazil.
Correo-e: antorani@yahoo.com.br | ORCID iD: https://orcid.org/00000-0003-1792-0383

Recibido: 08/03/23; 2ª versión: 20/06/23; Aceptado 10/07/23; Publicado: 28/02/24

Cómo citar este artículo/Citation: Schvirck, E., Lievore, C., Rubbo, P., Herrera Cantorani, J. R., Pilatti, L. A. (2024). Invisible Publications: A Study of Academic Productivity in the Web of Science Database. *Revista Española de Documentación Científica*, 47 (1), e375. https://doi.org/10.3989/redc.2024.1.1454

Abstract: The purpose of this study is to examine the contemporary landscape of academic publishing, with a specific emphasis on the dissemination and evaluation of scholarly research, and the characteristics of articles that have remained uncited or have been infrequently cited, indexed in the *Web of Science* database within the timeframe of 2017 to 2021. Descriptive statistics were used to characterize the sample according to field of research and variable. A negative binomial regression method was used to estimate which variables influence the number of citations received by articles in the sample. We found that of the 256,524 articles that make up the research corpus, 39,469 (15.39%) received no citations and 91,963 articles (35.85%) received up to four citations. We conclude that the pressures of academic productivity create a vicious cycle in which studies are referenced based on their ranking in the struggle for privileged editorial space.

Keywords: Academic productivity; citation; Journal Citation Reports; Web of Science; invisible publications.

Publicaciones invisibles: un estudio de la productividad académica en la base de datos Web of Science

Resumen: El propósito de este estudio es examinar el panorama contemporáneo de la publicación académica, con énfasis específico en la difusión y evaluación de la investigación académica y en las características de los artículos indexados en la base de datos *Web of Science*, en el marco temporal de 2017 a 2021, que no han sido citados o han sido citados con poca frecuencia. Se utilizó la estadística descriptiva para caracterizar la muestra según el campo de investigación. Se utilizó un método de regresión binomial negativa para estimar qué variables influyen en el número de citas recibidas por los artículos de la muestra. Encontramos que de los 256.524 artículos que componen el corpus de investigación, 39.469 (15,39%) no recibieron citas y 91.963 artículos (35,85%) recibieron hasta cuatro citas. Concluimos que las presiones de la productividad académica crean un círculo vicioso en el que los estudios son referenciados en función de su posición en la lucha por un espacio editorial privilegiado.

Palabras clave: Productividad académica; citación; Journal Citation Reports; Web of Science; publicaciones invisibles.

1. INTRODUCTION

Although the main purpose of publication is to disseminate research results, it is also one of the primary ways in which academic success is assessed when evaluating researcher performance (Denning, 1997; De Rond and Miller, 2005; Fanelli, 2010; Lawrence, 2008; Suryani et al., 2015; Tian et al., 2016). Publications are also a key aspect in obtaining and/or maintaining research funding and are commonly used to improve a university's position in global rankings (Fanelli, 2010; Lawrence, 2008; Suryani et al., 2015).

In the academic-scientific community, the pressure and desire to publish and have one's research cited have become characteristic of the current context and the *modus operandi* of most researchers (De Rond and Miller, 2005; Denning, 1997; Tian et al., 2016). In recent decades, the academy has been marked by the now familiar maxim "publish or perish." This saying was first linked to academic and research contexts in 1942 (Garfield, 1998; 2006a; 2006b), but is regularly associated with current pressures to establish, maintain, and advance one's career or to simply the keep one's job (Alcadipani, 2011; Demir, 2018; De Rond and Miller, 2005; Fanelli, 2010; Furnham, 2021; Lawrence, 2008; Leite, 2017; Machado and Bianchetti, 2011; Tian et al., 2016). To some extent, this has devalued the original purpose of research in universities and has brought about worrying consequences, since it is becoming increasingly more difficult to distinguish serious research from that conducted merely to increase one's number of publications (Denning, 1997; Rego, 2014).

For Castiel and Sanz-Valero (2007), papers have become an academic commodity following the trend of "bibliographic Darwinism," a struggle for survival in which the academics who publish the most are those who stand out. Similarly, Alcadipani (2011) cites "academic productivity" as the emphasis on producing a large quantity of something that has little substance and, consequently, using quantity as a proxy for quality. For Godoi and Xavier (2012), the very definition of academic productivity is dialectical, as it engenders the very situation being criticized: a form of evaluation centered on the number of papers, which in general are rarely read or of no major scientific importance, that serve as a basic parameter for advancing one's academic career.

The issue of academic productivity has been widely addressed in the literature (De Rond and Miller, 2005; Lawrence, 2008; Machado and Bianchetti 2011; Tian et al., 2016; Zuin and Bianchetti, 2015). In the context of citations and self-citations, some authors have focused on analyzing the knowledge produced (Kacem et al., 2020; Szomszor et al., 2020), the differences between highly cited and uncited research (Suryani et al., 2015), and the consequences of the *ethos* of publish or perish for faculty in terms of their intellectual life, ethics, and work-related stress (De Rond and Miller, 2005; Leite, 2017; Tian et al., 2016).

Although the work by Fanelli (2010) supports the hypothesis that competitive academic environments increase scientific productivity, some academic debates strongly question the imperative of publish or perish. The major criticism lies in the fact that the core of this model considers how much the researcher publishes, to the detriment of the scientific quality, innovation, or social relevance of what is published – often reaffirming the obvious (De Rond and Miller, 2005; Demir, 2018; Denning, 1997; Dettori at al., 2019; Godoi and Xavier, 2012; Tian et al., 2016). For Lawrence (2008), the publication of research has always been crucial for scientists with the main objectives being to: disseminate new knowledge; allow other scientists to replicate or expand on studies; and justify financial or other types of support provided to the scientist. Once secondary, since the advent of academic productivity, financial support has become one of the main goals in a scientist's career (Lawrence, 2008).

The number of articles published in scientific journals not only brings researchers recognition in their field, but it has also resulted in another ambition: to be cited (De Rond and Miller, 2005; Garfield, 2006b; Tian et al., 2016). The obsession with receiving citations (citationism) creates a scenario in which journals and researchers focus their efforts on being citated as a means to improve their *h*-index and impact factor scores (Bornmann and Marx, 2020; Rego, 2014). As a result, many journals that publish review articles achieve higher impact factors compared to journals that publish original research (Moustafa, 2015). In this regard, Hicks et al. (2015) state that publication count has become an opportunity for self-promotion and a way to compete in the ranking of academic indices. This argument was outlined in the Leiden Manifesto, published in *Nature*, which presents ten principles for evaluating scientific research. Regarding the seventh principle "Base assessment of individual researchers on a qualitative judgment of their portfolio", which refers to the *h*-index, Hicks et al. (2015) argue that reading, evaluating, and judging a researcher's work, their experiences, activities, and skills is much more appropriate than relying on a number. From a complementary point of view, Moustafa (2015) argues that some editors build their editorial strategies to increase impact

factors, encouraging authors to cite their journal, thus creating an unfair and exclusionary selection policy. For Rego (2014), the misuse of the impact factor distorts the real meaning of the indicator, with negative consequences for the development of journals, as it encourages the adoption of practices that artificially increase citation rates, often disregarding the ethical standards established and shared in scientific communication.

As such, a frequently cited work does not necessarily represent advances in knowledge or high productivity in research (De Rond and Miller, 2005). Meanwhile, the pressure to publish ends up discouraging bold and original research, especially by younger researchers. As discussed by Smith (1990), when the focus is on publication, the risks associated with doing something unorthodox or that upsets deeply entrenched prejudices in a specific field are considerable. In this context, De Rond and Miller (2005), when analyzing the consequences of the pressure to publish, found a lack of intellectual daring and methodological rigor in research, the primacy of methodological conventionality, and negligence in innovation. These were identified as important limiting factors to produce innovative research. As a result of the current high productivity orthodoxy, more insignificant research is being generated.

An emphasis on citations as a means of evaluating the quality of an academic work can become a vicious cycle that can influence the processes of searching for and producing knowledge (Castiel and Sanz-Valero, 2007; Lawrence, 2008). This context, which encompasses academic productivity and citationism, has cast a shadow over academia causing irreversible damage, including: encouraging the precarities of an excess of redundant or superfluous information (Castiel and Sanz-Valero, 2007; Else and Van Noorden, 2021; Tian et al., 2016); promoting the publication of the same content in several articles across different journals, or "salami slicing", through the publication of underdeveloped, incomplete, split up, or repetitive articles (Castiel and Sanz-Valero, 2007; Lawrence, 2008; Rego, 2014); bombarding prestigious journals with submissions, occupying the time of editors and ad hoc reviewers in an exhausting review cycle (Rego, 2014); enabling the emergence of predatory publishers that have increased dramatically in number and wealth in recent years or companies that sell scientific articles to researchers (Rego, 2014; Else and Van Noorden, 2021); generating a stand-off between ethics and research as it forces scientists to produce "publishable" results at any cost, encouraging scientific misconduct (Castiel and Sanz-Valero, 2007; Fanelli, 2010; Moustafa, 2015; Rego, 2014; Rubbo et al., 2019; Sayão et al., 2021); and causing health issues among university professors (Leite, 2017; Tian et al., 2016).

As it stands, the scholarly publishing system has been perverted into a mechanism to advance one's career, swamping databases with countless mediocre, unread, unreadable, and useless articles (De Rond and Miller, 2005; Sykes, 1988).

Despite the recurrent mention in the literature of rarely cited or uncited articles, only a few studies specifically analyze this body of literature. One such article is that of Sayão et al. (2021). Although they take a triumphalist view of science, they show that an important proportion of scientific research produces negative results (inconclusive experiments or unexpected data), which almost as a rule, makes publication unfeasible or leads to publications that are rarely cited. This creates a paradox in that a key aspect in the development of scientific knowledge comes from the findings that reject consolidated hypotheses and drive scientific progress but are never published.

The purpose of this study is to examine the contemporary landscape of academic publishing, with a specific emphasis on the dissemination and evaluation of scholarly research, and the characteristics of articles that have remained uncited or have been infrequently cited, indexed in the *Web of Science* database within the timeframe of 2017 to 2021. The parameters for this examination will focus primarily on elements associated with the pursuit of publications and the volume of citations, which correspond to the principles of academic productivity and citationism respectively. Within this context, the *Journal Citation Reports* (JCR) represents a pivotal variable. It plays a determinative role in evaluating a journal's impact factor (Kaldas et al., 2020; Gilyarevskii et al., 2021; Kulczycki et al., 2021), and further acts as a propellant for the author, enhancing the visibility of their research (Haba-Osca et al., 2019; Tang, 2013). Corresponding to the *Journal Citation Reports* (JCR), the variables, namely Quartile and Language, also hold significant importance. Moreover, the variables Maturity and Funding have demonstrated relevance in this context. These independent variables will be subjected to an investigation in correlation with the dependent variable, which is identified as the "Number of Citations".

2. METHODOLOGY

The methodology section describes the data collection process and the statistical analysis used to examine the relationship between various

variables and the number of citations received. The research data were extracted from the *Web of Science* (WoS) database. The WoS is considered one of the most relevant and comprehensive databases of scientific publications, with almost 1.9 billion references cited in more than 171 million records, representing multidisciplinary and selective data (Pranckutė, 2021).

The study sample was defined based on a simple stratified random sample from 14425 journals listed in the WoS and on the *Journal Citation Reports* (JCR) website, on April 28, 2022, thus, in order to ascertain the statistical significance of the sample, Equation 1 was used based on Martins (2002) to define the sample size, considering the use of nominal variables in a finite study population, with a confidence level of 95%, the sample size was calculated as 373 journals.

Equation 1: Sample calculation for nominal variables in a finite population.

$$n = \frac{z^2.p.q.N}{d^2(N-1) + z^2.p.q}$$

The research areas were classified according to WoS, which, classifies existing content into five major fields of research, that are further divided into subfields. The sample was proportionally stratified to represent each major field of research and journals were distributed into basic classification quartiles. The stratified sample was defined as shown in Table I.

The data used in the research were collected from articles published between 2017 and 2021, for a total of 256524 articles, from which the variables used for analysis were extracted, as shown in Table II.

The dependent variable was defined as the number of citations, with independent variables being the JCR index, maturity, quartile, language of publication, and whether the research was funded.

For the statistical analysis of the data, the software IBM SPSS version 27 was used. The analysis was performed in two segments. Initially, using descriptive statistical techniques, the study sample was characterized considering the proportion of articles - by classification area - according to the number of citations, the JCR index, citation maturity, research funding, publication language and classification by quartile and also data related to the mean, standard deviation and data range.

Then, after the descriptive characterization of the sample, the negative binomial regression

Table I: Number of journals in the stratified sample.

Field of research	Quartile 1	Quartile 2	Quartile 3	Quartile 4	Total
Arts and Humanities	1	1	1	1	4
Life Sciences and Biomedicine	39	42	42	41	164
Physical Sciences	13	13	13	14	53
Social Sciences	20	21	21	21	83
Technology	15	18	18	18	69
Total number of journals	88	95	95	95	373

Table II: Analysis variables.

Variable	Metric
Number of citations	Number of citations registered in all databases, available in WoS
JCR	JCR ranking in year 2020
Maturity	Date of publication in relation to the year 2022: "1" for articles published in 2021; "2" for articles published in 2020; "3" for articles published in 2019; "4" for articles published in 2018; and "5" for articles published in 2017.
Quartile	Quartile ranking of the journal in the WoS database
Language	Language of publication of the article: "0" for English and "1" for others
Funding	"0" for unfunded research and "1" for funded research
Number of citations in the last 180 days	Number of citations registered in all databases, available in WoS

model was used to estimate which variables, among those listed in the study, may influence the number of citations of this sample. The negative binomial regression method was used due to the characteristic of the study data, since the dependent variable, citation count, has a Poisson probability distribution and presents great dispersion in the observations, as endorsed by Fávero and Belfiore (2021) the negative binomial regression model is appropriate for estimation when the dependent variable is quantitative and with whole and non-negative values (count data) and when there is great dispersion in the data.

Following Fávero and Belfiore (2021), the negative binomial regression is estimated as:

Equation 2: Negative binomial regression estimation model.

$$\lambda i = e^{(\alpha+\beta_1.\times_{1i}+\beta_2.\times_{2i}+\ \ldots\ \beta_k.\times_{ki}+\varepsilon i)}$$

3. RESULTS

The results section presents the inferences related to the proportion of uncited articles, the

Table III: Total number of published and uncited articles per Web of Science research field.

Research field	No. of articles	No. of uncited articles	% uncited articles
Arts and Humanities	615	289	46.99%
Life Sciences and Biomedicine	95074	13976	14.70%
Physical Sciences	55280	8370	15.14%
Social Sciences	27628	6139	22.22%
Technology	77927	10695	13.72%
Total	256524	39469	15.39%

Table IV: Description of the number of citations received by publications by WoS fields of research.

	Arts and Humanities	Life Sciences and Biomedicine	Physical Sciences	Social Sciences	Technology	Total
Mean	1.74	8.94	8.36	6.19	11.31	9.22
Median	1	4	4	3	5	4
Mode	0	0	0	0	0	0
Standard deviation	3.08	19.40	13.46	11.97	19.48	17.66
Minimum	0	0	0	0	0	0
Maximum	27	2095	1129	342	711	2095
Total no. of citations	1071	850298	462280	171056	881191	2365896
Total no. of articles	615	95074	55280	27628	77927	256524

Table V: Number of citations received by the sample considering the maturity of the article.

Number of citations	Maturity					Total (sample percentage)
	1	2	3	4	5	
0	22975	7541	3991	2713	2249	39469 (15.39%)
1	13128	7410	4532	3159	2575	30804 (12.01%)
2	7551	6462	4614	3234	2745	24606 (9.59%)
3	4482	5495	4242	3115	2680	20014 (7.80%)
4	2833	4316	3899	2922	2569	16539 (6.45%)
5	1827	3501	3484	2726	2342	13880 (5.41%)
6	1275	2840	3038	2420	2202	11775 (4.59%)
7	868	2318	2665	2305	2057	10213 (3.98%)
8	606	1972	2376	2071	1973	8998 (3.51%)
9	436	1584	2036	1812	1774	7642 (2.98%)
10	367	1290	1764	1682	1653	6756 (2.63%)

distribution of citations, and the impact of different factors on citation counts.

In Table III, the research corpus and the number and percentage of uncited articles are presented based on the fields of research classified by WoS.

The 256524 articles that make up the research corpus received 2365896 citations in the survey carried out. Table IV presents the description of these publications by field.

Adopting ten citations as a cutoff, we considered the number of citations received for the sample as a function of the maturity of the articles (Table V).

Table VI shows the articles that received no citations in the 180 days before data sampling.

Of the research corpus, four articles received more than 1000 citations (Table VII).

To measure the dependent variable (number of citations received), five independent variables were selected: **JCR**; **language** in which the article was published; existence of **funding**; distribution of articles in WoS **quartiles**; and **maturity** of the publication.

For the analysis, the research corpus was sectioned into deciles based on the JC. The range of the JCR index of the studied journals is from 0.089 (*Psycho-Oncologie*) to 24.88 (*Nature Neuroscience*). It is important to note that across all JCR scores, articles with no citations were identified. Table VIII shows the average JCR of

Table VI: Uncited publications in the last 180 days, as a function of maturity.

	1 year	2 years	3 years	4 years	5 years	Total
Total no. of articles	57934	53043	52638	46845	46334	256524
No. of uncited articles	13272	16903	19379	19758	22208	91520
Percentage of uncited articles	22.90%	31.86%	37.00%	42.18%	47.93%	35.67%

Table VII: Articles published between 2017 and 2021 in the *Web of Science* database that received 1000 citations or more.

Authors	Title (DOI)	Journal (JCR)	Year	Citations
Shi et al.	Association of cardiac injury with mortality in hospitalized patients with COVID-19 in Wuhan, China (10.1001/jamacardio.2020.0950)	JAMA Cardiology (14.676)	2020	2095
Guo et al.	Cardiovascular implications of fatal outcomes of patients with Coronavirus Disease 2019 (COVID-19) (10.1001/jamacardio.2020.1017)	JAMA Cardiology (14.676)	2020	2023
Bakker and Demerouti	Job demands-resources theory: Taking stock and looking forward (10.1037/ocp0000056)	Journal of Occupational Health Psychology (7.25)	2017	1209
Zhu et al.	Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources (10.1109/MGRS.2017.2762307)	IEEE Geoscience and Remote Sensing Magazine (8.225)	2017	1129

Obs.: Survey carried out on April 28, 2022.

Table VIII: JCR index of journals with uncited publications in the first and tenth deciles.

	JCR of 10th decile journals	JCR of 1st decile journals
Mean	2.862	4.972
Median	2.712	3.000
Standard deviation	1.9361	5.123
Interval	8.041	24.759
Minimum	0.408	0.125
Maximum	8.449	24.884
Number	37	37

Table IX: Uncited publications in the first and tenth deciles.

	Uncited articles from the 10th decile	Uncited articles from 1st decile
Mean	458.48	9.05
Median	329	9
Mode	803	14
Standard deviation	337.69	4.49
Interval	1433	14
Minimum	213	1
Maximum	1646	15
Total	16,964	335

Table X: Language of articles.

Language	Total no. of articles	No. of uncited articles	% of uncited articles
English	254972	38573	15.13%
Others	1552	896	57.73%
Total	256524	39469	15.39%

Table XI: Research funding.

Funding	Total no. of articles	No. of uncited articles	% of uncited articles
No	74366	15303	20.58%
Yes	182158	24166	13.27%
Total	256524	39469	15.39%

Table XII: Distribution of articles by quartile.

	Q1	Q2	Q3	Q4	Total
Total no. of articles	101850	73929	48518	32227	256524
No. of uncited articles	9038	12193	8009	10229	39469
% of uncited articles	8.87%	16.49%	16.51%	31.74%	15.39%

Table XIII: Uncited articles considering the maturity of the publications.

	1 year	2 years	3 years	4 years	5 years	Total
Total no. of articles	57934	53043	52638	46845	46334	256524
No. of uncited articles	22975	7541	3991	2713	2246	39469
% of uncited articles	39.66%	14.22%	7.58%	5.79%	4.85%	15.39%

the journals and the descriptive statistics for the first and tenth deciles. In Table IX, the descriptive statistics of the uncited articles is presented considering the same deciles.

To measure the variable **language** the number of uncited articles published in English and other languages was assessed (Table X).

Table XI shows the results related to the variable funding.

Journals in the WoS database were classified into quartiles. The journals with the highest JCR indices are in the first quartile and journals with lowest JCR indices are in the fourth quartile. Table XII presents the number of articles published and

the number of uncited articles in each quartile and their percentage.

The variable maturity is compared with the number of uncited articles in Table XIII.

Considering the relationships indicated in the descriptive analysis and a negative binomial regression, an estimation model for the dependent variable (number of citations) was developed based on the effect of the independent variables (JCR, language, funding, quartile, and maturity) (Table XIV).

To evaluate equality of parameters between variables, the omnibus test was applied. The result indicated that there are constant variables in the model with significance that explain the number of citations in journals, with a p-value (Sig.) <0.05 (Table XIV). Thus, the variables JCR, maturity, quartile, language, and funding were considered valid for estimating the dependent variable.

The statistical tests for validating the variables in the model showed that all variables are significant, with a p-value (Sig.) <0.05. All coefficients (B) are within the Wald confidence interval test, reinforcing the relevance of the variables as influencing the dependent variable.

As a result, considering the relevance of the variables and the coefficients presented in the model, a regression model that can estimate the possible number of citations of an article is structured according to Equation 3.

4. DISCUSSION

The objective of the study is to discuss the current state of academic publishing, particularly in the context of research dissemination and evaluation. It highlights the pressure and emphasis on publishing as a measure of academic success, researcher performance evaluation, and securing research funding. The text addresses the phenomenon of "publish or perish," where the quantity of publications becomes prioritized over the quality, innovation, and social relevance of the research. It also discusses the obsession with citations and the impact factor as measures of research impact and the negative consequences of this emphasis on quantity over quality. The text raises concern about the proliferation of redundant and useless articles, the publication of underdeveloped or repetitive content, the exploitation by predatory publishers, the ethical dilemmas faced by researchers, and the negative effects on the well-being of academics.

Equation 3: Negative binomial regression estimation model.

$$\lambda i = e^{(0.169 + JCR.0.159 \quad + language.-1.032 \quad + Quartile.-0.17 + Maturity.0.481 + Funding.0.02 + \varepsilon i)}$$

Table XIV: Omnibus test for equality of parameters between variables.

Chi-square likelihood ratio test	df	Sig.
137812.655	5	.000

Dependent variable: CiteAllBases
Model: (Interception), JCR, Language, Funding, Quartile, Maturity

Table XV: Estimates of parameters in a negative binomial regression.

Parameter	B	Standard Error	Wald test		Hypothesis test			Exp (B)	Wald Test	
			Inf.	Sup.	Wald Chi-square	df	Sig.		Inf.	Sup.
(Interception)	0.169	0.0104	0.149	0.189	266.46	1	0.00	1.184	1.160	1.208
JCR	0.159	0.0010	0.157	0.161	23961.86	1	0.00	1.173	1.170	1.175
Language	-1.032	0.0339	-1.098	-0.965	929.13	1	0.00	0.356	0.333	0.381
Funding	0.020	0.0047	0.011	0.029	17.81	1	0.00	1.020	1.011	1.029
Quartile	-0.170	0.0026	-0.175	-0.165	4190.18	1	0.00	0.844	0.839	0.848
Maturity	0.481	0.0016	0.477	0.484	90785.11	1	0.00	1.617	1.612	1.622
(Scale)	1a									
(Negative binomial)	0.896	0.0031	0.890	0.902						

Dependent Variable: CitedAllBases
Model: (Interception), JCR, Language, Funding, Quartile, Maturity
Fixed at display value

The analysis identified 256 524 articles, distributed across the five fields of research classified in the WoS, with a greater number of publications in the fields of Life Sciences and Biomedicine (37.06%), Technology (30.38%), and Physical Science (21.55%). Of the total articles identified, 15.39% received no citation (Table III). The fields of research with the greatest number of publications are also those with a below-average number of uncited articles.

The pooled data show that the articles received an average of 9.22 citations (Table IV). As a result, the highest average number of citations was found in the field of Technology (average of 11.31 citations per article and median of 5), followed by Life Sciences and Biomedicine (average of 8.94 citations per article and median of 4), and Physical Sciences (average of 8.36 citations per article and median of 4) (Table III).

In analyzing the frequency distribution of the sample, we found that 74.34% of the publications received up to 10 citations. In this same analysis, we also identified that 23.46% of the publications received between 11 and 50 citations, 1.84% between 51 and 100 citations, and 0.36% more than 100 citations. Only four articles were cited more than 1,000 times (Table VII). Of the highly cited articles, two were about COVID-19.

Establishing four citations as a cutoff point to identify rarely cited articles (less than half of the average), we found 91,963 articles (35.85%). If we add these almost invisible articles to the uncited (invisible) articles, there were 131,432 articles (51.24%). Considering only the 180-day period prior to data collection, 35.67% of the existing production did not receive any citation.

Although several studies have shown the negative consequences of academic productivism (Kacem et al., 2020; Lawrence, 2008; Leite, 2017; Machado and Bianchetti 2011; Miller, 2005; Suryani et al., 2015; Szomszor et al., 2020; Tian et al., 2016; Zuin and Bianchetti, 2015), no study to date has offered a similar comparison of the values as presented herein. The data presented here show elevated numbers of uncited and rarely cited articles, especially when considering the importance of the journals included in the study database and the analysis of the variables JCR, language, funding, WoS quartiles, and publication maturity.

To understand the importance of JCR in the received citations, the sample was sectioned into deciles. With data extracted for the first (decile 1) and last decile (decile 10), a significantly heterogeneous sample was observed. In the first decile, the lowest JCR found was 0.125, one of the lowest in the sample, and the highest was 24.884, the highest in the sample. In the tenth decile, the lowest JCR was 0.408 and the highest 8.449. The average JCR index of the first decile (4.972) was much higher than that of the tenth decile (2.862), but the median was very similar: 3.0 for the first and 2.71 for the tenth.

The data also indicate the existence of a very large distribution of JCR in the group of journals that form the first decile, which is confirmed by the value of the standard deviation (5.123). This situation suggests that a high JCR is not a determining factor for a lower number of publications with few or no citations. However, based on data from the negative binomial regression (Table XV), a high JCR positively influences the number of citations a publication receives.

Regarding the number of uncited publications divided into the first and tenth deciles (Table IX), greater homogeneity can be seen in the first decile, with a mean of 9.05, median of 9 uncited publications, and a range of only 14 publications. In this decile there is a minimum of one uncited publication and a maximum of 15, for a total of 335 uncited publications in a group of 37 journals. In the tenth decile, 16 964 publications received no citation. In this decile, the data showed greater heterogeneity; the average was 458 publications and the distribution, given by the standard deviation, was 337.69. This is evident in the interval between the minimum and maximum number of citations (213 and 1 646 respectively) in a group of 37 journals.

The JCR is considered the gold standard for evaluating a journal's impact factor (Kaldas et al., 2020). WoS is the best database system available and its publications are considered satisfactory in academia (Gilyarevskii et al., 2021; Kulczycki et al., 2021).

Despite criticism directed at the evaluation system that assesses the impact of a journal in relation to others (Dorta-González and Dorta-González, 2013; Gonzáles-Sala et al., 2022; Kim et al., 2018; Saeed et al., 2019), publishing in a journal with a consistently high impact factor on the JCR is a benefit for the author and increases the visibility of the research (Gilyarevskii et al., 2021; Haba-Osca et al., 2019; Tang, 2013). However, a high impact factor does not necessarily reflect the quality of a publication (Kaldas et al., 2020). Additionally, in very specific editorial contexts, issues such as self-citation and editorial pressure for authors to cite works from the same journal are common (Dettori et al., 2019).

Moreover, it is highly unlikely that a researcher will dismiss a reference in a high impact journal within his or her area of research (Saeed et al., 2019). This visibility increases the probability that an article will be cited more frequently, especially those published in the last two years, demonstrating the short-term academic influence of journals with a high impact factor (Guo et al., 2021; Krampl, 2019). This scenario explains the results found herein regarding the influence of high JCR as a predictor of citation.

Regarding language, we found that publication in a language other than English tends to have a negative impact on citation. As can be seen in Table X, less than 1% of the works published in the WoS database are written a language other than English. Of these, 57.73% received no citation in the period examined, while for publications in English the percentage of uncited articles is 15.13%.

The result in relation to language is consistent with studies that argue that articles published in non-English languages have a low number of citations (Liang et al., 2013; Tahamtan et al., 2016). Additionally, journals that publish in English are positively influenced in the JCR (Liu et al., 2018). Therefore, a self-perpetuating cycle ensures that English remains the *lingua franca* of international scientific communication.

Among the 256524 articles analyzed, 71.01% declared some type of funding to carry out the research. Of these, 13.27% were not cited during the study period. However, it is important to note that the proportion of uncited articles is more representative among research without funding (20.58%) (Table XI).

Funding represents an important input in the scientific process. Funding entities appear to be influential actors in scientific communication, often stimulating the type of knowledge produced (Álvarez Bornstein and Montesi, 2020). Funding happens at the level of individual researchers, with some concentration among elite scholars (Larivière et al., 2010). Underfunding affects researchers' ability to publish high-quality articles in high-impact journals (Asubiaro, 2019). From this perspective, we can assume that funding provides the conditions necessary to produce higher quality research, increasing the potential of a study to be published in better journals with greater visibility. This assumption explains the result found.

Nevertheless, in considering the stratified sample, or the proportion of journals per quartile, Table XII indicates that 39.70% of the articles analyzed are classified as Q1, while only 12.56% are in Q4. It appears that the highest incidence of uncited articles occurs in Q4, with 31.74% of publications without any citation. In Q1, the percentage of uncited articles is 8.87%.

When analyzing the citations received by quartile in 25 areas of research, Miranda and Garcia Carpintero (2019) found that, on average, Q1 publications received 65.3% of citations, Q2 received 20.6%, Q3 received 10.0%, and Q4 received 4%. These results are consistent with those found in the present study and support the argument that articles published in Q1 journals tend to be cited more often.

Regarding the maturity of the publication, we found that most of the uncited articles are the most recent (58.21%). Considering the established maturity levels, the greatest number of articles that received one or two citations are allocated to maturity level 1 or published in the year previous to data collection. Among the articles that received between three and five citations, the greatest number is found in maturity level 2 (published two years before data collection). Of the articles that received between six and ten citations, the largest number is in maturity level 3 or three years before data collection. Based on the results, we can infer that a greater number of citations received immediately after publication (maturity 1), indicates that it will take longer to reach the peak point of citation (half-life) and it is less likely that an article will become obsolete.

This finding is consistent with the literature. Price (1976) found that most citations are for recent articles because most articles are recent. He further added that in the first few years after the publication, a paper's relative citability declines slowly and parabolically as a function of time. The half-life is a significant predictor of total citations in the next two years (Diaz Ruiz et al., 2018). In this sense, it is important to clarify that the half-life of a study varies considerably (Faber et al., 2021) and tends to be longer in human and social sciences than in the hard sciences (Gilyarevskii et al., 2021).

The findings referring to the regression estimation model, presented in equation 3, the variable maturity (0.481) has the greatest influence on the number of citations of a publication. The variables JCR (0.159) and funding (0.020) positively influence the number of citations of a publication. Also significant are the variables quartile (-0.170) and language (-1.032) that have a negative influence on estimates of the number of citations.

In this sense, scholarly articles published in non-English languages and those appearing in top-

quartile journals tend to garner fewer citations. Conversely, the publication timing of the article emerged as the most significant variable affecting the number of publications—a result that aligns with expectations, given that an article's relevance generally increases over time. It is crucial, however, to note that the publication timeframe employed in this analysis was confined to a five-year period. Articles published in journals with a higher JCR index are perceived to be of superior quality, and thus, the JCR variable can also influence the number of citations an article receives.

When discussing how variables can influence the number of citations to a publication, it is important to consider how disciplinary differences in research practices affect this influence. Research practices vary across academic disciplines, affecting how publications are cited. Some factors include communication and collaboration - disciplines differ in how researchers communicate and collaborate. In some areas it is common for researchers to work in teams, forming large collaborative groups; speed of publication - the time between submission of a paper and its publication can vary widely across disciplines. In certain fields, articles are often expeditiously published in response to the pressing need to disseminate newfound knowledge; the prestige and influence of the journal where an article gets published can be subject to variations based on the discipline; dissemination methods - the approaches employed for the propagation of research findings can also differ across disciplines.

This set of factors may be objects for future research.

5. CONCLUSION

The study's results highlight the pressure to publish, academic productivism, the use of citations as a metric, the problems of the academic publication system, and also uncited or poorly cited articles (invisible or almost invisible). Regarding the variables analyzed (JCR index, maturity, quartile, language of publication, and research funding), we found that they all influence the number of citations of a publication. Considering the importance of the journals included in the studied database and that, of the 256 524 articles analyzed, 51.24% were not cited or rarely cited, we conclude that the results reflect the paradigm of academic productivity. This paradigm creates a vicious cycle in which a study's references are chosen based on a high ranking in the struggle for privileged editorial space and, thus, are more likely to be cited.

By approaching the concept of academic productivity, the study presents and discusses the emphasis given to the quantitative production of research to the detriment of quality, innovation, and social relevance. This excessive focus on the number of publications can result in less bold and original research.

Concerning the pressure to publish, the study portrays the pressure in the academic-scientific community to publish as a way of evaluating academic success and obtaining funding for research. This circumstance corresponds to the growing prominence of scholarly article publication as a performance metric for researchers.

In the case of citations as a metric, the study highlights that the number of citations received by an article is considered important to assess the quality and impact of an academic work. However, it is also emphasized that the obsession with receiving citations can lead to questionable practices, such as self-promotion, dividing a study into several smaller articles (salami slicing) and searching for publications in high-impact journals.

In this context, the study addresses the problems of the academic publishing system, resulting in the flooding of databases with irrelevant articles, the emergence of predatory publishers, the conflict between ethics and research due to the pressure to produce "publishable" results at any cost, and on the negative effects on the mental health of university professors.

Finally, with equal relevance, uncited or rarely cited articles are addressed. Even though they may be important for the advancement of scientific knowledge, these articles often occupy editorial space without being useful for questionable reasons. These aspects illustrate the challenges and distortions present in the academic publication system, reinforcing the need to rethink the emphasis on the quantity of publications to the detriment of the quality and impact of research.

6. CONTRIBUTION TO AUTHORSHIP

Conceptualization: Priscila Rubbo, Luiz Alberto Pilatti.

Data curation: Eliandro Schvirck, Priscila Rubbo.

Formal Analysis: Eliandro Schvirck, Priscila Rubbo.

Investigation: Priscila Rubbo.

Methodology: Eliandro Schvirck, Priscila Rubbo, Luiz Alberto Pilatti.

Project administration: Priscila Rubbo.

Resources: Caroline Lievore, Eliandro Schvirck, Priscila Rubbo.

Supervision: Priscila Rubbo, Luiz Alberto Pilatti.

Validation: Caroline Lievore, José Roberto Herrera Cantorani, Luiz Alberto Pilatti.

Visualization: Priscila Rubbo, Luiz Alberto Pilatti.

Writing – original draft: Caroline Lievore, Eliandro Schvirck, Priscila Rubbo, Luiz Alberto Pilatti, José Roberto Herrera Cantorani.

Writing – review and editing: José Roberto Herrera Cantorani, Priscila Rubbo, Luiz Alberto Pilatti.

7. REFERENCES

Alcadipani, R. (2011). Resistir ao produtivismo: uma ode à perturbação acadêmica. *Cadernos Ebape.br*, 9(4), 1174–1178. DOI: https://doi.org/10.1590/S1679-39512011000400015.

Álvarez Bornstein, B., & Montesi, M. (2020). Funding acknowledgements in scientific publications: A literature review. *Research Evaluation,* 29(4), 469–488. DOI: https://doi.org/10.1093/reseval/rvaa038.

Asubiaro, T. (2019). How collaboration type, publication place, funding & author's role affect citations received by publications from Africa: A bibliometric study of LIS research from 1996 to 2015. *Scientometrics,* 120, 1261–1287. DOI: https://doi.org/10.1007/s11192-019-03157-1.

Bornmann, L., & Marx, W. (2020). Thomas theorem in research evaluation. *Scientometrics*. 123(1), 553–555. DOI: https://doi.org/10.1007/s11192-020-03389-6.

Castiel, L. D., & Sanz-Valero, J. (2007). Entre fetichismo e sobrevivência: o artigo científico é uma mercadoria acadêmica? *Cadernos de Saúde Pública*, 23(12), 3041–3050. DOI: https://doi.org/10.1590/S0102-311X2007001200026.

De Rond, M., & Miller, A. N. (2005). Publish or perish: Bane or boon of academic life? *Journal of Management Inquiry*, 14(4), 321–329. DOI: https://doi.org/10.1177/1056492605276850.

Demir, S. B. (2018). Predatory journals: Who publishes in them and why? *Journal of Informetrics*, 12(4), 1296–1311. DOI: https://doi.org/10.1016/j.joi.2018.10.008.

Denning, P. J. (1997). A new social contract for research. *Communications of the ACM*, 40(2), 132-134. DOI: https://doi.org/10.1145/253671.253755.

Dettori, J. R., Norvell, D. C., & Chapman, J. R. (2019). Measuring academic success: The art and science of publication metrics. *Global Spine Journal*, 9(2), 243–246. DOI: https://doi.org/10.1177/2192568219831003.

Diaz Ruiz, A., Orbe Arteaga, U., Rios, C., & Roldan Valadez, E. (2018). Alternative bibliometrics from the web of knowledge surpasses the impact factor in a 2-year ahead annual citation calculation: Linear mixed-design models' analysis of neuroscience journals. *Neurology India,* 68(1), 96–104. DOI: https://doi.org/10.4103/0028-3886.222880.

Dorta-González, P., & Dorta-González, M. I. (2013). Hábitos de publicación y citación según campos científicos: Principales diferencias a partir de las revistas JCR. *Revista Española de Documentación Científica*, 36(4), en012. DOI: http://dx.doi.org/10.3989/redc.2013.4.1003.

Else, H., & Van Noorden, R. (2021). The fight against fake-paper factories that churn out sham science. *Nature*, 591(7851), 516–519. DOI: https://doi.org/10.1038/d41586-021-00733-5.

Faber, F. T., Eriksen, M. B., & Hammer, D. M. G. (2021). Obsolescence of the literature: A study of included studies in Cochrane reviews. *Journal of Information Science*, 0(0). DOI: https://doi.org/10.1177/01655515211006588.

Fanelli, D. (2010). Do pressures to publish increase scientists' bias? An empirical support from US States Data. *PloS one*, 5(4), e10271. DOI: https://doi.org/10.1371/journal.pone.0010271.

Fávero, L. P., & Belfiore, P. (2021). *Manual de Análise de Dados: Estatística e Modelagem Multivariada com Excel®, SPSS® e Stata®*. Elsevier.

Furnham, A. (2021). Publish or perish: rejection, scientometrics and academic success. *Scientometrics*, 126(1), 843–847. DOI: https://doi.org/10.1007/s11192-020-03694-0.

Garfield, E. (2006a). Citation indexes for science: A new dimension in documentation through association of ideas. *International Journal of Epidemiology,* 35(5)*,* 1123–1127. DOI: https://doi.org/10.1093/ije/dyl189.

Garfield, E. (1998). Long-term vs. short-term journal impact: does it matter. *Physiologist,* 41(3), 113–115. PMID: 9652169.

Garfield, E. (2006b). The history and meaning of the Journal Impact Factor. *JAMA,* 295(1), 90–93. DOI: https://doi.org/10.1001/jama.295.1.90.

Gilyarevskii, R. S., Libkind, A. N., Libkind, I. A., & Bogorov, V. G. (2021). The obsolescence of cited and citing journals: Half-Lives and their connection to other bibliometric indicators. *Automatic Documentation and Mathematical Linguistics,* 55, 152–165. DOI: https://doi.org/10.3103/S0005105521040026.

Godoi, C. K., & Xavier, W. G. (2012). O produtivismo e suas anomalias. *Cadernos Ebape.br*, 10(2), 456–465. DOI: https://doi.org/10.1590/S1679-39512012000200012.

González-Sala, F., Silivestru, Y., Osca-Lluch, J., & Martí-Vilar, M. (2022). Journal self-citations included in the same thematic category of the Clarivate Journal Citation Reports (JCR): the case of the journal Adicciones. *Revista Española de Documentación Científica,* 45(3), e331. DOI: https://doi.org/10.3989/redc.2022.3.1886.

Guo, X., Li, X., & Yu, Y. (2021). Publication delay adjusted impact factor: The effect of publication delay of articles on journal impact factor. *Journal of Informetrics,* 15(1), 101100. DOI: https://doi.org/10.1016/j.joi.2020.101100.

Haba-Osca, J., Gonzalez-Sala, F., & Osca-Lluch, J. (2019). Education journals worldwide: an analysis of the publications included in the 2016 Journal Citation Reports (JCR). *Revista de Educación*, 383, 113–131. DOI: https://doi.org/10.4438/1988-592X-RE-2019-383-403.

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

Kacem, A., Flatt, J. W., & Mayr, P. (2020). Tracking self-citations in academic publishing. *Scientometrics*. 123(2), 1157–1165. DOI: https://doi.org/10.1007/s11192-020-03413-9.

Kaldas, M., Michael, S, Hanna, J., & Yousef, G. M. (2020). Journal impact factor: a bumpy ride in an open space. *Journal of Investigative Medicine,* 68(1), 83–87. DOI: http://dx.doi.org/10.1136/jim-2019-001009.

Kim, B., Kim, Y., & Kang, J. (2018). Analysis of the citation impact of national journals toward SCIE journals on JCR ranking. *Malaysian Journal of Library & Information Science,* 23(2), 1–24. DOI: https://doi.org/10.22452/mjlis.vol23no2.1.

Krampl, A. (2019). Relatórios de citação de periódicos. *Jornal da Associação de Bibliotecas Médicas*, 107(2), 280–283. DOI: https://doi.org/10.5195/jmla.2019.646.

Kulczycki, E., Hołowiecki, M., Taşkın, Z., & Krawczyk, F. (2021). Citation patterns between impact-factor and questionable journals. *Scientometrics*, 126, 8541–8560. DOI: https://doi.org/10.1007/s11192-021-04121-8.

Larivière, V., Macaluso, B., Archambault, E., & Gingras, Y. (2010). Which scientific elites? On the concentration of research funds, publications and citations. *Research Evaluation*, 19(1), 45–53. DOI: https://doi.org/10.3152/095820210X492495.

Lawrence, P. A. (2008). Lost in publication: how measurement harms science. *Ethics in science and environmental politics*, 8(1), 9–11. DOI: https://doi.org/10.3354/esep00079.

Leite, J. L. (2017). Publicar ou perecer: a esfinge do produtivismo acadêmico. *Revista Katálysis*, 20(2), 207–215. DOI: https://doi.org/10.1590/1982-02592017v20n2p207.

Liang, L., Rousseau, R., & Zhong, Z. (2013). Non-English journals and papers in physics and chemistry: bias in citations? *Scientometrics*, 95, 333–350. DOI: https://doi.org/10.1007/s11192-012-0828-0.

Liu, F., Hu, G., Tang, L. & Liu, W. (2018). The penalty of containing more non-English articles. *Scientometrics,* 114, 359–366. DOI: https://doi.org/10.1007/s11192-017-2577-6.

Machado, A. M. N., & Bianchetti, L. (2011). (Des) fetichização do produtivismo acadêmico: desafios para o trabalhador-pesquisador. *Revista de Administração de Empresas*, 51(3), 244–254. DOI: https://doi.org/10.1590/S0034-75902011000300005.

Martins, G. A. (2002). *Estatística geral e aplicada*. (2nd ed.). Editora Atlas.

Miranda, R., & Garcia Carpintero, E. (2019). Comparison of the share of documents and citations from different quartile journals in 25 research areas. *Scientometrics,* 121, 479–501. DOI: https://doi.org/10.1007/s11192-019-03210-z.

Moustafa, K. (2015). The disaster of the impact factor. *Science and engineering ethics*, 21(1), 139–142. DOI: https://doi.org/10.1007/s11948-014-9517-0.

Pranckutė, R. (2021). Web of Science (WoS) and Scopus: The titans of bibliographic information in today's academic world. *Publications*, 9(1), 12. DOI: https://doi.org/10.3390/publications9010012.

Price, D. J. S. (1976). A general theory of bibliometric and other cumulative advantage process. *Journal of the American Society for Information Science,* 27(5-6), 292–306. Available at: http://garfield.library.upenn.edu/price/pricetheory1976.pdf.

Rego, T. C. (2014). Produtivismo, pesquisa e comunicação científica: entre o veneno e o remédio. *Educação e Pesquisa*, 40(2), 325–346. DOI: https://doi.org/10.1590/S1517-97022014061843.

Rubbo, P., Pilatti, L. A., & Picinin, C. T. (2019). Citation of retracted articles in engineering: A study of the *Web of Science* database. *Ethics & Behavior*, 29(8), 661–679. DOI: https://doi.org/10.1080/10508422.2018.1559064.

Saeed, S., Jhanjhi, N. Z., Naqvi, M., Malik, N. A., & Humayun, M. (2019). Disparage the Barriers of Journal Citation Reports (JCR). *International Journal of Computer Science and Network Security,* 19(5), 156–175. Available at: http://paper.ijcsns.org/07_book/201905/20190523.pdf.

Sayão, L. F., Sales, L. F., & Felipe, C. B. M. (2021). Invisible science: publication of negative research results. *Transinformação*, 33, e200009. DOI: https://doi.org/10.1590/2318-0889202133e200009.

Smith, P. (1990). *Killing the spirit: Higher education in America*. Viking Penguin Publishing.

Suryani, I., Yaacob, A., & Aziz, N. H. A. (2015). Introduction sections of research articles with high and low citation indices. *Pertanika Journal of Social Sciences & Humanities*, 23(4), 1139–1152. Corpus ID: 64552981.

Sykes, C. J. (1988). *Profscam: Professors and the demise of higher education*. Regnery Publishing.

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

Tahamtan, I., Safipour Afshar, A. & Ahamdzadeh, K. (2016). Factors affecting number of citations: a comprehensive review of the literature. *Scientometrics,* 107, 1195–1225. DOI: https://doi.org/10.1007/s11192-016-1889-2.

Tang, L. (2013). Does "birds of a feather flock together" matter—Evidence from a longitudinal study on US–China scientific collaboration. *Journal of Informetrics,* 7(2). 330–344. DOI: https://doi.org/10.1016/j.joi.2012.11.010.

Tian, M., Su, Y., & Ru, X. (2016). Perish or publish in China: Pressures on young Chinese scholars to publish in internationally indexed journals. *Publications*, 4(2), 9. DOI: https://doi.org/10.3390/publications4020009.

Zuin, A. A. S., & Bianchetti, L. (2015). O produtivismo na era do "publique, apareça ou pereça": um equilíbrio difícil e necessário. *Cadernos de Pesquisa*, 45(158), 726–750. DOI: https://doi.org/10.1590/198053143294.