Do impact-factor journals make blacklisted journals visible? A systematic citation study

Emanuel Kulczycki[1*], Marek Hołowiecki[1], Zehra Taşkın[1], Franciszek Krawczyk[1]

[1] *Scholarly Communication Research Group, Adam Mickiewicz University in Poznań (Poland)*
* emek@amu.edu.pl

Abstract

Predatory publishing has recently become the most popular and controversial subject in academia. The main reason is the efforts of policymakers to draw a line between predatory and legitimate publishing. While the journals included in the citation indexes are accepted as 'legitimate' by policymakers, the journals in two lists (Beall's and Cabell's), one of which has not been updated for years, are considered blacklisted. Our main aim in this study is to reveal the contribution of the journals accepted as legitimate by policymakers to the visibility of predatory journals. The research in progress presents preliminary findings regarding the first stage of a two-part study. In the first stage, the contribution of impact-factor journals to the visibility of predatory journals will be revealed, and in the second, citations will be examined in terms of their contents. In this study, the indexes (JCR, ESCI or A&HCI) and impact factors of the citing journals are presented. According to the preliminary results of the study, the results confirm the necessity of citation content analysis.

Introduction

Predatory publishing is one of the most discussed topics regarding journal publishing in academia, which crosses over narrow fields of bibliometrics, scientometrics and academic-publishing studies (Frandsen, 2017; Xia et al., 2015). This topic related to publishing in so-called questionable or low-quality journals attracts much attention not only in academia but also outside it (Bohannon, 2013; Sorokowski et al., 2017). Predatory journals, accused of damaging science and diminishing the quality of scholarly communication and trust in science, are trying to be classified and listed. In recent years, the most famous attempt to list predatory journals was initiated by Jeffrey Beall, whose list (henceforth: Beall's list) gained attention from scientists and the media (Krawczyk & Kulczycki, 2020). The second well-known approach is done by the company Cabell's International (henceforth: Cabell's list), which not only lists predatory journals but also offers another product listing 'reputable' journals. Thus, journals listed on Beall's or Cabell's lists are called in this study blacklisted journals in contrast to whitelisted journals (listed for instance in reputable international indexes like Scopus or Web of Science Core Collection).

This study is the first extensive study looking at citations from Web-of-Science-indexed journals to papers published in blacklisted (or predatory) journals. Demir (2018) pointed out the big difference between the largest citation databases: Scopus indexes 53 predatory journals from Beall's lists, but Web of Science indexes only three such journals, and Somoza-Fernández et al. (2016) reported that this difference is small but still visible. This could also suggest a difference in citations to predatory journals in these databases, but previous attempts at analysing such citations were based mostly on Google Scholar data (due to an easier data-acquisition procedure) or Scopus.

In this paper, we will provide a systematic-citation study focussing on journals covered by the WoS products and on citers in terms of their country affiliation. The study aim is two-fold: The first is to analyse the visibility of blacklisted journals, and the second is to understand citation contexts. This paper is the first paper of a bigger endeavour aiming to investigate how papers published in blacklisted journals from the field of social sciences are cited by papers published by journals indexed in the Web of Science Core Collection. This paper will be followed by an upcoming study on the content-based analysis of citations of the blacklisted papers to evaluate the citations in terms of their content (meaning, purpose, shape, array). This is the key reason

why we focus on social-sciences journals here: evaluating the content of citations will require expertise in the field (Cano, 1989).

Data and Methods

We have used two blacklists: Beall's and Cabell's lists. Moreover, we have collected journals' ISSNs from their websites and used the ISSN Portal to find variants of journal titles and ISSNs as well as to provide data of the country of publishing. The data on citations of papers published in selected journals were obtained from the Web of Science Core Collection (WoS) using the Cited Reference Search. We focussed on three main WOS products: Journal Citation Reports (JCR) based on the Science Citation Index Expanded and the Social Sciences Citation Index, the Arts and Humanities Citation Index (A&HCI) and the Emerging Sources Citation Index (ESCI). We downloaded PDF files of either paper from blacklisted journals and WoS-indexed journals to collect and verify the country affiliations of the authors. Using the blacklisted journal websites, we collected the data on the number of papers published by each journal.

We have decided to include in the analysis only active journals that have fulfilled the set of criteria: they published at least one paper in each year of the 2012–2018 period, their websites were active at the moment of the start of this study (May 2019), were not indexed in WoS even only for one year in at least one of three selected WoS products during the years 2012–2018, and were classified (on expert-scope decisions of the authors of this study) as social-sciences journals according to their titles and aims and scopes published on journals' websites. In the final sample, we have 74 unique blacklisted journals (37 indexed in Beall's and 37 indexed in Cabell's). Ten of those journals are indexed on both lists.

We have prepared two datasets to analyse citations. The first one consists of the bibliographic data and PDF files of the papers published in the years 2012–2018 in the social-sciences blacklisted journals included in this study and were cited by journals indexed in WoS (henceforth: cited papers). The other dataset consists of data on the papers (i.e., all journal publication types) published in journals indexed in WoS in the years 2012–2019 (henceforth: citing papers) that cited papers from the first dataset. The articles from 9 of 74 journals were not cited by articles from WoS-indexed journals, or their PDF files were missing on journals' websites. This allows us to analyse 3,234 unique cited papers from 65 blacklisted journals and 5,964 unique citing papers (6,750 citations of cited papers) from 2,338 WoS journals. The list of analysed blacklisted journals is in Appendix.

Results

Share of cited papers from blacklisted journals

Sixty-five analysed journals published 25,146 papers between 2012 and 2018, of which 3,234 (13%) were cited by WoS journals.

Figure 1 shows the highest share, i.e., 19%, in 2012 and 2013. The mean number of papers published by a single journal in the analysed period is 53.5 (min=43, max=2,176). On average, 11% of papers published by a journal were cited by WoS journals. The highest shares were found for journals that published 1,748 and 259 papers. The shares are 36% (635 papers) and 35% (91 papers), respectively. In the analysed sample, 8,327 papers were published in journals listed in Beall's list (10.3% of them were cited) and 13,910 papers from journals indexed in Cabell's list (14.7% were cited). 5,587 papers were published by journals indexed in both Beall's and Cabell's lists.

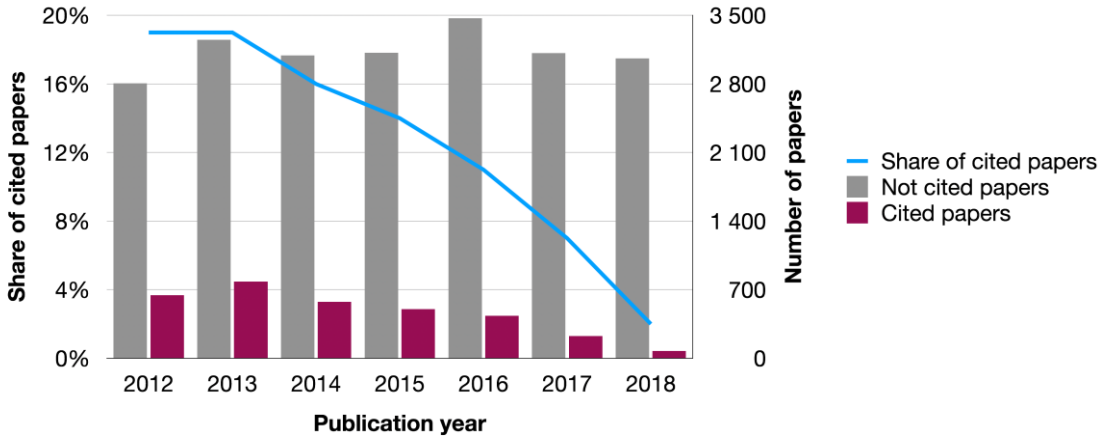

Figure 1. Share and Number of Cited Papers from Blacklisted Journals by WoS Journals

We checked whether 65 analysed journals were covered by Scopus, as this could potentially influence the share of citations. We found that five of 65 were or have been covered: one has been covered by the whole analysed period (24% of papers from this journal were cited), one removed from Scopus before the analysed period (23%), one covered and removed in the analysed period (36% papers), one covered in the last year of the analysed period (19%) and one covered before the analysed period and removed during the period (8%).

Web of Science journals citing blacklisted journals

We found that 2,338 unique WoS journals cited 3,234 blacklisted papers 6,750 times. The average number of citations per blacklisted journal from WoS journals is 2.88 (median=1, minimum=1, maximum=218). Half of the citations were from 261 journals. Eighty-nine of 2,338 journals cited papers from blacklisted journals at least 10 times, and four WoS journals cited over 100 blacklisted papers. In the analysed period, one WoS journal published 183 papers, which cited blacklisted papers from our sample 218 times (all except one published in one blacklisted journal). One of these blacklisted papers was cited 36 times by this WoS journal. We analysed in which WoS product (JCR, ESCI, A&HCI) a journal was indexed when a citing paper was published. We consider the publication year and whether a journal was included in a WoS product in the year in question. We found that 1,152 of 2,338 journals were indexed in ESCI, 35 in A&HCI and 1,047 in JCR. One-hundred and four journals that published 366 citing papers were neither in ESCI, A&HCI nor in JCR, which means that they were either in SCIE or in SSCI indexes but not yet JCR (e.g., waiting for calculation of their impact factor) or dropped from the indexes because of quality issues or manipulations such as citation stacking or excessive self-citation rates. Figure 2 shows how blacklisted journals were cited by WoS journals. Of the 6,750 citations, 2,502 (37%) were from JCR journals, 3,821 (56.6%) were from ESCI journals and 61 (0.9%) were from A&HCI journals.

Impact-factor journals citing blacklisted journals

The fact that a journal has a valid impact factor or is included in citation indexes is used by policymakers and managers to determine the level of that journal. We analysed the relationship between the impact-factor journals and their citations to blacklisted journals. To be able to make accurate statistical analyses, the impact factor of all journals cited in blacklisted journals were gathered with yearly changes. For example, if two articles in the same journal cited the blacklisted journals in 2018 and 2019, JCR 2017 and JCR 2018 were used. As a result, 1,600 impact factors for 1,047 IF journals were obtained.

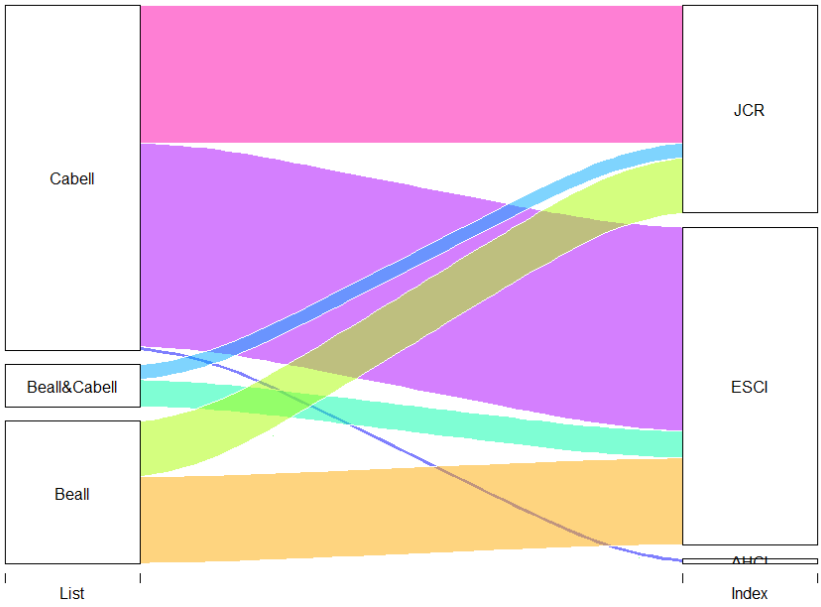

Figure 2. WoS Journals Citing Blacklisted Journals According to Blacklists

Before presenting the comparisons between impact factors and citations to blacklisted journals, it is worth mentioning those journals dropped from citation indexes. Twenty impact-factor journals that cited blacklisted journals 125 times were dropped from JCR or WoS for different reasons. Fifteen of them were dropped from the index without listing any unethical concerns. This means that coverage of the journals did not meet the WoS selection criteria (Clarivate, 2018). *Scientific World Journal* was suppressed from JCR based on citation stacking and four journals (*Business Ethics: A European Review*, *Environmental Engineering and Management Journal*, *Eurasia Journal of Mathematics Science and Technology Education* and *Industria Textila*) were dropped for their excessive self-citation rates. These five journals cited blacklisted journals 39 times. Furthermore, although they were not indexed in JCR and did not have an impact factor, 15 journals were excluded from ESCI after being indexed for a couple of years in ESCI. All these findings can be commented as questionable journals in WoS cited blacklisted journals; however, statistical tests did not confirm this comment.

The Spearman's Rho correlation coefficient shows that the correlation between journal impact factors and the number of citations to blacklisted journals is very low, at a 99% confidence level (r=0.090, p<0.001). Also, according to the Kruskal-Wallis test results, the differences between journal impact-factor quartiles and the number of citations to blacklisted journals were not significant (χ^2= 7.785, df=3, p=0.051). However, the Mann Whitney U test revealed that the only differences were found between Q1 and Q4 journals' number of citations to blacklisted journals (U=72661.500, Z=-2.648, p=0.008).

The impact-factor range of blacklisted journal citers is from 0 to 27.604 (mean=1.689, median=1.378, SD=1.471, 25%=0.745, 75%=2.252), while the minimum impact factor of the whole JCR between 2011 and 2018 is 0 and the maximum is 244.585 (mean=2.072, median=1.373, SD=3.310, 25%=0.704, 75%=2.462).

Eighty percent of the journals in JCR cited blacklisted journals only one time, and there is a significant difference between the impact factors of one-time citers and the others (U=174977.000, Z=-3.668, p<0.001). However, the surprising result is that the average impact factor of journals that cited blacklisted journals more than once is 1.896 (median=1.634), and this is higher than that of one-time citers (mean=1.639, median=1.318).

Table 1 shows the main features of 10 impact-factor journals that cited blacklisted journals more than 20 times.

Table 1. Ten Impact-Factor Journals that Cited Blacklisted Journals more than 20 Times

Journal name	*IF 2019*	*N of articles citing blacklisted journals*	*N of citations to blacklisted journals*	*Publisher country*	*Journal self-citation rate of the journal*
ACTA MEDICA MEDITERR	0.249	25	126	Italy	63.7%
SUSTAINABILITY-BASEL	2.576	56	56	Switzerland	38.9%
SAGE OPEN	0.715	16	45	USA	2.8%
SYSTEM	1.979	31	34	England	12.4%
FRONT PSYCHOL	2.067	24	26	Switzerland	15.5%
EGIT BILIM	0.493	19	25	Turkey	14.0%
PLOS ONE	2.740	25	25	USA	5.1%
COMPUT EDUC	5.296	21	22	England	11.6%
INT J BANK MARK	2.800	11	21	England	33.7%
COMPUT HUM BEHAV	5.003	20	21	USA	11.4%

All the test results on impact-factor journals prove that it is impossible to evaluate the blacklisted journals by looking at the impact factors or impact-factor percentiles of the journals because no pattern is identified. The impact factor is neither a descriptor of the quality of a paper nor the quality of citation. Therefore, it reveals the importance of content-based analysis in understanding the purpose of citations to blacklisted journals.

Discussion and Conclusions

The main aim of our study is to reveal the contributions of citation indexes, which are accepted as the authority in research evaluations, to the visibility of blacklisted journals, whose scientific levels are always considered quite low in academia. According to the results, 13% of the blacklisted articles were cited by Web of Science journals and 37% of these citations came from the impact-factor journals. If we accept being cited from authority-citation indexes as a tool for visibility, it is obvious that the indexes help the blacklisted journals to be visible regardless of the name of the index, whether SSCI, A&HCI or ESCI. The question to be asked at this point is: Do citations to the blacklisted journals make citation indexes questionable, or do these citations require a closer look at articles published in blacklisted journals? It is easy to accept all the papers published in blacklisted or questionable journals as low-quality, but without answering the question, it is difficult to draw a boundary for the definition between high and low quality. The findings show that there are no significant differences between the impact factors and the number of citations to the blacklisted journals. All the statistical tests conducted using journal-level statistics in our study confirm the emerging need to analyse the citations to blacklisted journals at the article level.

This is the first study in which a large-scale analysis of citations to predatory journals is conducted using WoS. When compared to the different results present in the citation studies based on Scopus, it is difficult to assess differences in terms of citing predatory journals in these two databases. However, taking into account that only 13% of articles in our study are cited, we can be sure that citations to the predatory journal are much more frequent in Google Scholar because it was reported there that 43% of articles were cited (Björk et al., 2020). Since we did not assess the quality of cited papers published in blacklisted journals, there are two possible interpretations of the main result of our study: 1) up to 13% of worthless articles in predatory journals can still leak to the mainstream literature legitimised by WoS, and 2) up to 13% of papers published in blacklisted journals are somehow important for developing scholarly legitimate discussion in social science. Unlike Oermann et al. (2020), we are not so sure that the important conclusion of the studies on predatory journals is to stop citing them completely. We prefer to leave the question raised by the result of our study open.

Our results indicate that there is no connection between the value of JIF of a given journal and this journal's citations to predatory journals. Although the number of such citations is relatively small, it could be another argument against treating JIF as a measure of journals' quality.

Acknowledgements

The work was financially supported by the National Science Centre in Poland (Grant Number UMO-2017/26/E/HS2/00019).

Appendix

The list of 65 analysed blacklisted journals is available here:
https://figshare.com/articles/dataset/Appendix_–_List_of_65_blacklisted_journals/13560326

References

Björk, B.-C., Kanto-Karvonen, S., & Harviainen, J. T. (2020). How Frequently Are Articles in Predatory Open Access Journals Cited. *Publications*, *8*(2), 17. https://doi.org/10.3390/publications8020017

Bohannon, J. (2013). Who's Afraid of Peer Review? *Science*, *342*(6154), 60–65. https://doi.org/10.1126/science.342.6154.60

Cano, V. (1989). Citation behavior: Classification, utility, and location. *Journal of the American Society for Information Science*, *40*(4), 284–290.

Clarivate. (2018, June 27). *Web of Science: Editorial statement about dropped journals*. https://support.clarivate.com/ScientificandAcademicResearch/s/article/Web-of-Science-Editorial-statement-about-dropped-journals?language=en_US

Demir, S. B. (2018). Scholarly databases under scrutiny. *Journal of Librarianship and Information Science*, 096100061878415. https://doi.org/10.1177/0961000618784159

Frandsen, T. F. (2017). Are predatory journals undermining the credibility of science? A bibliometric analysis of citers. *Scientometrics*, *113*(3), 1513–1528. https://doi.org/10.1007/s11192-017-2520-x

Krawczyk, F., & Kulczycki, E. (2020). How is open access accused of being predatory? The impact of Beall's lists of predatory journals on academic publishing. *The Journal of Academic Librarianship*, 102271. https://doi.org/10.1016/j.acalib.2020.102271

Oermann, M. H., Nicoll, L. H., Ashton, K. S., Edie, A. H., Amarasekara, S., Chinn, P. L., Carter-Templeton, H., & Ledbetter, L. S. (2020). Analysis of Citation Patterns and Impact of Predatory Sources in the Nursing Literature. *Journal of Nursing Scholarship*, *52*(3), 311–319. https://doi.org/10.1111/jnu.12557

Somoza-Fernández, M., Rodríguez-Gairín, J.-M., & Urbano, C. (2016). Presence of alleged predatory journals in bibliographic databases: Analysis of Beall's list. *El Profesional de La Información*, *25*(5), 730. https://doi.org/10.3145/epi.2016.sep.03

Sorokowski, P., Kulczycki, E., Sorokowska, A., & Pisanski, K. (2017). Predatory journals recruit fake editor. *Nature*, *543*(7646), 481–483. https://doi.org/10.1038/543481a

Xia, J., Harmon, J. L., Connolly, K. G., Donnelly, R. M., Anderson, M. R., & Howard, H. A. (2015). Who publishes in "predatory" journals? *Journal of the Association for Information Science and Technology*, *66*(7), 1406–1417. https://doi.org/10.1002/asi.23265