

Engineering Literature Retractions: Applications to Scholarly Communications Training

Mrs. Daniela Solomon, Case Western Reserve University

Daniela Solomon is Research Services Librarian for Engineering & Chemistry at Kelvin Smith Library, Case Western Reserve University. She is interested in standards, bibliometrics, data management, and library instruction.

Christopher Heckman, St. Mary's College of Maryland

Research and Instruction Librarian at St. Mary's College of Maryland.

Engineering Literature Retractions: Applications to Scholarly Communications Training

Abstract

This study presents the preliminary results of an investigation of retracted publications in the engineering literature to identify scholarly communications topics that librarians need to address in instruction settings. We used Elsevier's Scopus abstract and citation database to collect engineering publications marked as "retracted articles" as of November 2019. We found that the most common reasons for retraction were plagiarism and authorship issues, while copyright infringement, errors, and lack of reproducible results were less common. The results of this study will be of particular interest to educators teaching scholarly communications practices to engineering graduate students by illustrating the common mistakes found in the engineering literature.

Introduction

For scientific publications to contribute to the advancement of knowledge, it is necessary that researchers be objective, complete, fair, and accurate in reporting the results of their work, and to act with integrity throughout the publishing process. The authors of scholarly publications bear the responsibility for adhering to ethical norms, including those related to authorship and attribution, disclosure of conflict of interest, and the integrity of the peer-review process. As many graduate engineering programs require journal publications and/or conference presentations for graduation, it is critical that students learn about best practices in scholarly communications and are cognizant of the ethical aspects of scientific publishing in their field.

The focus of this study is to find common reasons for retraction in engineering scholarship in order to identify points of need in education about the publishing cycle. This will help educators to teach engineering students seeking to publish how to avoid common publishing pitfalls. Graduate programs play a critical role in preparing students for an academic career. Together with faculty, librarians can work to give students a grounding in the fundamentals of publishing ethics due to their combination of scholarly communications expertise and hands-on experience with the information behavior of students at their institutions.

Literature review

In recent years, academic libraries have reported a major shift in focus towards scholarly communications and research management services [1]. Craft and Harlow [2] observed increased requests from graduate students for scholarly communications training in a variety of topics, with the top choices being “publishing tips” and “post-dissertation publishing” [3]. This demand for training reflects the need among students for guidance in navigating the publication process. Many doctoral programs, especially in the sciences and engineering, require students to publish in order to graduate. As relative novices in the complicated, often opaque arena of academic publishing, many of these students struggle with the task and would benefit from guidance in navigating the publication process. The analysis of retracted publications can help identify the critical scholarly communications issues researchers struggle with. Focusing scholarly communication training on the mistakes that lead to retractions could help researchers avoid upsetting and damaging situations in their future careers.

Retraction is defined as “a mechanism for correcting the literature and alerting readers to articles that contain such seriously flawed or erroneous content or data that their findings and conclusions cannot be relied upon” [4]. Papers can be retracted for a number of reasons, including academic dishonesty and unintentional errors. Over the last several decades, there has been observed a worrisome phenomenon of substantial increase in the number of articles retracted from academic journals across a wide range of disciplines [5] - [7]. While the major cross-disciplinary studies of the increasing retraction phenomenon were conducted between eight and ten years ago, there have been more recent studies that have investigated retractions in the medical [8-10] and dental [11] fields, and investigations into scientific studies conducted in China [12] that have found higher rates of retraction than in previous decades. Rubbo et al.’s paper on retractions in the engineering literature also found higher rates of retraction between 2008 and 2015 than in prior decades [13]. “Publishing misconduct” is determined to be the most common reason for retraction in Grieneisen and Zhang’s study, a category that can include plagiarism of other authors, data fabrication, falsification, or omission, faked peer-review, and listing of false authors [7]. Redundant publication, a practice in which researchers publish all or part of the same paper more than once is also a common reason for retraction [8]-[9], [11], [13]. Notably, many studies describe redundant publication as “self-plagiarism,” and it is frequently found to be one of the top reasons for retraction [7]-[8], [13], [15]. Papers are sometimes retracted for reasons other than academic dishonesty, such as the discovery of an error in a study’s methodology or analyses [5], [16] or the inability of other researchers to reproduce the results.

When an article is retracted, the publisher issues a retraction notice to announce the article’s withdrawal. However, the retracted article is not removed from the publisher’s

website or from any databases that index the publication. Instead, the retraction is indicated in a variety of ways depending on the source. The most detailed retraction notice is usually available on the publisher's website. Despite the efforts of organizations like the Committee on Publication Ethics, policies vary significantly in terms of how retracted status is indicated, whether the retraction itself is cataloged separately from the original article, and whether an explanation for the retraction is made available [4]. As a result, it is often difficult to identify retracted papers and to determine the reason for their retraction [17] - [19].

The inconsistency with which publisher websites and research databases indicate the retracted status of an article represents an obvious problem, since the retracted articles potentially may continue to influence subsequent research and thinking on a given topic and distort the scientific record [16], [20] - [23]. The extent of this distortion depends on the reason for the retraction so it is important that scholars are aware of the common reasons for retraction in the scholarly literature of their field and of the ways various sources of information indicate retracted publications.

While there have been many studies written on retraction in medical fields, there has only been one study written to date on retractions in the engineering field [16]. In that study, Rubbo et al focused on creating a comprehensive profile of retracted engineering publications indexed by Web of Science, identifying 238 retracted articles published between 1945 and 2015. The most common reason for retraction in that study was “unethical research,” defined by the authors as “publication without the author’s consent, lack of references in the article, error in figure description, error in article editing, unacknowledged financing source, data misuse, and data dispute.”

The present study investigates retractions in engineering literature indexed by Scopus in order to provide educators with an overview of some common reasons for retraction in the engineering literature.

Methodology

In November 2019, we searched Scopus for engineering journal articles and conference papers marked as retracted by the end of 2018. We chose to use Scopus because its coverage of journal titles and conferences in engineering and its overall citation numbers are better than that of Web of Science [24]. Retraction notices available in Scopus can take various forms such as “retraction to,” “retraction of,” “retracted article,” “retraction,” or “article retraction.” We limited our search to articles with the phrase “retracted article” in the title field and used the Subject filter to limit the results to “Engineering” in order to run a pilot test for our analysis. We retrieved a total of 3,216 documents that we then exported to an Excel spreadsheet. These documents were

labeled as article, erratum, editorial, note, retracted, or conference paper. We consolidated the document types into *journal article* and *conference paper* based on the publication type. After eliminating one journal article that was erroneously indexed in the Engineering category, we conducted analyses on the remaining 3,215 documents of which there were eighty journal articles and 3,135 conference papers.

For each retracted journal article, we collected the publication year and retraction year, the total number of citations, the number of citations after retraction, and the retraction note on Scopus, the publishers' website and Google Scholar. Analysis of the data allowed us to determine how quickly journal articles were retracted, the most common reasons for retractions, and how clearly the retraction status was marked in Scopus, the publisher's website, and Google Scholar. The analysis of citations received by journal articles post-retraction illustrates whether retracted publications continue to influence subsequent research and whether the inconsistencies in marking retraction status have an impact on the overall citation performance of retracted publications.

For each retracted conference paper, we collected the number of citations, publisher, and the reason for retraction, but a comprehensive analysis was not possible because very few details on the reasons for retraction were available. Intriguingly, only IEEE conference papers were retracted during this period.

Results and Discussion

The eighty retracted journal articles identified were published by forty-three journal titles from twelve different publishers. Most of these journal titles (27 out of 43) had only one retraction each but there were also several journals with multiple retractions (see Figure 1). This suggests that, while there is room for improvement in editorial practices, the issues that led to these retractions are likely not systemic problems.

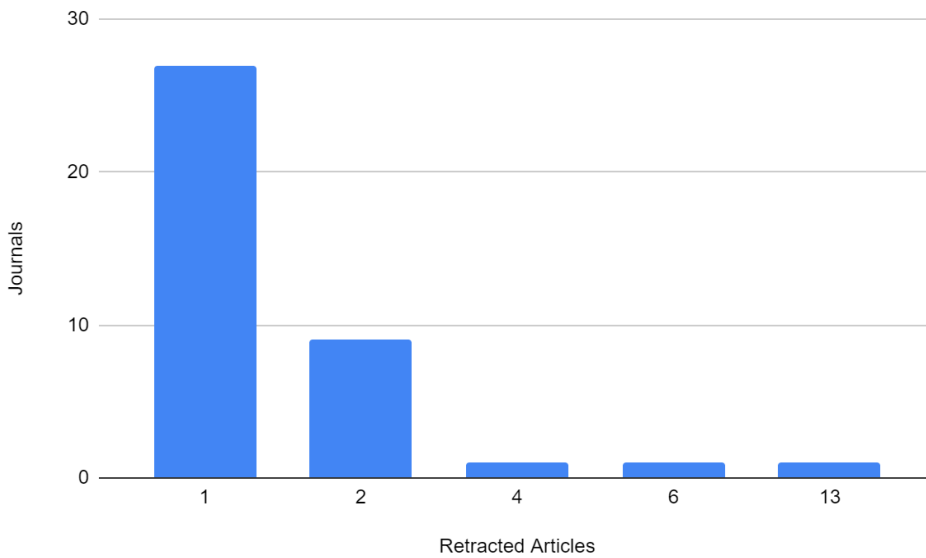


Figure 1. Number of retracted articles per journal titles

The 3,135 retracted conference papers identified were published in forty-eight conference proceedings, all by IEEE. It was intriguing to observe that only IEEE conference papers were included in this dataset, which raises questions about the reliability of the conference papers data. We found it troublesome that several conferences have retracted a large number of papers or even the entirety of the proceedings and can only speculate about why this happened. However, it was encouraging to see that after the spike in 2011 the number of retracted conference papers has decreased significantly (see Figure 2). A more in depth analysis of the reasons for retractions for conference papers was not possible because IEEE does not list any reason beyond “violation of IEEE's Publication Principles.

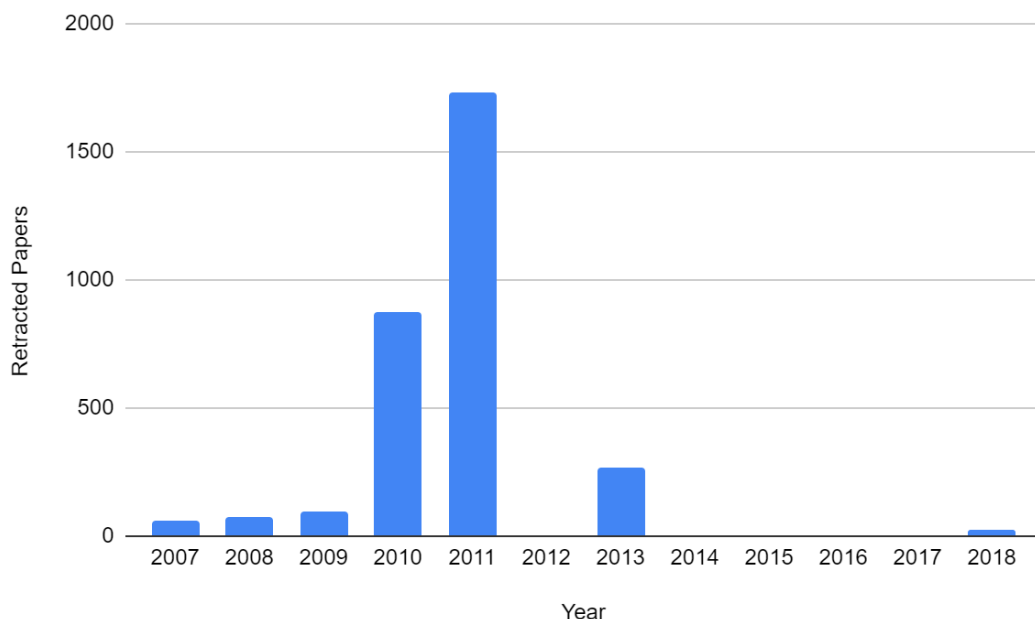


Figure 2. Retracted conference papers over the time period of the study

The analysis of retraction notices for journal articles revealed the common themes for journal publications. Similar to prior retraction studies, we found that the top reasons for retraction were self-plagiarism and plagiarism (see Figure 3). Self-plagiarism, or duplicate publication, is when an author publishes a paper that he or she has previously published, in a different venue or sometimes in a different language, without updating the content or crediting the original publication. It is important for graduate students to understand that this practice is unethical. As engineering students often present the results of their work at conferences and then go on to publish those same results in journals, theses, or dissertations, it is important to be aware of the need for self-citation. As educators, it is our responsibility to make students aware of the expectation that these publications include sufficient novel content to be more than a mere copy of their previous work, and that they need to cite their own data and previous publications in order to avoid plagiarism.

The second most common reason for retraction was the authorship category, which included several issues related to disputes between the co-authors. The list of issues included disputed authorships, listing of coauthors without consent, and differences in data interpretation. Authorship is critical to an academic career, and it is necessary that students understand both the importance of a correct record of authorship and their rights and responsibilities as authors. The first step is to understand what constitutes authorship so that the list of co-authors includes only those with substantial

contributions to the publication. The order in which authors are listed is also important, and students should be aware of the practice in their discipline.

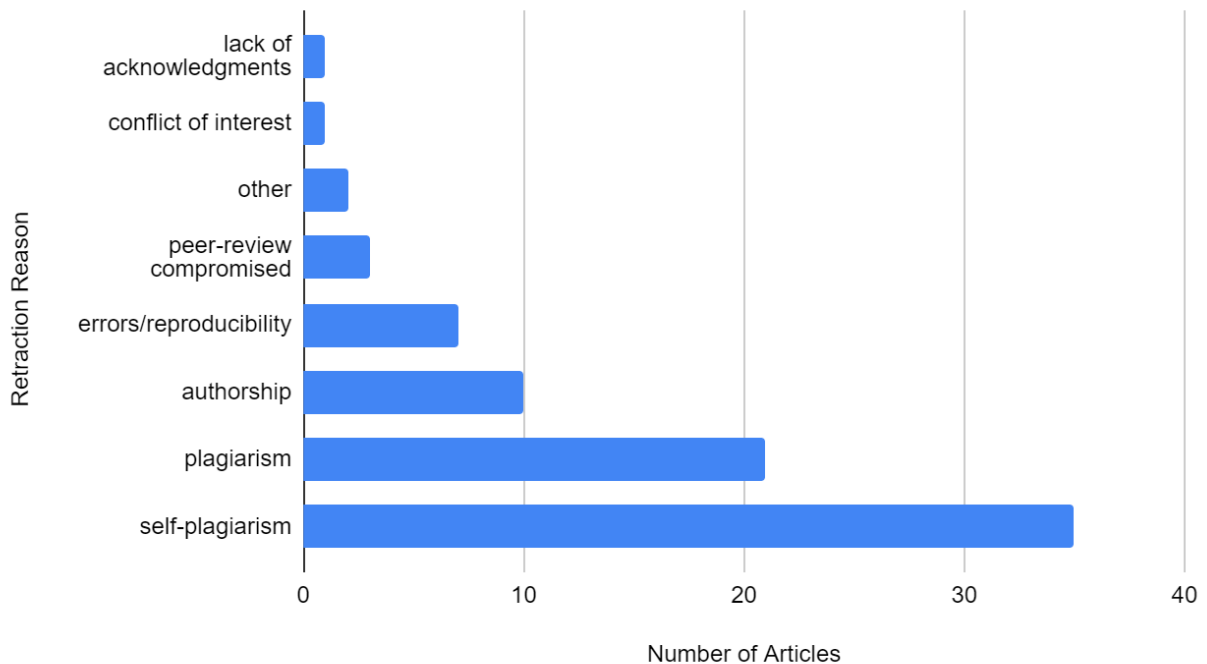


Figure 3. Retraction reasons for journal articles

Errors and reproducibility issues, compromised peer-review, lack of acknowledgments, and conflict of interest were the least common reasons for retractions. Unintentional errors were not as common as plagiarism and authorship issues, and, depending on the type of error, these can sometimes be corrected without retracting the publication. In our sample, the only retraction for reproducibility issues was due to missing lab notebooks. It is critical for engineering students to learn the best practices for maintaining lab notebooks and to apply data management principles to document their research and curate the resulting data. Research funders may also require these practices so it is important for students to understand the value of good data management. The retraction notices for the case of compromised peer-review did not include specific details on the nature of the problem. Similarly, there were no details provided for the articles retracted for lack of acknowledgements and conflict of interest. However, it was encouraging to observe that there were no retractions due to data fabrication or image manipulations. These represent significant breaches of research ethics, and while it is reassuring not to see examples of them in our sample, it is nonetheless important to ensure that students are aware of their severity.

Identifying the reason for retractions was time-consuming and not straight-forward. Articles marked with "retracted article" in Scopus do not necessarily include a clear statement that describes the reason for withdrawal. Following the DOI link on an article record in Scopus, one can eventually discover the retraction notice on the publisher's website in cases where it has been made available. The existing retraction notes are often vague or misleading and do not always include sufficient information to fully understand the context of the retraction. While checking for all of the journal article titles in this sample on Google Scholar, we found that almost half were either not marked as retracted at all or required several steps of searching to uncover their retracted status. This is worrisome since many engineering students prefer to use Google Scholar instead of databases when searching for literature. We also observed that some publishers share better metadata with Google Scholar than others, resulting in better identification of their retracted articles.

One of the consequences of this ambiguity is that scholars may have difficulty identifying a retracted article and continue to cite it. We found 186 citations to the retracted journal articles included in this study. These citations were accumulated by thirty-six articles, which represent slightly less than half of the retracted articles (45%), meaning that the rest were not cited at all (see Figure 4).

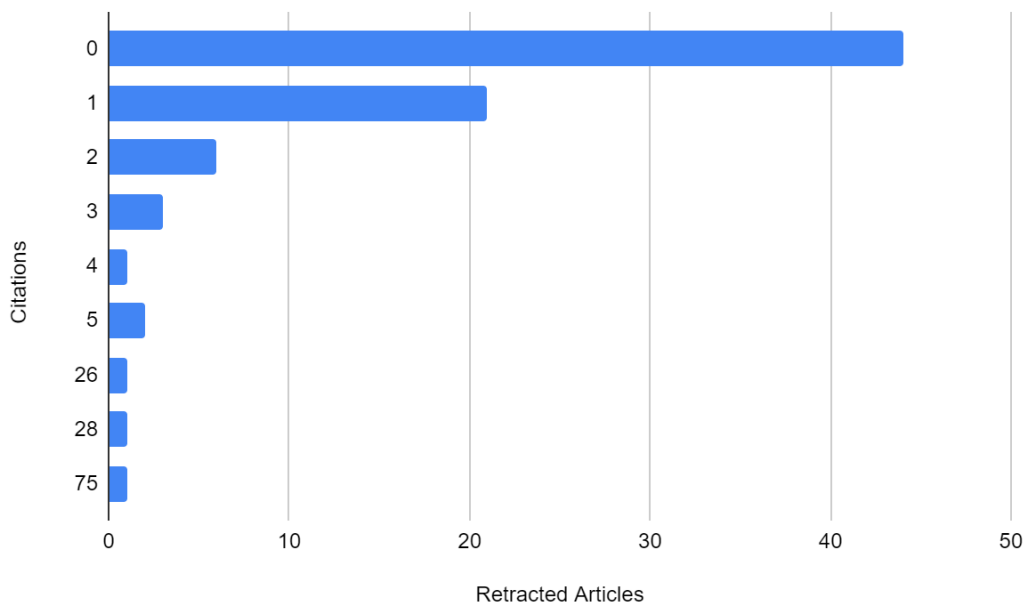


Figure 4. Total citation numbers for retracted journal articles

The percentage of cited conference papers is even smaller, with only 22.5% of the 3,135 total papers being cited and the most highly cited paper receiving a total of fourteen citations (see Figure 5).

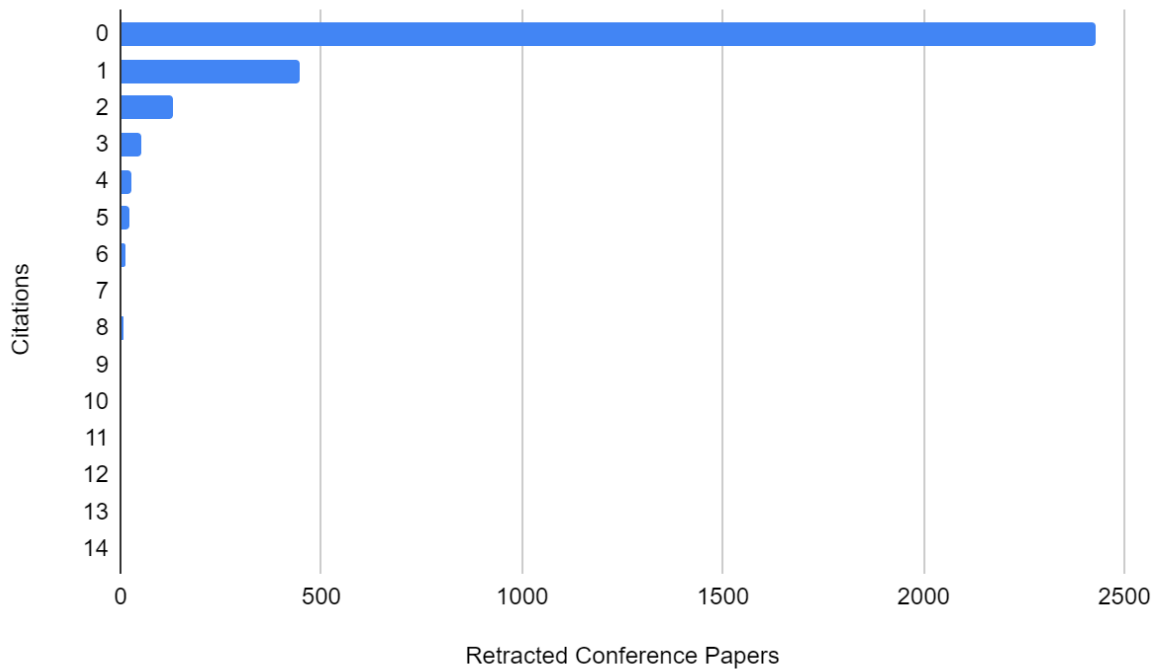


Figure 5. Total citation numbers for retracted conference papers

The thirty-six cited journal articles received a total of seventy-two citations before retraction and 144 citations post-retraction. Post-retraction citations were accumulated by thirty of the thirty-six articles and ranged from one to twenty-four. Most of the thirty journal articles cited post retraction (90%) accumulated fewer than five citations after retraction, while three articles received a high number of citations (see Figure 6).

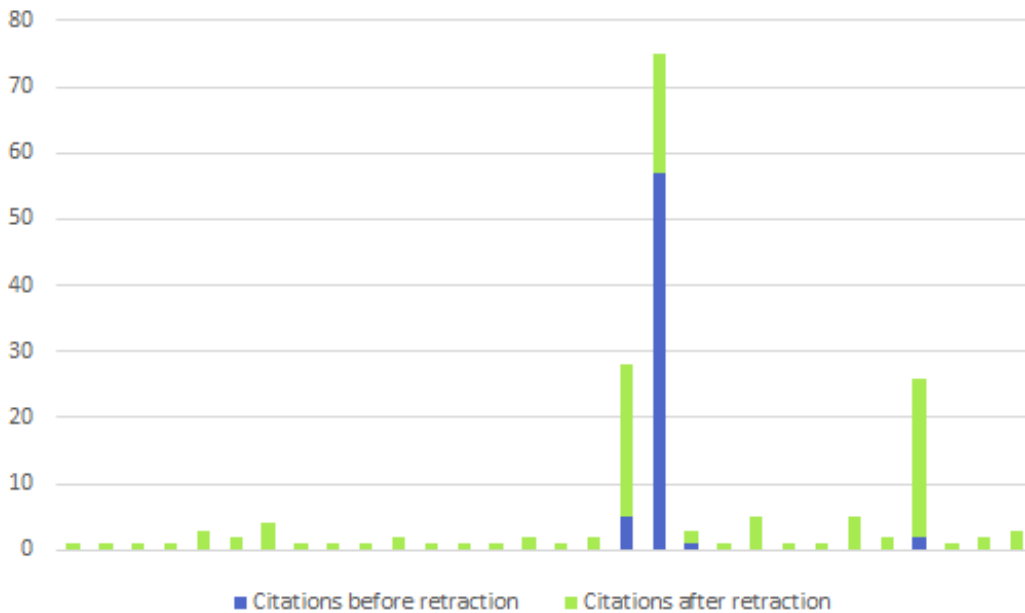


Figure 6. Citations pre- and post-retraction for journal articles

The highest citation number received by an article was seventy-five: fifty-seven before retraction and eighteen afterward. The article was retracted four years after publication when other scholars challenged the results, and the authors were unable to produce lab notebooks. Its retracted status is not clearly indicated on Google Scholar. The next highly cited article received a total of twenty-eight citations. It was retracted within a year of publication but twenty-three of the twenty-eight citations occurred post-retraction. The reason for retraction was an error in the results section, and it is alarming that so many scholars did not observe it. This article is not clearly marked as retracted on Google Scholar either. Both these articles were published in *Nature Biotechnology*, a highly respected journal with a high impact factor. The third most highly cited article was published by *Strength of Materials* which received a total of twenty-six citations. It was retracted within a year of publication due to self-plagiarism, and twenty-four of the twenty-six total citations occurred post-retraction. Its retracted status is not marked on Google Scholar.

Conclusion

The U.S. Office of Research Integrity defines misconduct as “...*fabrication, falsification, or plagiarism in proposing, performing or reviewing research, or in reporting research results.*” Similar to retraction research in other disciplines, this study found that the most common reason for retraction in engineering literature is self-plagiarism. These results

suggest that engineering students may need more explicit and robust instruction in publishing ethics and scientific integrity. Equipped with this knowledge, engineering librarians can plan for focused scholarly communication training, alone or in collaboration with other library departments. Due to their expert knowledge of engineering sources and literature, engineering librarians are well positioned to contribute to the scholarly communication system practice by the engineering scientists and graduate students. While undergraduate students do not conduct and publish research as often as graduate students, introducing them to the principles of research ethics early in their academic careers can help them to develop the habits they will need to be effective researchers later.

As reflected by the high number of retractions due to plagiarism and self-plagiarism, the topic most misunderstood is credit attribution, and students would benefit from instruction in the ethics and best practice of giving credit. A better understanding of the authorship concept would help students avoid upsetting situations and possible damage to their career. Instruction in the ethics of the research cycle would also be helpful in developing an appreciation for the value of proper data recording and preservation, which would help prevent future reproducibility challenges. Ideally, instruction in these areas should be targeted at both undergraduate and graduate students with the goal to build practical knowledge for an ethical approach to scholarly communications overall.

Despite the fact that less than half of the retracted journal articles were cited overall, with most papers accumulating a small total number of citations, the number of citations accumulated post-retraction is higher than the number of citations accumulated pre-retraction. This difference may be the result of the difficulty in identifying retracted status and/or length of the time between publication and retraction. It was interesting to observe that two of the retracted papers published by a well-regarded journal accumulated the larger number of citations pre- and post retractions. The reproducibility issues of one of these articles took several years to be identified and that could explain the accumulation of citations. The errors in the other article were discovered within one year, but it was discouraging to see how many scholars did not observe it sooner and cited the paper. Other unintentional errors were identified within two years of the publication date, but the articles accumulated no citations. This situation provides a strong argument for students' need to develop excellent critical reading skills. Librarians can contribute to these efforts through information literacy instruction and research support.

Limitations to this study include the small sample of publications selected and the lack of details available for the retractions of IEEE conference proceedings. In addition, the incongruous policies among publishers on how much information to include in

retractions, how to indicate an article's status as retracted, and where to make retraction notices available complicates the conclusions we can draw. However, the results are similar to those of prior retraction studies and support our hypothesis that retraction analysis can help to identify critical topics for scholarly communication education. Further investigation is needed to create a comprehensive understanding of Scopus's policy for indexing retraction notices, to determine the differences in labeling retractions between Web of Science, Scopus, Google Scholar, and other increasingly popular scholarly search engines like Dimensions and Microsoft Academics, and to measure the impact of these practices on the scholarly record.

Acknowledgements

We thank the reviewers for their feedback and comments on our manuscript.

References

- [1] A. Hutchinson, *Science Libraries in the Self Service Age: Developing New Services, Targeting New Users*. Chandos Publishing, 2018.
- [2] A. R. Craft and S. Harlow, "Scholarly communications training: Professional development for the next generation of scholars," *Ser. Rev.*, vol. 46, no. 3, pp. 175–183, Jul. 2020, doi: 10.1080/00987913.2020.1806651.
- [3] B. L. Fong, M. Wang, K. White, and R. Tipton, "Assessing and serving the workshop needs of graduate students," *J. Acad. Librariansh.*, vol. 42, no. 5, pp. 569–580, Sep. 2016, doi: 10.1016/j.acalib.2016.06.003.
- [4] COPE Council, "COPE Retraction guidelines — English.," Version 2, Nov. 2019. [Online]. Available: <https://doi.org/10.24318/cope.2019.1.4>.
- [5] R. G. Steen, "Retractions in the scientific literature: is the incidence of research fraud increasing?," *J. Med. Ethics*, vol. 37, no. 4, pp. 249–253, Apr. 2011, doi: 10.1136/jme.2010.040923.
- [6] R. G. Steen, A. Casadevall, and F. C. Fang, "Why has the number of scientific retractions increased?," *PLOS ONE*, vol. 8, no. 7, p. e68397, Jul. 2013, doi: 10.1371/journal.pone.0068397.
- [7] M. L. Grieneisen and M. Zhang, "A comprehensive survey of retracted articles from the scholarly literature," *PLoS ONE*, vol. 7, no. 10, p. e44118, Oct. 2012, doi: 10.1371/journal.pone.0044118.

- [8] B. D. Cassao, F. A. M. Herbella, F. Schlottmann, and M. G. Patti, "Retracted articles in surgery journals. What are surgeons doing wrong?," *Surgery*, vol. 163, no. 6, pp. 1201–1206, Jun. 2018, doi: 10.1016/j.surg.2018.01.015.
- [9] S. Huh, S. Y. Kim, and H.-M. Cho, "Characteristics of retractions from Korean medical journals in the KoreaMed database: A bibliometric analysis," *PLoS ONE*, vol. 11, no. 10, pp. 1–7, Oct. 2016, doi: 10.1371/journal.pone.0163588.
- [10] R. M. Varnier Rodrigues de Almeida, F. Catelani, A. J. Fontes-Pereira, and N. de S. Gave, "Retractions in general and internal medicine in a high-profile scientific indexing database," *Sao Paulo Med. J.*, vol. 134, no. 1, pp. 74–78, Feb. 2016, doi: 10.1590/1516-3180.2014.00381601.
- [11] T. E. Nogueira, A. S. Gonçalves, C. R. Leles, A. C. Batista, and L. R. Costa, "A survey of retracted articles in dentistry," *BMC Res. Notes*, vol. 10, no. 1, p. 253, Jul. 2017, doi: 10.1186/s13104-017-2576-y.
- [12] L. Lei and Y. Zhang, "Lack of Improvement in Scientific Integrity: An analysis of WoS retractions by Chinese researchers (1997-2016)," *Sci. Eng. Ethics*, vol. 24, no. 5, pp. 1409–1420, Oct. 2018, doi: 10.1007/s11948-017-9962-7.
- [13] P. Rubbo, C. L. Helmann, C. B. dos Santos, and L. A. Pilatti, "Retractions in the engineering field: A study on the Web of Science database," *Ethics Behav.*, vol. 29, no. 2, pp. 141–155, Feb. 2019, doi: 10.1080/10508422.2017.1390667.
- [14] J. Wang, J. C. Ku, N. M. Alotaibi, and J. T. Rutka, "Retraction of neurosurgical publications: A systematic review," *World Neurosurg.*, vol. 103, pp. 809+, Jul. 2017, doi: 10.1016/j.wneu.2017.04.014.
- [15] R. M. V. R. Almeida, K. de Albuquerque Rocha, F. Catelani, A. J. Fontes-Pereira, and S. M. R. Vasconcelos, "Plagiarism allegations account for most retractions in major Latin American/Caribbean databases," *Sci. Eng. Ethics*, vol. 22, no. 5, pp. 1447–1456, Oct. 2016, doi: 10.1007/s11948-015-9714-5.
- [16] J. D. Mena, M. Ndoeye, A. J. Cohen, P. Kamal, and B. N. Breyer, "The landscape of urological retractions: the prevalence of reported research misconduct," *BJU Int.*, vol. 124, no. 1, pp. 174–179, Jul. 2019, doi: 10.1111/bju.14706.
- [17] C. R. Madlock-Brown and D. Eichmann, "The (lack of) impact of retraction on citation networks," *Sci. Eng. Ethics*, vol. 21, no. 1, pp. 127–137, Feb. 2015, doi: 10.1007/s11948-014-9532-1.

- [18] C. Bakker and A. Riegelman, "Retracted publications in mental health literature: discovery across bibliographic platforms," *J. Librariansh. Sch. Commun.*, vol. 6, no. 1, p. eP2199, Jan. 2018, doi: 10.7710/2162-3309.2199.
- [19] J. A. Teixeira da Silva and J. Dobránszki, "Notices and policies for retractions, expressions of concern, errata and corrigenda: Their importance, content, and context," *Sci. Eng. Ethics*, vol. 23, no. 2, pp. 521–554, Apr. 2017, doi: 10.1007/s11948-016-9769-y.
- [20] J. Bar-Ilan and G. Halevi, "Post retraction citations in context: a case study," *Scientometrics*, vol. 113, no. 1, pp. 547–565, Oct. 2017, doi: 10.1007/s11192-017-2242-0.
- [21] H. Bornemann-Cimenti, I. S. Szilagyi, and A. Sandner-Kiesling, "Perpetuation of retracted publications using the example of the Scott S. Reuben Case: Incidences, reasons and possible improvements," *Sci. Eng. Ethics*, vol. 22, no. 4, pp. 1063–1072, Aug. 2016, doi: 10.1007/s11948-015-9680-y.
- [22] R. Jan and T. Zainab, "The impact story of retracted articles altmetric it!," in *5th Int. Sym. on Emerging Trends and Tech. in Lib. and Inf. Serv. (ETTLIS)*, Feb. 2018, pp. 1–5, doi: 10.1109/ETTLIS.2018.8485245.
- [23] J. L. Furman, K. Jensen, and F. Murray, "Governing knowledge in the scientific community: Exploring the role of retractions in biomedicine," *Res. Policy*, vol. 41, no. 2, pp. 276–290, Mar. 2012, doi: 10.1016/j.respol.2011.11.001.
- [24] A. Martín-Martín, M. Thelwall, E. Orduna-Malea, and E. Delgado López-Cózar, "Google Scholar, Microsoft Academic, Scopus, Dimensions, Web of Science, and OpenCitations' COCI: a multidisciplinary comparison of coverage via citations," *Scientometrics*, vol. 126, no. 1, pp. 871–906, Jan. 2021, doi: 10.1007/s11192-020-03690-4.