Scientometrics (2017) 110:365–370
DOI 10.1007/s11192-016-2178-9

Why do some retracted papers continue to be cited?

Jaime A. Teixeira da Silva[1] · **Helmar Bornemann-Cimenti**[2]

Received: 17 October 2016 / Published online: 11 November 2016

Abstract Perpetuation of retracted publications is an ongoing and even increasing problem in the scientific community. In addition to the direct distortion of scientific credibility, the use of retracted findings for interpretation und discussion in subsequent publications poses the risk of drawing false and, for example in medical research, even harmful conclusions. One major contributor to this development is that many authors are not aware of the retraction status of a paper they cite. COPE guidelines state that the "retracted status should be indicated as clearly as possible", but this is definitely not true for many retracted publications. Likewise, databases do not consequently link retracted articles with the notice of retraction. Furthermore, many papers are deposit in the "original", i.e. pre-retraction version on personal or institutional websites or online repositories. Similarly, printed "stock files" are obviously unaffected by a retraction. Clear identification of a retracted article using a watermark and in databases is a crucial step while incorporation of an electronic "retraction check" in reference management software and during the online submission is necessary to detect and avoid citing retracted literature. Solving this problem needs the close attention of everybody involved in the publishing process: authors, reviewers, and publishers.

Keywords Citations · COPE · Eric Poehlman · PubMed · Scott Reuben

Disclaimer The first author is not associated with any academic institute, blog or web-site. Screen-shots used in Fig. 1 used under the fair-use agreement.

✉ Jaime A. Teixeira da Silva
jaimetex@yahoo.com

✉ Helmar Bornemann-Cimenti
helmar.bornemann@medunigraz.at

[1] Miki-cho Post Office, Ikenobe 3011-2, P. O. Box 7, Miki-cho, Kagawa-ken 761-0799, Japan

[2] Department of Anaesthesiology and Intensive Care Medicine, Medical University of Graz, Auenbruggerplatz 29, 8036 Graz, Austria

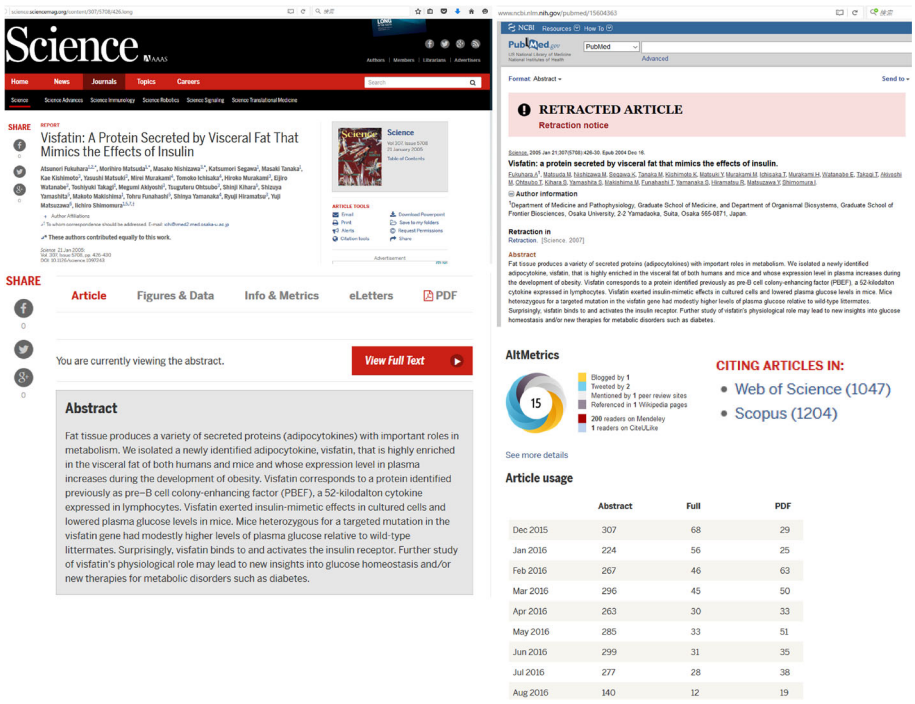

Fig. 1 The Science web-site (http://science.sciencemag.org/content/307/5708/426.long) of the Fukuhara et al. (2005) paper, which was retracted in 2007, does not show the retracted status of the paper, and this may mislead readers who rely exclusively on the title and abstract, into believing that the paper is fine, and can thus be cited. The PubMed page (http://www.ncbi.nlm.nih.gov/pubmed/15604363), in contrast, labels the paper as retracted. Data correct until August 15, 2016

The issues and case studies

In general, a retracted scientific paper should not be used, or cited. Whether retracted for misconduct, duplication or error, a paper that has been officially marked as retracted by the publisher has in essence been struck in full from the academic record. Despite this, some papers, particularly papers that have had a profound impact on research in the same filed, continue to be cited, even though they have been retracted. A citation represents the continued lifeline of a scientific paper, and in general shows that the work is being used, or appreciated, by those who cite it. When used correctly, a citation serves as a valuable tool to support a statement, method or hypothesis. Save for citation cartels and individuals who selfishly abuse citations in non-academic ways, a citation can serve as a crude barometer of a manuscript's popularity. Thus, a retracted scientific paper is a shock to a scientist and their research institute due to the financial and intellectual investment that goes into publishing a scientific paper. The often acclaimed self-correcting nature of science, in which scientists avoid citing work that has proven unreliability, thus phasing out such papers over time, might not be possible when retracted papers continue to be positively cited, i.e., in a praise-worthy manner. The topic of citing retracted papers is not new and has been debated for some time, for example, early concerns by Pfeifer and Snodgrass

(1990). Why then, if the issue has already been known for so long, and with such technological advances in information sharing on the internet, blogs, social media and data-bases, do retracted papers continue to be cited? This paper aims to examine some possible reasons why such papers continue to be cited, despite having been officially struck from the academic record.

Citing a retracted paper may pose some risks. For example, if methodology is flawed, then the results are most likely also suspect, or unreliable. Unfortunately, problems with one part of the data set casts doubt on the remainder of the data set, thus rendering the entire manuscript unfit for citation. In other cases, where conclusions drawn are incorrect, then there may be dangerous consequences, especially if the take-home message affects human health. In other cases, such as the duplication of a review, the effect may be close to innocuous. In all cases, however, a retraction casts a negative light and thus retracted papers should not be cited, unless to highlight an aspect about the retracted paper, such an analysis of its flaws. In the latter case, however, both the original paper and the retraction are cited, so it is ultimately not an edifying academic act.

In Budd et al. (1999) showed how 235 retracted papers had been cited 2034 times after retraction, and only 275 of those implicitly indicated that the articles had been retracted. The remainder positively cited retracted papers, even though the majority were clearly marked as "retracted" on MEDLINE. Neale et al. (2007) noted that 102 papers that had been corrected or retracted from 1991 to 2001 had, until May 2005, been cited 5393 times, with a median citation of $n = 27$ per retracted article. Couzin and Unger (2006), who identified several studies that were cited almost as much after retraction as they were before retraction, noted that once a paper is published in print, the printed version is almost never retracted. This is particularly true when copies are distributed in libraries worldwide, but a problem that is gradually becoming eliminated in newer digital-only online journals that do not make print copies available to their readership. It is also not uncommon for scientists to maintain their own personal paper repository or electronic reference manager in which PDF files of untainted (i.e., uncorrected) science remain intact and in their original published state. More out of habit rather than an issue of evasion, scientists will then cite a paper from their "stock" repository, possibly unaware that it was corrected, or retracted. Bornemann-Cimenti et al. (2016) found that almost half of Scott Reuben's retracted articles continued to be cited, even 5 years after retraction. Only a quarter of the citations clearly indicated that the work was retracted. Particularly problematic is the fact that this percentage has been decreasing over the years, to nearly 5%, 5 years after the article's retraction. Van der Vet and Nijveen (2016) conducted a case study which showed that propagation of retracted research is mainly due to direct citation. In their network analysis they found no case of indirect citation of a retracted paper.

Dealing with the problem and finding solutions

The number of insufficiently marked retractions on data-bases such as PubMed or on the online platforms of mainstream publishers, such as Elsevier's Scopus or sciencedirect.com, should theoretically be decreasing if one considers that information sharing, awareness and the speed of word of mouth have all increased in recent years with faster and more efficient internet tools, data-bases and social media. For example, Hilda Bastian indicated that all retractions listed on PubMed have recently been highlighted by a pink bar that is clearly visible to any reader (Oransky 2016). The Committee on Publication Ethics (COPE)

elaborated guidelines for journals editors on how to retract articles (Wager et al. 2009), including a recommendation to clearly identify the retracted status of an article to readers and electronic databases in all sources. Elia et al. (2014) demonstrated, however, that only 6% of retracted articles fulfilled all criteria adequately while only 83% had the full text marked as retracted. Wright and McDaid (2011) compared the results of three databases when searching for retracted articles. While MEDLINE indicated all papers as retracted, CENTRAL identified 80% correctly, while EMBASE displayed a notice of retraction in only 6% of cases.

The lag between the issuing of a retraction notice and its registration on data-bases such as PubPeer is decreasing, thus reducing the chance that the cause for citing such papers is a result of poor or lethargic dissemination of retraction-related information. Why then do some scientists continue to cite retracted papers? Why would a scientist cite a paper from PubMed, for example, that is clearly demarcated as retracted?

There are already robust policies in place by most mainstream publishers that deal with retractions, corrections, errata and expressions of concern, although several of these policies differ (Teixeira da Silva and Dobránszki 2016). Whose responsibility is it then to alert downstream readers after a retraction, correction, erratum or expression of concern? Who takes care of cleaning up the downstream literature that has cited retracted papers and relied upon potentially invalid findings to support their conclusions (Teixeira da Silva 2015a, b)? Even in the case of silent retractions, where the entire paper is struck or removed from the academic record (Teixeira da Silva 2016a), a risk of being cited continues to exist since authors may deposit copies on self-archives such as individual webpages or blogs, or on sites such as Mendeley, Academia.edu or ResearchGate. Currently, 79% of publishers formally allow authors some form of self-archiving (RoMEO 2016). Davis (2012) showed that 16.2% of retracted articles could be found on non-publisher websites and only 5% included retraction statements. Thus, other scientists who have perhaps done a Google Scholar search on a topic may be directed towards the intact copy of the original paper, rather than the original site, totally unaware that the original study was retracted. In some ways, the greater facilitation and expansion of information sharing, including by sites like SciHub that may harbor intact (i.e., in the unretracted state) copies of a manuscript, may be fueling an increase in the citation of retracted papers simply because the copies uploaded and retrieved by SciHub may be pre-retraction versions.

One sector of the community believes that the original authors whose paper has been retracted have the responsibility of contacting journals in which papers have cited a retracted or corrected paper. Such journals, together with corresponding authors, should then be responsible for correcting the literature, with appropriate notices of correction. Even when a notice is provided by non-authors of the retracted papers, such as regular readers or even anonymous commentators, journals have the responsibility of correcting the literature if they are aware that the citation list is erroneous. Should journal editors and publishers, who have been formally notified of such retractions and who are requested to correct the academic record (reference list in this case), but who actively ignore such notices, be slapped with a misconduct notice and somehow be penalized (e.g., suspension of impact factor) for actively ignoring their academic responsibilities? In such a case, the debate regarding the scope and fine-scale nature of what constitutes an "error" and whether it should be corrected, needs to be expanded (Teixeira da Silva 2016b). Understandably, in the case of deceased scientists (Teixeira da Silva and Dobránszki 2015), the responsibility falls on the shoulders of their research institutes and journals/publishers to correct the academic record. As Harry Klee, the then editor-in-chief of *The Plant Journal*,

correctly stated to Couzin and Unger (2006), "The journal is the primary point of enforcement against fraud. In the end, it's our process that got that work into publication."

To further reduce the risk of citing retracted papers, Sox and Rennie (2006) suggest that "journals should require authors to attest that they have checked their manuscript's reference list." However, it is likely that with an already over-burdened peer reviewer pool and non-remunerated editor system in place in most journals, adding additional requirements and checks-and-balances will burden the publishing system even further. As Sox and Rennie (2006) pointed out, with—at that time—3007 citations having been assigned to 186 Eric Poehlman papers, "who will take responsibility?" Ten years later, the summative citation of his papers exceeds 10,000. This may be the Achilles' heel in correcting the scientific literature that continues to cite retracted or corrected papers. This impasse was also pointed out by Wright and McDaid (2011).

How does one go about stopping scientists from citing retracted papers? One clue to the answer may lie in the Fukuhara et al. (2005) *Science* paper retraction, which was retracted in 2007, but which has gathered many-fold more citations from 2008 to 2016 than from 2005 to 2007, 1220 versus 27 citations, respectively in Elsevier's Scopus (or 1047 and 124, respectively in Thomson Reuters' Web of Science™), when the paper was academically valid. However, close examination of the *Science* web-site indicates that nowhere on the top page of the article does the word "retracted" appear, misleading readers into believing that the paper is still intact (i.e., not retracted), even if the PubMed page prominently labels the paper as retracted (Fig. 1). Article usage statistics indicate that dozens of PDF downloads continue to be made, even almost a decade after retraction, although it is unclear (at present) how many of those downloads eventually translate into citations. Eisenach (2009) suggested that new papers that challenge the premises made in retracted papers only be noted as footnotes rather than as citable references in the reference list.

Bornemann-Cimenti et al. (2016) suggested incorporating an electronic "retraction check" in reference management software. This could be based on services such as CrossMark, a database on retractions and updates of scientific articles. Unfortunately, such a software-plug in is not readily available. Furthermore, they propose an automated reference check during the review process. Van der Vet and Nijveen (2016) emphasized the service of Retraction Watch, a blog devoted to tracking and spreading information on retractions in the scientific literature. They also suggested that authors of previously published papers should be contacted in the case of a retraction of one of their citations, giving them the opportunity to revise their own articles. Thereby, the influence of retractions on the conclusions of subsequent publications can be minimized.

Acknowledgements The authors thank Judit Dobránszki (University of Debrecen, Hungary) for providing data from Elsevier's Scopus and Thomson Reuters' Web of Science™ pertaining to the Fukuhara et al. (2005) paper.

Compliance with ethical standards

Conflict of interest The authors declare no conflict of interest.

References

Bornemann-Cimenti, H., Szilagyi, I. S., & Sandner-Kiesling, A. (2016). Perpetuation of retracted publications using the example of the Scott S. Reuben case: Incidences, reasons and possible improvements. *Science and Engineering Ethics,*. doi:10.1007/s11948-015-9680-y.

Budd, J. M., Sievert, M. E., Schultz, T. R., & Scoville, C. (1999). Effects of article retraction on citation and practice in medicine. *Bulletin of the Medical Library Association, 87*(4), 437–443.

Couzin, J., & Unger, K. (2006). Scientific misconduct. Cleaning up the paper trail. *Science, 312*, 38–43.

Davis, P. M. (2012). The persistence of error: A study of retracted articles on the Internet and in personal libraries. *Journal of the Medical Library Association, 100*(3), 184–189. doi:10.3163/1536-5050.100.3.008.

Eisenach, J. C. (2009). Data fabrication and article retraction. *Anesthesiology, 110*, 955–956. doi:10.1097/ALN.0b013e3181a06bf9.

Elia, N., Wager, E., & Tramèr, M. R. (2014). Fate of articles that warranted retraction due to ethical concerns: A descriptive cross-sectional study. *PLoS ONE, 1*, e85846. doi:10.1371/journal.pone.0085846.

Fukuhara, A., Matsuda, M., Nishizawa, M., Segawa, K., Tanaka, M., Kishimoto, K., et al. (2005). Visfatin: A protein secreted by visceral fat that mimics the effects of insulin. *Science, 307*(5708), 426–430. doi:10.1126/science.1097243.

Neale, A. V., Northrup, J., Dailey, R., Marks, E., & Abrams, J. (2007). Correction and use of biomedical literature affected by scientific misconduct. *Science and Engineering Ethics, 13*, 5–24. doi:10.1007/s11948-006-0003-1.

Oransky, I. (2016). How to better flag retractions? Here's what PubMed is trying. http://retractionwatch.com/2016/07/20/how-to-better-flag-retractions-heres-what-pubmed-is-trying/. Last Accessed 18 Oct 2016.

Pfeifer, M. P., & Snodgrass, G. L. (1990). The continued use of retracted, invalid scientific literature. *Journal of the American Medical Association, 263*, 1420–1423. doi:10.1001/jama.1990.03440100140020.

RoMEO. (2016). Statistics for the 2264 publishers in the RoMEO database. http://www.sherpa.ac.uk/romeo/statistics.php. Last Accessed 18 Oct 2016.

Sox, H. C., & Rennie, D. (2006). Research misconduct, retraction, and cleansing the medical literature: Lessons from the Poehlman case. *Annals of Internal Medicine, 144*, 609–613.

Teixeira da Silva, J. A. (2015a). For whom the bell tolls: Downstream effects of retractions and the bump-on effects of post-publication peer review. *International Journal of Plant Biology & Research, 3*(4), 1050.

Teixeira da Silva, J. A. (2015b). The importance of retractions and the need to correct the downstream literature. *Journal of Scientific Exploration, 29*(2), 353–356.

Teixeira da Silva, J. A. (2016a). Silent or stealth retractions, the dangerous voices of the unknown, deleted literature. *Publishing Research Quarterly, 32*(1), 44–53. doi:10.1007/s12109-015-9439-y.

Teixeira da Silva, J. A. (2016b). An error is an error… is an erratum. The ethics of not correcting errors in the science literature. *Publishing Research Quarterly, 32*(3), 220–226. doi:10.1007/s12109-016-9469-0.

Teixeira da Silva, J. A., & Dobránszki, J. (2015). The authorship of deceased scientists and their posthumous responsibilities. *Science Editor (CSE), 38*(3/4), 98–100.

Teixeira da Silva, J. A., & Dobránszki, J. (2016). Notices and policies for retractions, expressions of concern, errata and corrigenda: Their importance, content, and context. *Science and Engineering Ethics,*. doi:10.1007/s11948-016-9769-y.

van der Vet, P. E., & Nijveen, H. (2016). Propagation of errors in citation networks: A study involving the entire citation network of a widely cited paper published in, and later retracted from, the journal Nature. *Research Integrity and Peer Review, 1*, 3. doi:10.1186/s41073-016-0008-5.

Wager, E., Barbour, V., Yentis, S., & Kleinert, S. (2009). Retractions: Guidance from the Committee on Publication Ethics (COPE). *Croatian Medical Journal, 50*(6), 532–535. doi:10.3325/cmj.2009.50.532.

Wright, K., & McDaid, C. (2011). Reporting of article retractions in bibliographic databases and online journals. *Journal of the Medical Library Association, 99*(2), 164–167. doi:10.3163/1536-5050.99.2.010.