23rd International Conference on Science and Technology Indicators
"Science, Technology and Innovation Indicators in Transition"

STI 2018 Conference Proceedings

Proceedings of the 23rd International Conference on Science and Technology Indicators

All papers published in this conference proceedings have been peer reviewed through a peer review process administered by the proceedings Editors. Reviews were conducted by expert referees to the professional and scientific standards expected of a conference proceedings.

Chair of the Conference

Paul Wouters

Scientific Editors

Rodrigo Costas
Thomas Franssen
Alfredo Yegros-Yegros

Layout

Andrea Reyes Elizondo
Suze van der Luijt-Jansen

The Schön case: Analyzing in-text citations to papers before and after retraction

Marc Luwel, Nees Jan van Eck and Thed N. van Leeuwen

luwel@cwts.leidenuniv.nl; ecknjpvan@cwts.leidenuniv.nl; leeuwen@cwts.leidenuniv.nl
Centre for Science and Technology Studies, Leiden University, PO Box 905, 2300 AX Leiden (The Netherlands)

Introduction

Although only a small fraction of all scientific publications is retracted for misconduct, it has a profound impact on the research community, policy makers and the public at large. Indeed, over the last decades scientific integrity became a hot issue in science policy (Tchao, 2014). Moreover, papers retracted for fraud or other reasons should not be cited or used anymore. Although in most cases they are earmarked as retracted in bibliographic databases and in the electronic version of journals, they often remain cited a long time after the retraction notice is published. The misconduct case of Jan Hendrik Schön is a well know example. It attracted a lot of attention due to the renown of the researcher's institute, the claims and the sheer number of publications involved. This researcher co-authored more than 100 papers and rose through prominence at the turn of the century with a number of apparent ground breaking discoveries in materials science (Reich, 2009).

In September 2002, an investigation commissioned by Bell Labs (Beasley et al., 2002), his employer, concluded that 17 papers contained manipulation and misrepresentation of data. They were retracted along with an additional 14 papers based on them (Table 1). For all these publications the retraction notices were collected. The retracted work was published in top journals such as *Science* (8 papers), *Nature* (7 papers), *Physical Review B* (5 papers) and *Applied Physics Letters* (4).

Table 1. Number of retracted and non-retracted articles and letters co-authored by Schön per publication year in the WoS.

Year	No. of publications		
	Retracted	**Non-retracted**	**All**
1994	0	1	1
1996	0	2	2
1997	0	7	7
1998	1	8	9
1999	1	9	10
2000	11	16	27
2001	16	26	42
2002	1	2	3
All	30	71	101

Other papers raised suspicion among the co-authors and the scientific community. Indeed, the former expressed concerns about the data in several non-retracted papers (Bao, 2003).

Subsequently, after a long legal battle, Schön's alma mater the University of Koblenz revoked his PhD degree due to "dishonorable conduct".

Notwithstanding all the rumors provoked by this scandal in the physics community, the retracted papers remained cited even several years after they were removed from the body of literature. As shown in Figure 1, within the journals included in the Web of Science (WoS), an international bibliographic database owned by Clarivate Analytics, between 2004 and 2016 the retracted papers received around 50 citations annually. This is roughly the same number of citations that is received annually by the larger set of non-retracted papers. As in the rest of this study, the Schön-Schön self-citations (Schön citing work he co-authored) are not counted.

Figure 1. Number of citations per year received in the WoS by retracted and non-retracted publications co-authored by Schön.
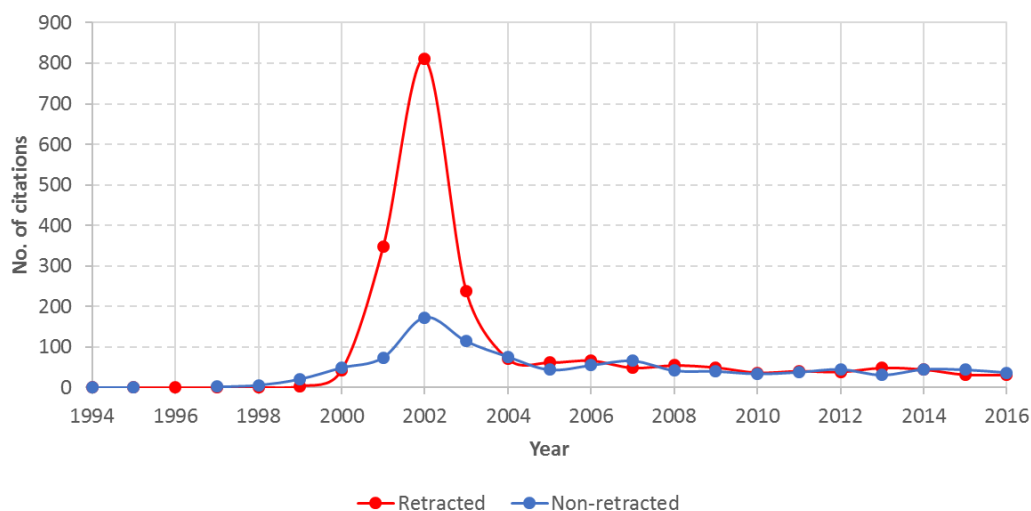
To get insight in the role of citations to Schön's oeuvre in citing publications, natural language processing techniques on publications' full text are used. The main questions are: to what extent, in what sections in the citing publications and in what context are papers co-authored by Schön cited. We make a distinction between retracted and non-retracted papers and between the citations that papers receive before and after the results of the fraud investigation became publicly available.

Methodology and data

Our analysis is based on a database that we constructed of the full text of publications citing Schön's oeuvre and that are published in Elsevier journals. The data collection is limited to publications in Elsevier journals because for the full text analysis a machine readable version of the publications is required. Elsevier provided CWTS access to the journal papers in such a format. To collect the full text data, Elsevier's Science Direct Article Retrieval API is used. Publications that cite at least one of the papers co-authored by Schön and that are published between 1998 and 2016 are taken into account. The full text of these publications is downloaded in the XML format. Special-purpose software was developed to identify within the full text of each citing publication the exact position of the citations to Schön's work, further called in-text citations. In addition, the software automatically extracts:
- The title of the section in which each in-text citation is mentioned; and
- The sentence containing the in-text citation as well as the sentence before and after it.

A large majority of the publications in Elsevier journals are not structured along the IMRaD model (Introduction, Methods, Results, and Discussion). The titles of the sections containing the in-text citations were manually standardized in four categories: 'Introduction', 'Body of the text', 'Conclusions', and 'Null', where the latter applies to letters not using section titles.

The impact of the misconduct review report (Beasley et al., 2002) and the subsequent retraction notices published in 2002 and early 2003 is studied by comparing the distribution over the standardized sections of in-text citations referring to retracted papers and in-text citations referring to non-retracted papers. A distinction is made between two periods:
- The pre-retraction period: period from 1998 (first publication year for which we have full text publications data available) until and including 2004; and
- The post-retraction period: period from 2005 until and including 2016.

To delimit the pre- and post-retraction period, 2004 was chosen to take the publication delay of the citing publications into account. As the last retraction notices were published in 2003, 2004 is chosen as the year to delimit the two periods to take the publication delay into account. The sensitivity of the results on this choice was tested. These distributions are also compared with the distribution of in-text citations in a set of almost 5 million publications (Boyack et al., 2018). To correct for possible changes over time, we included only in-text citations that refer to publications published between 1994 and 2002, the period Schön's papers were published. To count the in-text citations, 2004 again is used as a citing publication year to delimit the pre- and post-retraction period.

Finally, the precise mention of each in-text citation was manually identified and classified by analyzing the sentences around them and where necessary the full section or paper. For the classification of the mentions, the category 'Retracted/fraud' was added to the scheme proposed by Bar-Ilan and Halevi (2017):
- *Neutral*: the retracted publication is mentioned as a publication that appears in the literature without a judgment on its validity;
- *Positive*: the retracted publication is cited as legitimate work and used to corroborate the authors' current research;
- *Negative*: the retracted publication is mentioned as such and its findings as inappropriate and/or questionable; and
- *Retracted/fraud*: the word 'retraction' or 'fraud' is mentioned in the citing publication.

To make the classification as robust as possible, the four-eyes principle was used for assigning the in-text citations. As for the in-text citations, this classification is made for the different sections and for the two periods. As the number of in-text citations in the concluding section is rather small, they are amalgamated with those in the body of the text under the heading 'Non-introduction' section.

Results

In the WoS, 30 articles co-authored by Schön can be found that are labelled as retracted. These articles are published between 1998 and 2002 (Table 1, see above). It turns out that one article is actually a correction and that it is not retracted. For the consistency of the analysis, the WoS labelling is maintained. Of these 30 articles, 21 are cited in Elsevier journals. Between 1994 and 2002, Schön co-authored also 70 articles and 1 letter indexed in the WoS that are not retracted. 49 of these papers are cited in Elsevier journals. As the publications in Elsevier journals represent only about 22% of the total number of publications included in the WoS (Boyack, et al., 2018), the number of citations Schön's work receives in Elsevier journals is considerable lower than in the WoS. For example, for the retracted articles

published in 2000 and 2001 the number of WoS citations till 2016 is 1876 with 345 (18.4%) in Elsevier journals. For the non-retracted papers published in these two years the numbers are 569 and 84 (14.7%). Each cited paper can be mentioned more than once in the full text: retracted articles published in 2000 and 2001 are mentioned 490 times in Elsevier journals; for the not retracted paper this number is 125.

As the retracted papers co-authored by Schön are cited long after their retraction and even today, the full text analysis makes it possible to detect changes over time in the distribution of in-text citations. Given the relative low number of (in-text) citations it was opted to use the above mentioned two periods.

Table 2. Number of in-text citations in the Elsevier dataset to retracted and non-retracted publications co-authored by Schön. A breakdown is provided for in-text citations that are given in publications published in the pre-retraction period (1998-2004) and post-retraction period (2005-2016) and for in-text citations that are mentioned in the introduction section and other sections.

Section	No. of in-text citations			
	Retracted		Non-retracted	
	1998-2004	2005-2016	1998-2004	2005-2016
Introduction	208 (61%)	106 (74%)	84 (50%)	70 (46%)
Non-Introduction	134 (39%)	37 (26%)	85 (50%)	81 (54%)
All	342 (100%)	143 (100%)	169 (100%)	151 (100%)

As shown in Table 2, for the two periods the distribution of the in-text citations referring to the non-retracted papers over the introduction and the other sections combined is very similar. In the pre-retraction period, the number of in-text citations to the retracted papers in the introduction section is about 10% higher compared to the non-retracted papers. But in the post-retraction period, in-text citations to the retracted papers are more strongly concentrated in the introductory section: 74% versus 46% for the non-retracted papers.

Table 3. A breakdown of the number of in-text citations in a set of almost 5 million Elsevier publications in the pre-retraction period (1998-2004) and post-retraction period (2005-2016) and in the Introduction and Non-Introduction section. The in-text citations refer to papers published between 1994 and 2002.

Section	No. of in-text citations	
	1998-2004	2005-2016
Introduction	8,046,152 (36%)	19,455,498 (38%)
Non-Introduction	14,322,319 (64%)	31,089,735 (62%)
All	22,368,471 (100%)	44,542,724 (100%)

Combining the results of Tables 2 and 3, the most striking observation is the strong concentration of the in-text citations in the introduction section referring to retracted papers compared to the full Elsevier dataset, especially in the post-retraction period.

To check the influence on the results of the choice of 2004 as reference year to delimit the pre- and post-retraction period, a sensitivity test was done: using 2003 and 2005 as reference year resulted in marginal changes of around 1%.

Table 4. Number of in-text citations in the Elsevier dataset to retracted and non-retracted publications co-authored by Schön. For the different sections the precise mention of each in-text citation was classified in one of four categories (neutral, positive, negative, or retraction). A further breakdown is provided for in-text citations that are given in publications published in the pre-retraction period (1998-2004) and post-retraction period (2005-2016).

| Section | Mention | No. of in-text citations | | | |
| | | Retracted | | Non-retracted | |
		1998-2004	2005-2016	1998-2004	2005-2016
Introduction	Neutral	206 (99.0%)	105 (99.1%)	84 (98.8%)	67 (95.7%)
	Positive	1 (0.5%)	0 (0%)	1 (1.2%)	3 (4.3%)
	Negative	0 (0%)	0 (0%)	0 (0.0%)	0 (0.0%)
	Retraction	1 (0.5%)	1 (0.9%)	0 (0.0%)	0 (0.0%)
Non-Introduction	Neutral	83 (61.9%)	25 (67.6%)	14 (16.7%)	22 (27,2%)
	Positive	26 (19.4%)	7 (18.9%)	68 (81.0%)	58 (71.6%)
	Negative	6 (4.5%)	0 (0%)	2 (2.3%)	1 (1.2%)
	Retraction	19 (14.2%)	5 (13.5%)	0 (0.0%)	0 (0.0%)
All	Neutral	289 (84.5%)	130 (90.9%)	98 (58.0%)	89 (58.9%)
	Positive	27 (7.9%)	7 (4.9%)	69 (40.8%)	61 (40.4%)
	Negative	6 (1.8%)	0 (0.0%)	2 (1.2%)	1 (0.7%)
	Retraction	20 (5.8%)	6 (4.2%)	0 (0.0%)	0 (0.0%)

Table 4 shows that the in-text citations to retracted and non-retracted papers are not only primarily concentrated in the introduction, but that their precise mention is mostly neutral. This could be expected, as in this section authors generally present an overview of earlier work used in their own research. In the other sections combined, 68% of the mentions of the retracted work is also neutral and 14% indicates that the article is retracted and/or earmarked as fraudulent. This pattern is in sharp contrast with the in-text citations to non-retracted work where even in the pre-retraction period a large majority of the mentions is positive and remains positive in the post-retraction period.

Discussion

The fact that fraudulent results remain cited after the publications have been retracted, is a reason of concern not only for the research community, but also for the public at large. Confidence in the creation of scientific knowledge could be eroded and in the case of medical research, public health in general could be at stake (e.g., consider the case of Dutch cardiovascular researcher and medical doctor Poldermans; Erasmus MC Follow-up Investigation Committee, 2012).

To understand the use of retracted work in later research one has to go beyond citation counting and analyze the distribution of the in-text citations and their precise mention. The increased availability of the full text of publications in machine readable format and the progress in natural language processing makes it possible to largely automate the underlying processes.

This approach is illustrated by the study of Schön's oeuvre. The retracted papers co-authored by Schön remain cited even 15 years after the retraction notices were published at the same annual rate as the much larger set of his non-retracted papers. In contrast to the neutral mentions of both the retracted and the non-retracted work in the introduction of publications in Elsevier journals, there is a large difference in the precise way in which Schön's work is cited in the other sections of scientific publications. In the case of the retracted publications, in more than 60% of the cases the citing authors just indicate that the papers co-authored by

Schön exist. In the case of the non-retracted publications, however, a large majority of the citing authors is confirming the results presented in the papers co-authored by Schön. This could indicate that even before retraction researchers could not replicate the results or were somewhat suspicious about the reported outcomes without openly stating this. However, before and after retraction in the non-introductory sections, around 20% of the mentions to retracted work are positive. A majority are in publications not reporting on experimental work but on theoretical models using the parameters Schön claimed to have measured, such as the temperature at which superconductivity was observed. In fabricating his data, Schön cleverly made estimated guesses about phenomena outside the boundaries of previous work that at first view were spectacular but not that unrealistic at all. In some cases, his 'experimental data' fitted reasonably well with existing theoretical models. Other out of the blue claims could later be corroborated by genuine experimental work.

Our next efforts will be focused on getting a better insight in various other aspects of this case. We are working on partitioning the in-text citations in classes: a retracted Schön paper cited individually or in a cluster together with two or more papers from other authors. Another venue for further work is enriching the full text dataset with journals from other publishers. Work has already been done on the much smaller PLOS dataset that is publicly available (Bertin et al., 2016).

References

Bar-Ilan, J. & Halevi, G. (2017). Post retraction citations in context: a case study. *Scientometrics*, 113(1), 547-565.

Bertin, M., Atannassova, I., Gingras Y., & Lariviere, V. (2016). The invariant distribution of references in scientific articles. *Journal of the Association for Information Science and Technology*, 67(1), 164-177.

Beasley, M.R., Datta, S., Kogelnik, H., Kroemer, H., Monroe, D. (September 2002). *"Report of the Investigation Committee on the possibility of Scientific Misconduct in the work of Hendrik Schön and Coauthors".* Bell Labs.

Bao, Z., Batlogg, B., Hadziioannou, G., Kloc, C., Meng, H., & Wildeman, J. (2003). Retraction. *Advanced Materials*, 15(6), 478.

Boyack, K.W., Van Eck, N.J., Colavizza, G., & Waltman, L. (2018). Characterizing in-text citations in scientific articles: A large-scale analysis. *Journal of Informetrics*, 12(1), 59-73

Erasmus MC Follow-up Investigation Committee (2012). Report on the 2012 follow-up investigation of possible breaches of academic integrity. Retrieved from https://cardiobrief.files.wordpress.com/2012/10/integrity-report-2012-10-english-translation.pdf

Tschao, R. (2014). Disciplinary impact and new model scholarship. A need to archive correct biomedical scientific data and to prevent continued citation of retracted scientific publications. *International Journal of Humanities and Arts Computing*, 8 (Supplement), 29-37