Correcting the Scientific Record: Retraction Practices in Chemistry and Materials Science

Peer-reviewed articles, published by scholarly journals, currently form the cornerstone of the modern scholarly publication system and guarantee the dissemination of research findings through the worldwide, ever-increasing community of researchers. Collectively these published works, stamped with the seal of approval of a review by the authors’ peers, form the scientific record—the record of knowledge accumulated by mankind. It is the duty of every scholar to add knowledge to this record by publishing but also to ensure the integrity of the existing works by critically assessing them: before publication, acting as a reviewer or editor, and post-publication, by building upon existing works, improving them, and checking their reproducibility.

The means of post-publication peer review of articles, which was once limited to formally published comments (“Comment on...”), journal clubs and conference coffee breaks, are rapidly expanding through the use of Internet and social media. Discussion of published papers regularly takes place on Twitter and through blog posts and preprints, as well as in structured discussions: comments on the webpage on published papers (e.g., PLOS One and Frontiers journals), indexing servers (PubMed Commons, now closed1), or dedicated websites (such as PubPeer2). Critique of published articles is a necessary and healthy part of the advancement of science. Sometimes, it can lead to the identification of serious flaws in the data or authors’ analysis, so that the findings or the conclusions published cannot be trusted anymore. In such cases, the paper may be corrected or retracted, i.e., expunged from the scientific record.

COPE, the Committee on Publication Ethics, publishes a series of guidelines (policies and practices) that are considered the industry standard in publishing ethics. The areas covered include the handling of allegations of misconduct, complaints and appeals, data issues and reproducibility, and standards of authorship, as well as post-publication corrections and the retraction of papers. COPE guidelines give clear insights into the difference in nature between corrections and retractions.3 Articles should be corrected if “a small portion of an otherwise reliable publication proves to be misleading (especially because of honest error)”. On the other hand, “journal editors should consider retracting a publication if:

- they have clear evidence that the findings are unreliable, either as a result of misconduct (e.g., data fabrication) or honest error (e.g. miscalculation or experimental error),
- the findings have previously been published elsewhere without proper crossreferencing, permission or justification (i.e. cases of redundant publication),
- it constitutes plagiarism,
- it reports unethical research.”

Retractions thus ensure that the literature is corrected, alerting readers to the fact that a publication contains erroneous or unreliable data, and give clear insight into the nature of the issues.

Despite the healthy role of retractions in preserving the scientific record, and while erroneous data can be the result of a good faith mistake, there is definitely a stigma associated with the retraction of a paper. COPE guidelines state that “The main purpose of retractions is to correct the literature and ensure its integrity rather than to punish authors who misbehave,”3 but previous work has shown a notable resistance to admitting error in scientific papers.4 The term retraction is too often associated with research misconduct, giving it a negative connotation for authors.5 This is particularly true in a highly competitive environment, where academics are driven to publish often and produce high-impact papers: Jin et al. showed that retractions have a negative effect on citations for early career researchers.6 The same argument can also be made for the publishers, who may fear a dent in the reputation made for the publishers, who may fear a dent in the reputation.

In this context, and while the number of retractions is rising,7,8 there is relatively little information available about retractions and retracted papers, beyond the retraction notices infrequently published by journals. There is no central repository or authoritative database that can be easily queried—although the Retraction Watch website, which covers the topic of retractions and publication ethics in general, has been collating such a database.9 Previous systematic studies have focused on retractions in specific fields, and in particular in medicine10—13—with the notable exception of a study by Grieneisen et al. that spanned several fields of research.14 In order to better understand the existing practices for article retraction in the chemical sciences, I have performed a systematic survey of 331 papers retracted in 2017 and 2018 and their retraction notices, publicly available on the journals’ websites. This article looks at the statistics of retractions, their distribution per country, and the occurrence of multiple retractions. I also provide a classification of the reasons behind the retractions and the distribution of their occurrence.

How Many Papers Are Retracted?

A search of Elsevier’s Scopus database,15 for retraction notices, in the fields of chemistry, materials science, and chemical engineering, resulted in a list of 331 papers that were retracted during the years 2017 and 2018 (see the Methods section for details). This is a very small number compared to a total of 1,114,476 papers published in those fields in the same two years, with on average 3 papers retracted for 10,000 papers published.

The geographic distribution of the retracted papers is given in Figure 1. The distribution generally reflects the number of papers published by each country. We see that 31.4% of retracted papers involve a Chinese affiliation, which is in line with the fact that 30.9% of papers published in 2017–2018
have a Chinese co-author. The same is true of the United States, with 14.2% of retractions (15.1% of published papers), South Korea (4.5% vs 4.0%), and other countries lower in the list—although with smaller absolute numbers, the statistics become noisier. The ratio of retraction in China is of particular interest here, given the country’s cash-per-publication reward policy has been suspected to drive researchers toward, at best, the production of more papers of lower scientific significance, or at worst a negative impact on academic ethics. The data show that this does not, for the field and period studied, lead to an increase in the ratio of retraction. There are, however, countries whose “share” of retractions exceed their share of published papers. Indian affiliations account for 17.2% of retractions in our database, compared to 7.3% of published papers worldwide—a difference of a factor of 2.4. But the country featuring the largest rate of retraction is Iran, which is involved in 11.2% of retractions compared to 2.7% of published papers. The ratio of retracted papers from Iran is 4.1 times larger than from the rest of the world. I note that while a small number of authors contribute heavily to those numbers, they do not—in themselves—appear to account entirely for the higher rate. Analysis over a larger time period would be necessary to confirm these trends.

How long after their original publication are papers retracted? The distribution of retraction time, presented in Figure 2, is well-spread. 10% of papers are retracted within 3 months of the original publication. In several cases the issues were discovered even before papers were “formally” published as part of a journal issue, while the accepted versions of the manuscripts were posted online. On the other hand, many papers take longer to retract: the median time to retraction in this data set is 24 months. But the distribution has a long tail; the tardiest retraction observed is for a paper published in June 2000 and retracted 18 years later, in February 2018 (DOI: 10.1007/s12540-018-0048-0).

Looking at the distribution in citations of the retracted papers (Figure 3), we see also a wide variety. This is linked, to a large extent, with the large distribution of time to retraction. Since we look at papers retracted in 2017 and 2018, the vast majority of citations reported were made before the retraction—it would be interesting, in future work, to look at older retractions to see if those papers are still cited by the community. The median number of citations is quite low, 4, while the average is 10.7. The most cited paper in the data set was cited 121 times in the course of 13 years (DOI: 10.1016/j.biortech.2017.08.105) and was retracted as a duplicate of another publication by the same author (which itself has gathered 130 citations).

How common are multiple retractions?

The past few years have seen the revelation of a certain number of cases of high-profile scientific misconduct, where the discovery of unethical practices by an author or research group leads to retraction of several papers. From the 2017–2018 retractions analyzed here, we clearly identify some cases of large-scale retractions, all of which have been reported in specialized media. The largest “clusters” of retractions in that two-year period include 18 retractions for R. Madhuri and P.
K. Sharma (Indian Institute of Technology, Indian School of Mines in Dhanbad),18 10 for G. Tang and X. Ye (Tsinghua University in Beijing),19 9 for A. S. Elahi and co-workers (Islamic Azad University in Tehran),20 7 for S. Thakur and R. B. Tokas (Bhabha Atomic Research Centre in Mumbai),21 and 7 also for S.-G. Shin (Kangwon National University in Chuncheon).22 These retractions were, in every case, linked to ethical shortcomings.

These cases of large-scale retractions are well publicized, and they account for a nonnegligible fraction of the retracted papers. We count 88 retracted papers involved in multiple retractions (3 or more papers retracted with at least one common co-author), which makes for 27% of the retractions. On the other hand, out of 1093 authors of retracted papers in 2017 and 2018, 92% (1003 authors) are involved in a single retracted paper, and only 31 authors (3%) are involved in the 27% of multiple retractions.

WHY ARE PAPERS RETRACTED?

To see why papers are being retracted, I have classified the reasons being listed in each retraction notice (see Methods for details) into four main categories. The results are shown in Figure 4 and Table 1, and the data is available as Supporting Information—including, for each paper, the full retraction notice and my classification into categories. The most frequent reason behind paper retractions is plagiarism, involved in 42% of cases. It can take the form of duplicate publication (15% of cases; the direct reuse of an entire article, sometimes after translation), self-plagiarism (45%; the reuse of significant parts of text without proper citation), or plagiarism of previously published works by other authors (40%). We note that although nowadays publishers can now check for plagiarism with automated software tools upon submission of manuscripts,23 that was not always the case, and some papers retracted in 2017 and 2018 were published before this software was available. We cannot determine, from the published retraction notices, what is the main motivation of authors in plagiarizing, but previous studies have pointed to possible contributing factors: lack of academic infrastructure and mentoring24 and incentives based on the number of publications, whether in the form of cash bonuses17 or career advancement.25

The second largest root cause of retraction is the presence of ethical issues with the data reported, with 90 occurrences (27%). The terminology used by retraction notices—which can be written either by the publisher, the editor, or the authors—to discuss this data can be very varied. Some notices are very specific, providing a clear characterization of the issues: “research misconduct”, “data fabrication”, “data falsification”, “inappropriately and extensively digitally manipulated”, etc. Other notices are much less clear-cut, mentioning unspecific “problems”, “significant problems”, “extraordinary similarities”, “incorrect data”, “doubts cast on the data”, etc.

Table 1. Breakdown of the Reasons Listed for Retraction, Classified from the Text of the Retraction Notices

<table>
<thead>
<tr>
<th>Category</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plagiarism</td>
<td>139</td>
</tr>
<tr>
<td>including</td>
<td></td>
</tr>
<tr>
<td>Duplicate publication</td>
<td>21</td>
</tr>
<tr>
<td>Self-plagiarism</td>
<td>62</td>
</tr>
<tr>
<td>Plagiarism of others</td>
<td>56</td>
</tr>
<tr>
<td>Data issues</td>
<td>90</td>
</tr>
<tr>
<td>including</td>
<td></td>
</tr>
<tr>
<td>“Problematic” data</td>
<td>28</td>
</tr>
<tr>
<td>Falsified data</td>
<td>54</td>
</tr>
<tr>
<td>Ownership or disclosure issues</td>
<td>8</td>
</tr>
<tr>
<td>Authorship issues</td>
<td>62</td>
</tr>
<tr>
<td>including</td>
<td></td>
</tr>
<tr>
<td>Missing authorization to publish</td>
<td>30</td>
</tr>
<tr>
<td>Missing author</td>
<td>13</td>
</tr>
<tr>
<td>Unclear authorship</td>
<td>12</td>
</tr>
<tr>
<td>Affiliation error</td>
<td>7</td>
</tr>
<tr>
<td>Honest errors</td>
<td>54</td>
</tr>
<tr>
<td>including</td>
<td></td>
</tr>
<tr>
<td>Error on the theory or data analysis</td>
<td>21</td>
</tr>
<tr>
<td>Experimental error</td>
<td>17</td>
</tr>
<tr>
<td>Work not reproducible</td>
<td>16</td>
</tr>
<tr>
<td>Other factors mentioned in retraction notices</td>
<td></td>
</tr>
<tr>
<td>Publisher error</td>
<td>5</td>
</tr>
<tr>
<td>Copyright issue</td>
<td>3</td>
</tr>
<tr>
<td>Abuse during review</td>
<td>24</td>
</tr>
<tr>
<td>Unable to contact one or more co-authors</td>
<td>23</td>
</tr>
<tr>
<td>Institutional inquiry</td>
<td>19</td>
</tr>
<tr>
<td>Discussion of individual responsibility of authors</td>
<td>43</td>
</tr>
</tbody>
</table>

Figure 4. Breakdown of the reasons listed for retraction, classified from the text of the retraction notices. The sum of the four main categories exceeds the total number of papers, because retractions can list several contributing factors in the publisher’s decision.
I have therefore broken down the “data issues” category into three different subclasses. In a majority of 60% of the cases, the data is falsified, i.e., the unethical handling of data is clearly presented. This includes cases of fabrication, invalid duplication, digital editing, etc. and accounts overall for 16% of retracted papers. In a smaller number of cases (31% of data issues, 8% of rejections), the issues with the data are not clearly explained, and the unethical nature of the publication is merely hinted at through the language of the notice. Although this is not the majority of cases, it still appears quite high—such cases should be totally avoided, to provide maximum information to the scientific community and to obey COPE guidelines:3 “retraction notices should always mention the reason(s) for retraction to distinguish honest error from misconduct”. Finally, there were a few (8) cases where the cause of retraction is due to lack of data ownership, lack of authorization to publish the data, or disclosure of confidential or proprietary information.

The third cause of paper retraction is related to authorship issues: this is mentioned as a factor in 19% of retraction notices. In most cases, the issue is a lack of authorization to publish from one of the authors. In most cases, it is a co-author who has not been informed of the submission and publication of the paper—and quite often a senior co-author, supervisor or advisor, according to the retraction notices surveyed. In a few cases, this can mean the inclusion on the author list of a co-author who did not participate in the work at all, possibly adding a well-known name to the author list to improve a manuscript’s chances of success. Authorship disputes leading to retraction can also take the form of a missing co-author (13 cases), i.e. someone whose contribution to the work met the criteria for authorship, yet was not added to the paper’s author list—or cases where the contribution of different individuals is unclear or disputed (12 cases). In particular, there are signs in some retraction notices of authorship being sold, after review, a known practice in some countries:26 “the paper was submitted to the journal by the first author, who was not able to provide a reasonable explanation for the addition of the other seven authors to the revised version of the paper”. Finally, some papers were retracted due to the presence of an incorrect affiliation (7 cases), although this is often a contributing factor but not the sole reason for the retraction.27

It may be surprising to note that the least-common category of retractions is that of papers that were identified, after publication, to contain “honest errors”, i.e., errors that are not attributed to misconduct. This represents 16% of retractions in this corpus, and they are almost evenly split among (i) theoretical errors, in mathematical derivations or during data analysis; (ii) experimental errors, including experiments performed in invalid conditions or apparatus used in an unreliable manner; and (iii) results that could not be reproduced, by the authors or other researchers, in later works. It is interesting to note that these retractions tend to be published at the request of the authors and in many cases feature long retraction notices that fully detail the nature of the errors made (see, e.g., DOI: 10.1021/acs.macromol.7b00492 and DOI: 10.1038/nchem.2885) and the investigation. In some cases, these investigations into the reproducibility of the work have even been published, separately of the retraction, as a scientific article in their own right (see DOI: 10.1103/PhysRevB.98.079902 and ref 28). This really represents a “positive” type of retraction, which is to be encouraged, for it benefits greatly the community.

OTHER FEATURES OF RETRACTION NOTICES

In addition to these four main categories of retraction reasons, we note here a few other less-frequently mentioned retraction reasons. The first is the occurrence of a technical error at the publisher (5 cases, 1.5%), for example, where a paper was accepted without review by mistake or was published twice due to an error in the production chain. The second is the retraction of papers whose peer review was insincere: I list this separately here, because while it affects 24 papers (7.2%), it is actually not common and happened at a single publisher (Elsevier). An inquiry revealed that these papers had been accepted after a biased or fake peer review: authors had submitted potential reviewers’ contact details (including email), yet those email accounts were not associated with the listed individuals and instead manipulated (either by the authors or third parties).29 Most publishers have now put in place a verification mechanism, to ensure the authenticity of reviews.

In addition to the reasons stated as causes for retracting papers, the notices can feature additional information about the process of the retraction. For example, some notices state who requested the retraction (author or authors, editor, publisher)—but since not all notices do, it was impossible to gather good statistics on this. I have also noticed that many notices (23 out of 331) mention the publisher being unable to contact some or all of the authors. This occurs in particular in cases where the retraction is due to ethical misconduct, probably because authors either try to ignore the issue, hoping it will go away, or may have judged that silence is the best defense.

Two other significant features found in some (but not all) retraction notices are the discussion of the individual responsibility of the authors and the mention of possible links to institutional investigations, either pending or completed. Inquiries by academic institutions, usually the authors’ principal affiliation, are mentioned in 19 notices (5.7%). It is explicitly mentioned, in several cases, that the article is retracted following an institutional inquiry into research misconduct. The reverse is quite rare: it does not appear that journals or publishers routinely inform the authors’ academic institutions of the retraction proceedings.

Finally, 13% of the retraction notices studied discuss explicitly the individual responsibility of the authors of a retracted paper. This is rarely a systematic discussion of all co-authors, but more often done to exonerate one specific author (“X has stated that she was unaware of the submission”) or to assign the blame to a single author (“sole responsibility lies with the corresponding author”; “the corresponding author takes full responsibility and apologizes for the academic misconduct”). In other cases, such statements are used to indicate which authors endorse the retraction or not (“X and Y oppose the retraction”).

This analysis of the retraction factors is in generally in line with the findings reported by Grieneisen et al. across several fields14 (i.e., on all journals indexed by Web of Science).

LEVEL OF DETAIL IN RETRACTION NOTICES

The content of retraction notices can vary widely—in both size and information provided. We found at least one journal (Journal of the American Oli Chemists’ Society) that appears not to publish freely the text of retraction notices (DOI: 10.1080/10837450.2018.1451433; DOI: 10.1007/s11746-011-1789-9),
a practice contrary to the COPE Guidelines. Other times, retraction notices contain no actual information as to the motivating factors. As an example, one retraction states:

“We, the Editor and Publisher of Mineral Processing and Extractive Metallurgy Review are removing the following article: [article title]” (DOI: 10.1080/10337450.2018.1451433)

Another manages to give slightly more information while keeping readers in the dark:

“This article has been retracted at the request of the authors because of a business decision by their employer.” (DOI: 10.1016/j.jcrgyro.2018.09.045).

A likely contributing factor for such short notices is the pressure on journals and their editors to avoid using clear or harsh retraction language, for fear of being proven wrong in a further inquiry—or of fear of legal retribution, as some authors have been known to contest retractions in law courts.

At the other end of the scale, some notices are extremely detailed and provide a detailed analysis of the reasons behind the retraction. Many retraction notices for plagiarism provide a clear picture of what material was copied from the authors, and what were the original sources. Unsurprisingly, the notices of papers retracted for good-faith mistakes are often the most detailed, providing good insight to readers on the problems with the published work. For example, Lu et al. (DOI: 10.1021/acs.macromol.7b00492) retracted four papers due to incorrectly performed molecular simulations and gave a detailed account of the shortcomings of their simulation procedure in a full-page retraction notice. Another interesting inquiry, related to a retracted paper, is the work of Ryan et al., who repeated the chemical reactions of a retracted organic chemistry paper (DOI: 10.1021/acs.orglett.7b02161) and provided a full analysis of the shortcomings in the analysis of the original paper.

Some retraction notices even acknowledge whistleblowers (when they did not wish to remain anonymous, we suppose!) “for the considerable effort put into collecting the evidence and reporting this case of multiple publication”, which I see as an encouraging note.

■ CONCLUSIONS

The aim of this study was to provide an overview of the retraction practices in chemistry and materials sciences. By analyzing retraction notices for a two-year period (2017–2018), I provide quantitative data on a phenomenon that is relatively common for the field as a whole—although quite extraordinary at the individual level. The retraction across all countries is relatively homogeneous, with two countries having significantly higher retraction rates: India and Iran. The causes of retraction are mainly related to misconduct, with plagiarism accounting for 42% of retractions, and data issues for 27%. Authorship issues are also common, and in particular papers being submitted without the knowledge and authorization of all authors, often by younger co-authors: this probably indicates that better education or training is needed in the workings of the academic publishing system and its rules.

While this analysis only concerns one academic field, nothing in the trends observed appears to be particularly tied to the specificities of chemistry or materials sciences. Its findings are thus probably generalizable to neighboring fields in experimental sciences, including physics and biology—although this cannot be directly confirmed, as I could not identify any systematic study on retraction practices in other fields, except medicine. There, a previous study also identified plagiarism as the largest factor leading to retraction but did not identify conflicts of authorship and lack of authorization by co-authors as a significant factor. This is difficult to interpret, though, and may indicate either that co-authorship is more formalized in the medical field or that editors are more likely to correct a paper following authorship disputes, rather than retract it.

This quantitative analysis of the recently retracted papers is clearly only a first step in providing insights into the practices of retraction in chemistry and materials sciences. In order to better understand the current status of retractions in the publishing model, it will have to be complemented with sociological investigation into the views of all parties involved, the motivations behind these views, and surveys of academics in their various roles in the chain of publication: as authors, reviewers, editors, and publishers.

■ METHODS

I gathered data by a Scopus search on “retraction” or “retracted”, for the years 2017 and 2018, limited to the fields of chemistry, materials science, and chemical engineering (as identified by the Scopus “subject areas”). I curated by hand the 362 results obtained, to remove the use of “retraction” in other contexts (e.g., retraction of polymer chains), and the result was a list of 331 unique retracted papers. For each paper, I gathered using the Scopus API the DOI of the retraction notice, the DOI of the original paper (which are sometimes the same), the authors’ names, and the affiliations (and in particular the country of each affiliation).

Because there is no simple API for the retrieval of the text of retraction notices, I retrieved the notices manually from each article’s webpage. Once the whole data set was assembled, each retraction notice was manually classified to identify the following possible reasons stated as being the cause(s) for, or factor(s) in, the retraction:

- Authorship issues:
  - Lack of approval to publish from some of the co-authors
  - Author missing from the author list
  - Error in the affiliations listed
  - Authorship disputes or unclear authorship
- Errors in the science reported:
  - Experimental error
  - Error in the analysis, interpretation, or mathematical derivations
  - Work that was not reproducible by the authors or other researchers
- Issues with the data
  - Data falsification or mishandling, clearly identified as such
  - “Problems” or “issues” with the data, for unclear reasons
  - Issues with data ownership, or authorization to publish (confidentiality, etc.)
- Plagiarism
  - Duplicate publication, identical or near-identical papers
  - Self-plagiarism (unacceptable reuse of material previously published by some of the authors)
  - Plagiarism of other sources (not by the authors)
- Publisher error

DOI 10.1021/acs.chemmater.9b00897
Issues of copyright
Problems with the integrity of the review process
I also took note of the occurrence, in the retraction notices, of the following features:

• Authors that could not be contacted or did not respond during the retraction process
• Mention of an institutional inquiry into research misconduct by one or more of the authors’ affiliations
• Statement detailing the individual responsibility or role of the authors in the publication of the paper or during the retraction process.

François-Xavier Coudert
Chimie ParisTech, PSL University, CNRS, Institut de Recherche de Chimie Paris, Paris, France

ASSOCIATED CONTENT

Supporting Information
The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.chemmater.9b00897.

Classification of each retracted article (XLSX)
All data gathered (XLSX)

AUTHOR INFORMATION

ORCID
François-Xavier Coudert: 0000-0001-5318-3910

Notes
Views expressed in this editorial are those of the author and not necessarily the views of the ACS.
FXC is Associate Editor for Adsorption Science & Technology (SAGE Publications) and Editorial Board Member for Communications Chemistry (Springer Nature). He has not, in these capacities, handled any retraction procedure.

All data gathered, including the classification of each retracted article, is available in the repository at https://github.com/fxcoudert/citable-data/ as well as in the Supporting Information.

ACKNOWLEDGMENTS

Nicolas Jonckheere is gratefully acknowledged for critical reading and comments on a draft of this paper.

REFERENCES

(2) Available online at https://pubpeer.com/.
(17) Fuyuno, I.; Cyranoski, D. Cash for papers: putting a premium on publication. Nature 2006, 441, 792.
(18) Oransky, I. Pair of nanotech researchers up to at least two dozen retractions. RetractionWatch, available online at https://retractionwatch.com/2018/12/04/pair-of-nanotech-researchers-up-to-at-least-two-dozen-retractions/.
(20) Stern, V. After Elsevier knew an author faked reviews, it kept accepting his papers for more than a year. RetractionWatch, available online at https://retractionwatch.com/2018/01/04/elsevier-knew-author-faked-reviews-kept-accepting-papers-year/.
(21) Stern, V. Top physicist loses another paper; tally now up to 7. RetractionWatch, available online at https://retractionwatch.com/2017/05/12/top-physicist-loses-another-paper-duplication-tally-now-7/.
(27) In the case of a simple error on the author list or affiliation, the COPE guidelines recommend a correction, if the validity of the content of the article is not in doubt.