Promoting Integrity as an Integral Dimension of Excellence in Research

***

Individual Case Studies

DOCUMENT DESCRIPTION

Deliverable Number	D3.5
Work Package	WP III – What Happens in Practice? Institutional Responses to Misconduct.
Task	T III.3 – Inventory: in-depth misconduct case studies
Type	Report
Version	Final-reviewed
Number of Pages	176
Due Date of Deliverable	Month 18, 02/017
Actual Submission Date	Month 27, 20/11/2017 Resunmitted Jan 2019
Dissemination Level	Public

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 665926.


PRINTEGER

Table of Contents

Place holder

The cases studies involving individual cases of misconduct were performed and shared among project researchers. However, since these case descriptions included sensitive information about these individual cases, we could not share these descriptions publicly. A sumarising report of the case studies is available as a separate deliverable.

In this document, we only report the studies of journals, which did not include sensitive information about individual researchers.

Leiden University – Case Study 1
Predatory Publishing and the Imperative of International Productivity

Authors: Sarah de Rijcke (Centre for Science and Technology Studies, Leiden University), Tereza Stöckelová (Institute of Sociology, Czech Academy of Sciences)

1. Summary

This report discusses a case of how predatory publishing, gaming of metrics, and exploitation of the model of gold open access can be partly understood as a logical response to the requirement of internationalization. Predatory publishing enacts an alternative mode of internationalization for those researchers and institutions who fail – for better or worse—within the established mode of international, with its epistemic and economic centers in the global, Anglophone North / West. A recent misconduct case in the Czech Republic shows how the imperative of internationalization and productivity inscribed in the country's research assessment framework impinges on institutional and individual publication strategies, and produces a market for gaming in the academy. The report concludes that purification and policing efforts are often based on the ideal of a unified science system, with internationally shared views 'from nowhere' about what constitutes 'bad' and 'proper' scientific conduct. This ideal is flawed, because different actors within science systems create and re-enforce distinctive normative hierarchies between the international, the national, and the local: journals, databases, evaluators, consultants, publishers, and also researchers.

2. Introduction

The requirement to 'be international' cannot be overlooked in current European research policy and research evaluation. The imports of 'internationalization' figure prominently in how value is given to "international visibility", "international impact", or to the international character of publication venues. The international is used as a trope in EU funding schemes, in project goals that guide national assessment exercises, in output measurements, in the formulation of institutional research missions, and in tenure-track criteria. Particularly in smaller countries (e.g. the Netherlands and the Czech Republic) the international is often taken as a proxy for quality, proving impact beyond the "academic pods." Consequently, the international, the national and the local constitute a clear normative hierarchy. For example, it is taken for granted that international excellence encompasses national excellence and (as such) is supposedly more valuable.

In this report we discuss a case that demonstrates how predatory publishing, gaming of metrics, and exploitation of the model of gold open access (Beall 2012), can be partly understood as a logical response to the imperative of internationalization. It enacts a different, yet debatable alternative mode of internationalization for those researchers and institutions who fail—for better or worse—within the established mode of international, with its epistemic and economic centers in the global, Anglophone North / West. We zoom in on a recent misconduct case in the Czech Republic to show how the imperative of internationalization and productivity inscribed in the country's research assessment framework impinges on institutional and individual publication strategies, and produces a market for gaming in the academy.

3. Gaming the system

In 2015, a major debate on publishing and research evaluation surfaced in the Czech academy. It was triggered by controversy over a very productive junior researcher at the Faculty of Social Sciences of Charles University. At first sight he seems a paradigmatic case of a successful scholar with a long list of international publications, collaborations and co-authorships - exactly what the current research policy in the Czech Republic holds as a normative ideal. However, on second sight and when some of his colleagues started to closely scrutinize his production, the case turned out to be something different: an attempt to game the current research assessment system on various levels - or rather, to take the imperative of the system to the extreme by some perfectly legitimate and some less legitimate ways. To understand what happened, the genesis and current state of research assessment in the country needs to be described first.

Various forms of competitive funding of academic research were introduced with the the establishment of the Czech Grant Agency in 1992. This was followed by the introduction of a new methodology for the quantitative assessment of institutional-level research performance in 2001. This was largely at the initiative of a few natural scientists who came back to the Czech Republic in 1990 after spending several years in the West. Its impact on research funding of academic institutions and on performance assessment of individuals has since then gradually increased. The central building blocks of the Evaluation Methodology are so-called RIV-points (RIV standing for "Information Register of R&D results"[i]), assigned to pre-defined types of outputs (journal articles, monographs, patents, prototypes et cetera) and meant to reflect their academic and user value (Office of the Government of the Czech Republic 2013).[ii] One of the key claimed rationales of the Evaluation Methodology was to create objective criteria that would increase the transparency of the research system and depoliticize its governance. However, during the last 15 years the Methodology developed into a convoluted metrics-based amalgamation with many unclear algorithms and weights that are far from transparent not only for 'ordinary' researchers but even for research policy managers at the national level (cf. Miholová and Majer 2016). At present the Evaluation Methodology's criteria for 'quality recognition' soak through the entire system. They have a significant - even if at times indirect - impact on academic hiring and promotion procedures, individual research grant endowment, and the funding allocation of public research institutions.

A key trope of the research policy reforms since the 1990s has been internationalization, and this trope is also inscribed into the current Evaluation Methodology. This is understandable in a small country where many disciplines tended to operate in closed circles consisting of local scholars. However, it is more problematic that the international oftentimes stands as a value in itself—unquestioned and undisputed. E.g. there is currently nearly no peer review evaluation of journal articles within the national evaluation framework (a peer review evaluation of a limited number of outputs submitted by research organizations as 'excellent' was introduced in 2015) and the journal impact metrics provided by Web of Science and Scopus are taken for granted as proxies for international recognition and quality. This is the context in which junior academics start to build their publication record and careers.

We now return to the controversy. Having gained his PhD in 2007, the academic in question has claimed to have (co)authored or (co)edited seventeen 'scientific monographs' between 2011-2013 and more than eighty journal articles between 2006-2015.[iii] Apart from the extreme productivity, four aspects of his CV are noteworthy. Firstly, the author also acts as an editor in chief, editorial board member, and even publisher of some of the 'European' or 'International' journals listed on his CV.

[iv] All these journals are English language and target an international audience, have an international review board and international pool of authors. Secondly, even if in SCOPUS, some of the journals on his publication list were also listed in Jeffrey Beall's database of predatory journals.[v] Thirdly, some of the co-authors on these articles in predatory journals were colleagues from the faculty - including the current head of the department.[vi] And, finally, as the author later confirmed, one of his co-authors was discovered to be a fictional character supposedly affiliated with prestigious Western European universities (first the University of Strasbourg and later the University of Cambridge).

While some of the academic's actions were rather extreme, they stayed in line with the current imperative of internationalization. The researcher tried to gain 'Western' recognition and certification (listing on the WoS and Scopus databases) for his publishing activities as an author, editor, editorial board member, and publisher based in the East. Interestingly, he not only strove to gain a position in the existing international playing field (which is what the research policy framework in fact tries to encourage) but also, as a skillful academic entrepreneur, to rework and reorder the field at one go by creating new journals and forging new East–West alliances (even if at times with fictitious co-authors). He also specifically offered his teaching and publication 'services' to researchers from Russia and Eastern Europe in relation to whom he positioned his activities as international. Apparently, he aimed at the enactment of a different international than the one of the current global science, in which the international in fact equals the West.

As a result of a major controversy at the faculty level, during which 'whistle-blowing' colleagues from the department filed a complaint to the Ethical Commission of Charles University (the complaint was deferred[vii]), and following the publication of a number of articles in national public media, the author's contract was terminated in September 2015. In response to the increasing media and academic community pressure, the Faculty openly distanced itself from unethical publishing practices connected with the case by issueing 'publication rules' which warned against predatory journals and vanity press publishers. Some other faculties and universities in the country followed suit.

Interestingly, the 'international' standards for quality assessment did not seem to count equally for all involved. A few weeks after the termination of the perpetrator's contract, the contract of the main whistle-blower was not renewed either. The faculty chiefly adhered to a 'bad apple' approach, a relatively common strategy in misconduct cases in the sense that measures are often taken mainly at the level of individuals.

4. Qui bono?

Calls for more transparent, trustworthy quality control mechanisms and more open infrastructures for communicating and publishing research are currently widely heard in European science policy. The European Commission has introduced several framework programs that focus in particular on responsible research and innovation, and on 'open science'. In 2020 all scientific and scholarly output should be freely available by way of open access. Another important aim for 2020 is a fundamentally novel approach to data (re)use, based on open data models. But change will not come easily, with vested interests of established academic elites and large commercial actors with their entrenched infrastructures for publishing and evaluating research. Paradoxically, part of the answer seems to lie in the hands of the exactly these commercial parties. At present they appear to be the ultimate gatekeepers of the 'international'. The critique of predatory journals inadvertently makes a very strong case for the value added by corporate, indexed outlets and black-boxed, commercially

endorsed algorithms. Predatory journals seem to play right in the hands of corporate publishers as a confirmation of the dangers of uncontrolled open access.

At the same time the predatory publishing industry managed to develop a business model that taps into both the 'open science' *and* the 'commercial' publishing models and normative frameworks. Evidently, some of the appeal of predatory journals and vanity publishers lies in their offering cheap, accessible vehicles for the 'international'—certainly when compared to the costlier 'gold' open access publications, with quality control and more or less US- and Eurocentric gatekeepers. Also, the predatory publishing business model closely mimics and reproduces the standards and incentive structures of the 'global', dominant publishing industry. This is an industry in which the journal and the journal article are the most valuable means of communication for international recognition and visibility, within a 'market world of justification' (Boltanski and Thevenot 2006) which is enacted, among other things, through indicators such as the Journal Impact Factor (Rushforth and De Rijcke 2015, Muller and De Rijcke 2017). Publication practices of predatory publishers are being linked to the most important and profitable value systems of the dominant publishing industry and the indicator production market. As such, predatory publishing and its concomitant practices are not outside of the research system but emerge at the heart of them and are embedded within them. These practices in effect drive the existing evaluation logic to the extreme. A crucial question then becomes: *qui bono* (cf. Star 1995), who actually benefits from this industry?

In the CR and further East, the predatory journals and vanity presses play a role in further empowering skillful local researchers who used the new industry to boost their publication records, international visibility and the financial status of their institutions (for instance by gaining RIV points for books published by international 'vanity' presses). The at the first sight useful term 'predatory publishing' or 'predatory journals' may be largely misleading, because it obscures much the agency of individual actors in using these outlets to their advantage. In the case at hand scholars were hardly 'prey', as they found clever ways of gaming the assessment system.

The Czech case makes clear how the predatory publishing industry thrives mainly by being successfully parasitic on existing forms of conduct and material infrastructures for publishing and evaluating research—without fully incorporating its quality control mechanisms (absence of 'proper' peer review, fake editorial boards, et cetera). But this lack of explicit quality control procedures should not be overemphasized. Some of them apparently have some quality control and rather than belonging on a blacklist they operate in a grey zone—into which some established quality journals may now be falling as well with the increased global pressures on production and auditable performance, which deprives the publication system of available competent reviewers and editors. We think the *excessive parasitism* of the 'predatory' journals is much more crucial. Many of them deliberately operate on the edges of dominant publication and citation infrastructures, hosted by big commercial publishers. A lot of these journals originate from the 'East', and these journals permeate the 'global' publishing industry when they are indexed in the WoS and—particularly—Scopus.[viii] The latter's reputation is based on being the 'largest abstract and citation database of peer reviewed literature', providing a 'comprehensive overview of the world's research output'.[ix] This is a problematic statement: the company cannot in practice control this international certification, and is nonetheless taken as proxy for quality in many evaluation systems.

5. Conclusion

Although the critique of predatory publishing does indeed lead to some sanitization efforts (codes of conduct, blacklists and whitelists) thus far it has not triggered any serious kind of more radical reform of the publishing and evaluation infrastructure. This may partly be because it is too soon. It could also be due to the fact that purification and policing efforts are often based on the ideal of a unified science system, with internationally shared views 'from nowhere' about what constitutes 'bad' and 'proper' scientific conduct. Such an ideal is doomed to fail when we see how different actors within science systems create and re-enforce distinctive normative hierarchies between the international, the national, and the local: journals, databases, evaluators, consultants, publishers, and also researchers. Some assessment systems are in fact beginning to recognize the need for contextual evaluation (in terms of disciplines and fields) and the complex relation between the international, national and local. But there still is a long way to go before the research policy and wider academic communities acknowledge that the more, the faster and the more international need not be always the better.

References

Beall, J. 'Predatory publishers are corrupting open access', *Nature* 489 (2012): 179, doi: 10.1038/489179a.

———. "What I learned from predatory publishers." *Biochemia Medica* 27, no.2 (2017): 273-9. https://doi.org/10.11613/BM.2017.029.

Boltanski, Luc, and Thévenot, Laurent*On Justification. The Economies of Worth*. Princeton: Princeton University Press, 2006 [1991].

Good, Barbara, Niki Vermeulen, Brigitte Tiefenthaler, and Erik Arnold. "Counting quality? The Czech performance-based research funding system." *Research Evaluation* 24, no. 2 (2015): 91-105, https://doi.org/10.1093/reseval/rvu035

Lin, Wen-yuan, and John Law. "Making things differently: on 'modes of international'." CRESC Working Paper No. 129. Centre for Research on Socio-Cultural Change (CRESC), Faculty of Social Sciences: The Open University, 2013. Available at: http://hummedia.manchester.ac.uk/institutes/cresc/workingpapers/wp129.pdf.

———."A correlative STS: Lessons from a Chinese medical practice." *Social Studies of Science* 44, no. 6 (2014): 801–824.

Linková, Marcela and Tereza Stöckelová. "Public accountability and the politicization of science: The peculiar journey of Czech research assessment." *Science & Public Policy* 39, no.5 (2012): 618-629.

Macháček, Vít, and Martin Srholec. 'Predatory journals in Scopus', *Study* 2/2017, IDEA – CERGE-EI, Praha. Available at: https://idea-en.cerge-ei.cz/files/IDEA_Study_2_2017_Predatory_journals_in_Scopus/mobile/index.html

Miholová, Katerina, and Ondrej Majer. *Analýza mechanismu Metodiky hodnocení výsledků [An analysis of the mechanism of the Evaluation methodology]*. Government of the Czech Republic, Section for Science, Research and Innovation. Unpublished report, 2016.

Office of the Government of the Czech Republic. *Methodology of Evaluation of Research Organizations and Evaluation of Finished Programmes (valid for years 2013 - 2015)*. Available at: http://www.vyzkum.cz/storage/att/A7FE6F4477F5064B57B683C62C4A4CFD/Methodology.pdf

Rushforth, Alexander D., and Sarah de Rijcke. "Accounting for impact. The Journal Impact Factor and the making of biomedical research in the Netherlands." *Minerva* 53: 117-139 (2015).

de Rijcke, Sarah, and Ruth Müller. "Thinking with indicators. Exploring the epistemic impacts of academic performance indicators in the life sciences." *Research Evaluation* (2017), doi: 10.1093/reseval/rvx023.

Silver, Andrew.. "Controversial website that lists 'predatory' publishers shuts down Librarian Jeffrey Beall won't say why he has unpublished his widely read blog. " *Nature*, 18 January, 2017. Available at: http://www.nature.com/news/controversial-website-that-lists-predatory-publishers-shuts-down-1.21328

Star, Susan Leigh. "The politics of formal representations: Wizards, gurus, and organizational complexity." In Susan Leigh Star (ed.) *Ecologies of knowledge. Work and politics in science and technology*. Albany: State University of New York Press, 1995.

Stöckelová, Tereza. "Immutable mobiles derailed: STS and the epistemic geopolitics of research assessment." *Science, Technology & Human Values* 37, no.2 (2012): 286-311.

Leiden University – Case Study 2
Analysis of papers retracted in the Web of Science

Authors: Thed van Leeuwen, Andrea Reyes Elizondo, Sarah de Rijcke (Leiden University)

1. Summary

Work Package III gathers indicators of the extent of misconduct and analyses how institutions respond to misconduct or deviance in science. Deliverable III.3 focuses on the occurrence of failures in the science system, in which material is being retracted from the pool of published knowledge. Various aspects of the retracted literature is analysed, such as geographic and disciplinary spreading of the phenomenon, as well as the reasons and initiators for retractions from the literature.

2. Introduction

Science seems to be in crisis, as can be concluded from the discussion on the reproducibility of scientific results, both in the literature as well as in the mass media. Several international conferences have dedicated attention to the phenomenon (e.g., 4S in Boston, August/September 2017, STI 2017 in Paris, and the NWBRP in Helsinki, 2017). However, it is not a new phenomenon under debate as already in the 1980's scientists wrote about the issue of replication (Broad & Wade, 1982, Collins, 1985). As is shown by the length of discussions on the topic, the talk about a crisis in replication is more prominent since the early 2010s, when a number of highly prolific fraud cases surfaced. Since the reproducibility of experiments is an essential element in the production of scientific knowledge, the inability to replicate previous studies is potentially harmful for many disciplines whenever theories in those disciplines are based upon on non-reproducible experiments. The inability to replicate research findings can be one of the many reasons why some research papers are perceived as possibly flawed, since non-reproducibility can be a consequence of human error or sloppiness. As sloppiness or human errors are reasons to retract literature from the public domain, more clear reasons for retraction are straightforward scientific misconduct, whenever we speak about the so-called FFP, fabrication, falsification and plagiarism, or about QRP, Questionable Research Practices. FFP is often considered as more serious as compared to the QRP part of studying science and the actors in the science system. FFP is considered as more serious as it involves incorrect outcomes from research, misleading the scientific audience, or stealing intellectual properties from colleagues in the research process, while QRP does not necessarily lead to incorrect outcomes, it is more about how these are generated.

By publishing their results in the scientific literature researchers claim the ownership of particular ideas, processes, and results, made public via journals, books, chapters, conference papers, and talks, to name a few. Prior to publication, the contents of scholarly manuscripts is made accessible for scrutiny to other researchers from the same discipline. Before a journal with an international standing accepts a manuscript it is reviewed by one or more reviewers. This peer review process, a process in which many submitted manuscripts can be rejected for various reasons, for example the research is not in the scope of the journal, the reported research is flawed, the work is reported in a bad or insufficient manner, etc.. Finally, manuscripts are either accepted or rejected by a journal. Once accepted, the manuscripts become included in the journal, and made publicly available. However, this peer review process in which manuscripts are being judged can be flawed. If

results presented in the scientific literature turn out to be not trustworthy, the publications are (partially) retracted. In principle these papers are removed from the common stock of knowledge. There are many reasons for retracting a paper, scientific misconduct being only one of them. However these cases attract a considerable amount of attention and undermine the public trust in science.

The scale enlargement through which modern science went over the last 40-50 years, with a development from smaller communities of scientists to the large-scale big science (Price, 1963) or techno-science (Ravetz, 1971, Latour, 1999) has been confronted with internal mechanisms for controlling the quality produced. The 'Publish-or-Perish' culture, that affected global science from the 1980's onwards, created an ever increasing growth of scientific outputs (not to the level predicted by Price, but still rapidly increasing). This growth of global science and its outputs has also had serious consequences for the quality control systems in science, as this has put increasing pressure on the peer review system. So while the Mertonian perspective of academia being a sphere consisting of an exchange of favors (as no money was involved, and the act of reviewing was considered as mutual favor), this changed due to the increasing pressure on the peer review system in particular, but on the scientific system as a whole. Next to this development, we now observe a further marketization or commodification of academic work (Mirowski, 2011, Radder, 2010). Increasingly, the production of new knowledge, and the way this is disseminated, is considered as a requirement to pursue an academic career. With outputs being rewarded, and reviewing not, a certain degree of misbalance might occur. This increasing pressure on peer review might result in a further decay of the quality control system.

Over the last couple of years the retraction of publications became a popular subject in science studies (Grieneisen & Zhang, 2012 and references therein). Authors studied mainly retractions of papers processed for MEDLINE or subsets of retracted papers published in journals processed for Clarivate Analytics' Web of Science (WoS). There is a general perception that the number of retracted papers is increasing over the years and thus scientific misconduct has become a more relevant topic (Steen, Casadevall & Fang, 2013 and references therein).

The study's aim is to provide a comprehensive analysis of all retracted papers published in journals processed for WoS between 1981 and 2015. A detailed classification is produced of the motives for the retraction as stated in the retraction notices and of the parties initiating the retraction form the journal, two subjects paid little attention to in previous work.

3. Methods and data

In the WoS database the suffix "Retracted article" is added to the title of articles that are officially retracted (Chen et al., 2012). Using this information on publications processed for WoS, we collected in total 3729 publications. The dataset collected that resulted from this step was used to make a bibliometric analysis of the retracted papers. When analysing the geographical spreading of retracted publications, we present the results of the distribution of the countries mentioned in the papers' address by-line. A full counting scheme at country level is applied.

The journals processed for the WoS are assigned at least to one subject category. The retracted papers are assigned to the subject category or categories of their journals. These WoS Journal Subject Categories form the basis for a classification of science on a somewhat higher level of aggregation, namely that of scientific disciplines. The distribution over the scientific disciplines of the

retracted papers is compared with the distribution of all papers processed for the WoS across those same scientific disciplines. Again a full counting scheme is used: a paper published in a journal assigned to several subject categories, is counted in full to each category, and thus to each discipline.

In the title of the retracted article, the suffix "Retracted article" is followed by the bibliographic data of the retraction notice (volume, page, year). We manually retrieved the pdf-version of the retraction notices from the Leiden electronic library or, if not available, we search for the hard copy. One notice in a journal may contain information on more than one retracted article in that same journal, solving various problems in one notification, which has the advantage of having to come clear only once, instead of in multiple notifications. Another entry on the data was formed by the doi's of the publications labelled as retractions. The doi, the digital object identifier was a helpful tool in a clear distinction on the notifications of the retractions.

A previous study of the current one (van Leeuwen & Luwel, 2014) was based upon a much smaller sample of retracted publications as compared to the current analysis. In that previous analysis, we also focused on both initiators and reasons for retractions. From the retraction notices the information on the party responsible for the retraction as well as the reason of retraction (further called retraction type) is retrieved. A classification scheme was developed for these items. What became apparent in the current study was that while we classified the reasons and initiators in a simple straightforward method (that is, one party initiating retraction, for one single reason), we observed this to be much more complex, in the sense that initiation for retraction occurs frequently based on two or more parties initiating the retraction, while the reasons for retraction can also be multiple. In that previous study we allowed for example the authors and editors to be labelled together, while we now observe that the initiating parties of retractions can be much more varied, and multiple (up to even four parties initiating or agreeing upon retraction). A similar issue is at stake for the reasons of retraction. In the previous study, we classified the reasons under one topic alone, while we now observed a much more varied situation, in which often various issues play a role at the same time. For example, fabrication, manipulation, and falsification of data are reported sometime similarly as reasons for a retraction. Similarly, issues around data manipulation, interpretational, and data unreliability are reported as reasons for a retraction. These combinations are multiple, and varied. So while we tried to classify the reasons for retraction under one umbrella in the previous analysis, we now observe these reasons often to coincide with other reasons. Important to notice here is the fact that the boundaries between FFP and QRP seem to be more fluid than often suggested. For example, if plagiarism is defined as the usage of text, procedures, outcomes, and ideas created by people who are not properly acknowledged, how does that relate to discussions on authorship, e.g., the exclusion of somebody from an author list whose work has substantially contributed to the final outcomes of a research project?

It is well known that retracted papers are cited often years after the retraction date (see e.g. Chen et al., 2012). Although interesting, the number of citations gives little information of the continuing impact of these publications. For a few highly publicised fraud cases we plan to carry out a sentiment analysis of citations to retracted papers before and after retraction (Li et al., 2013).

4. Results

In Figure 1 we show the development of publications published in the WoS covered journals, which get retracted later onwards. The trend that is visible indicates a sharp increase in the period between 1995/2007. From 2007 onwards, we observe a number of peaks, in 2007, in 2010, in 2012, and to a lesser extent, in 2014. The fact that we observe somewhat of a decreasing number of retractions is not a positive development, this simply relates to the fact that it takes some time before erroneous/fraudulent research outputs get recognized. But we will get into this later on in this section.

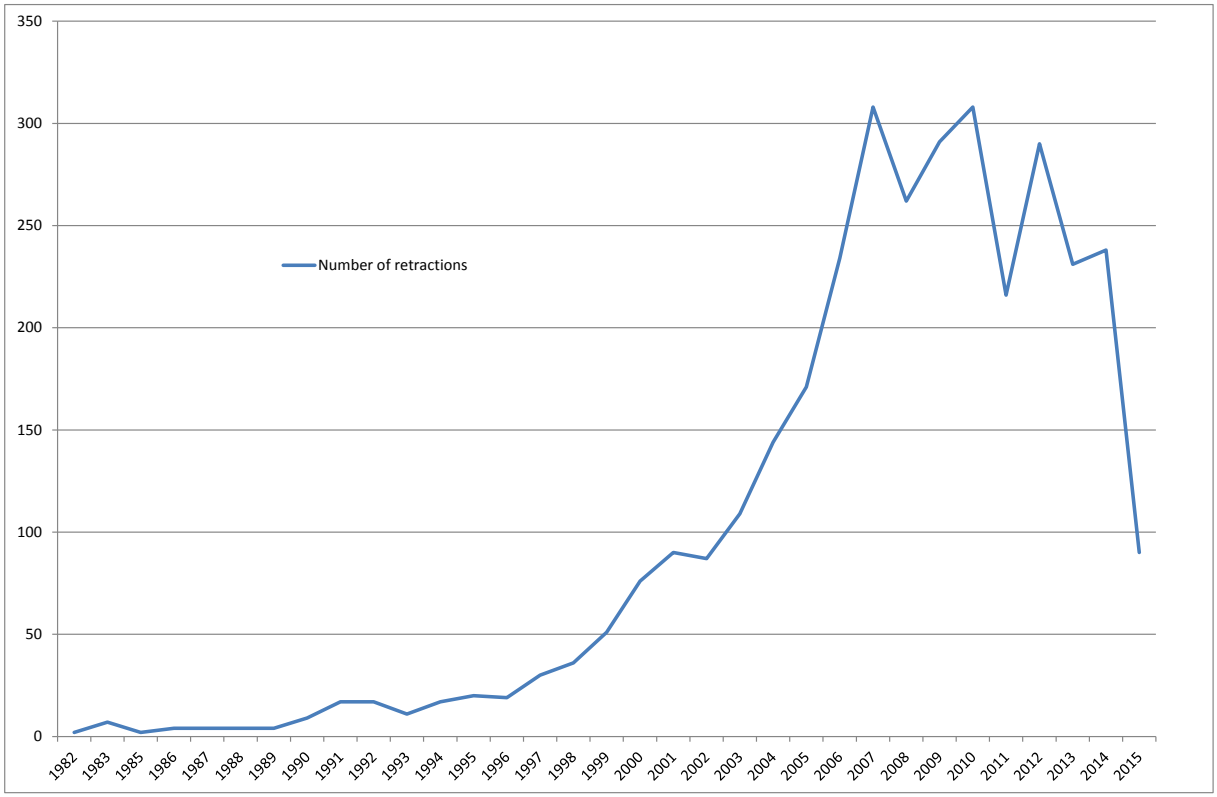

Figure 1: Number of retracted publications from WoS covered journals, 1981-2015

Figure 2 shows that the highest number of papers retracted occurs in the first year after publication, while we simultaneously also notice that a long tail of single occasions get retracted over a longer period. The accumulation of retractions over time shows that 90% of all publications that get retracted are retracted within a six year time frame.

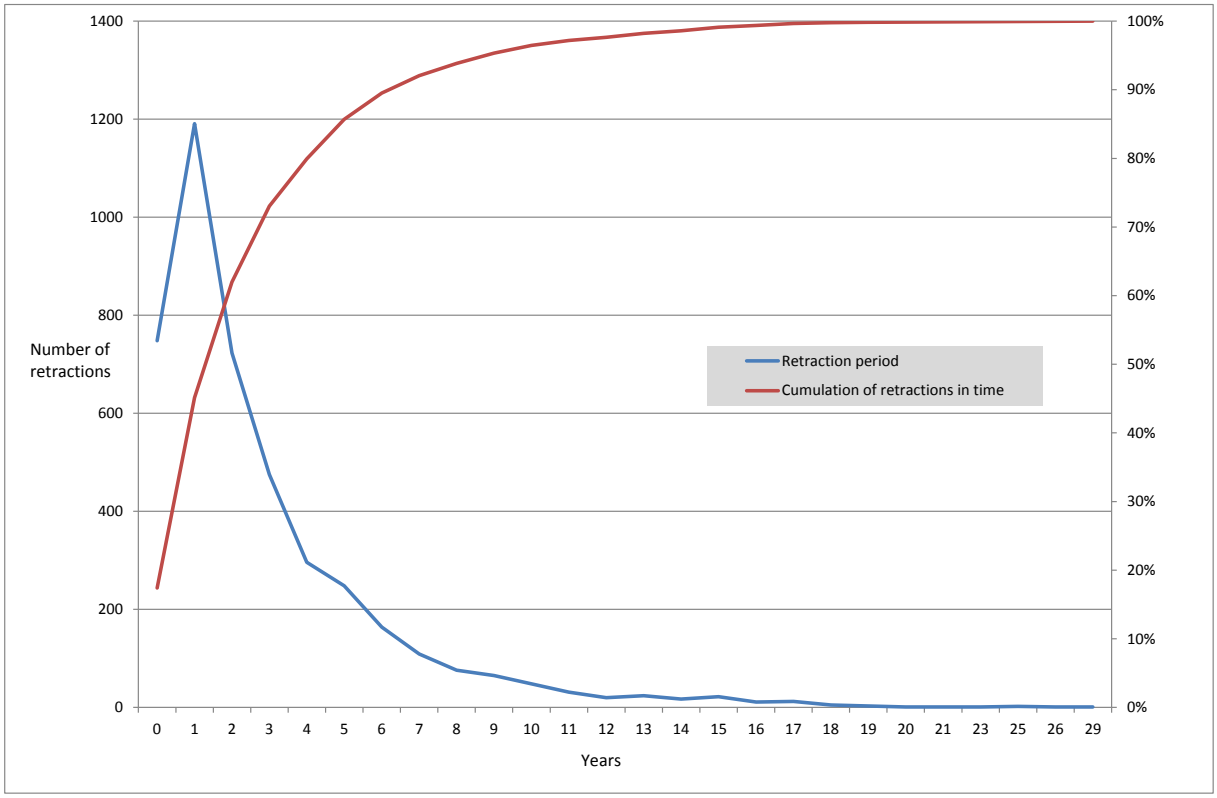

Figure 2: Period of retraction: time before publications get retracted from WoS covered journals, 1981-2015

Another characteristic of the total body of publications that is retracted relates to the actors involved. Often retractions and research integrity/scientific misconduct is put into the 'bad apple' metaphor, or sometimes as the 'lone wolf'. Here we analyse the involvement of more people per paper retracted (which does not mean these people are all involved in fraudulent behaviour, more people involved implies a higher chance of notifying the fact that something is wrong). So the number of authors involved is important, as a stepping stone towards the analysis of countries involved in retracted material as well as the various types of scientific activity involved.

In Figure 3 we present the distribution of the numbers of authors involved in retracted literature, as well as the cumulative shares of the total. In the line indicating the number of authors per paper, we find that most publications that are retracted do have three authors attached to them. Combining the two lines, we observe that 67% of all retracted publications have 1-5 authors, while the remaining 33 % of all retracted publications have up to 33 authors (95% of all retracted publications have up to 10 authors). A conclusion could be that the lone wolf as a single author causing retraction (we avoid for now the reasons for retraction) is only limited as not even 10% of all retracted publications have one single author attached to them, while in total only 40% of all publications that are retracted carry up to three author names.

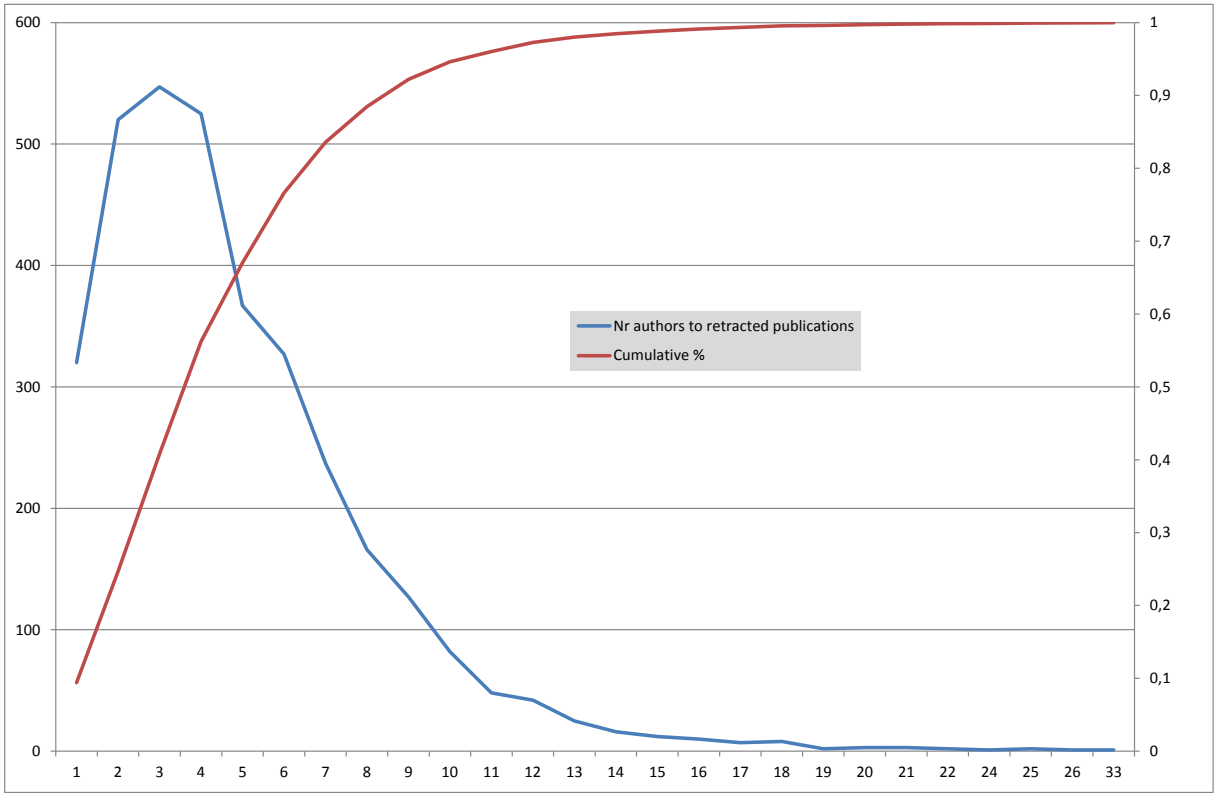

Figure 3: Distribution of number of authors over publications in WoS covered journals, 1981-2015

Next, we analysed the various types of scientific activity connected to the retracted literature. We defined three mutually exclusive types of scientific activity, that is papers that carry

only one address (so we then conclude that no scientific cooperation is at stake here), indicated as Single Institute or SI, publications that stem from international cooperation as can be concluded by the occurrence of two or more country names in the address by-lines of the papers (indicated as International Cooperation, or IC), and the remaining part, which carries multiple addresses, but all from one country (interpreted as national cooperation, indicated as NC). One should note that this latter category NC includes intramural research, while the IC category might also include various address from within one country (which means that IC then also partially covers NC). So mutual exclusiveness applies within the definition, reality is of course more complex, but for the sake of clarity, we defined these three types of scientific activity. In this analysis we integrated something from the previous analysis, underlying Figure 2, while introducing single authored publications in Figure 3.

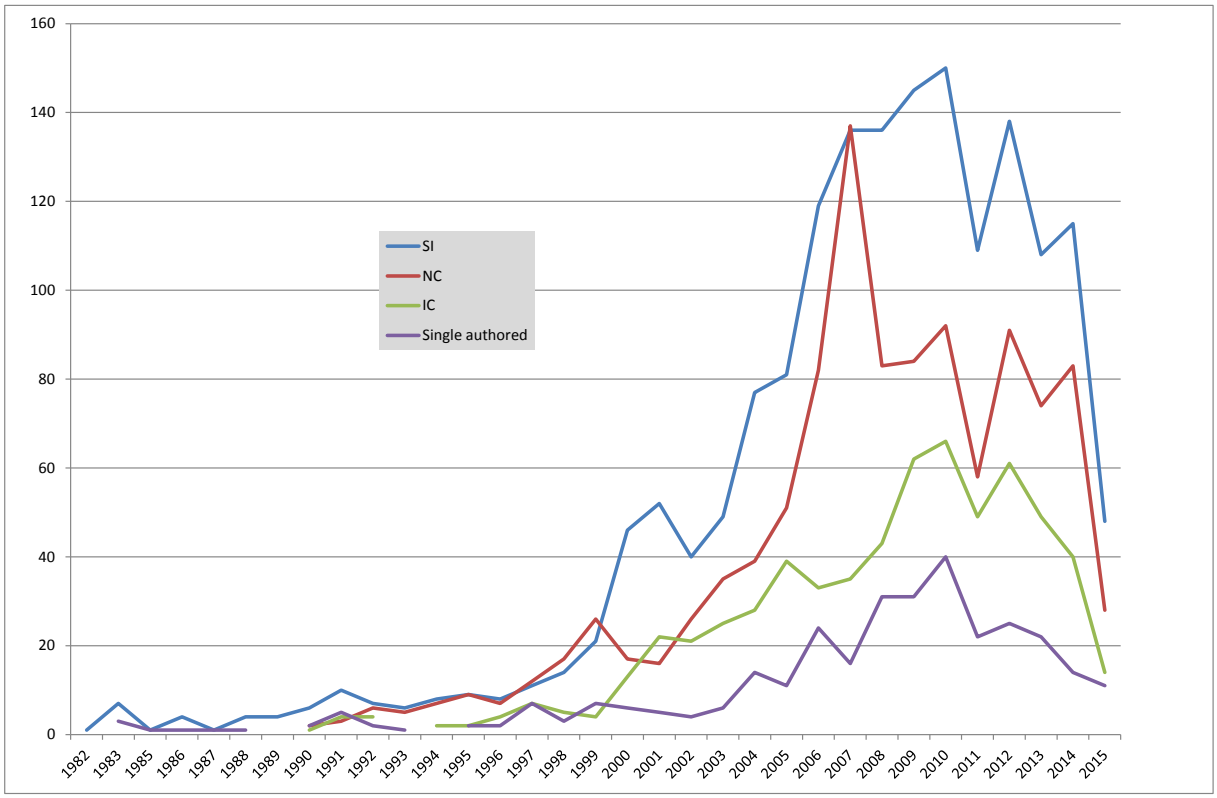

Figure 4: Distribution of number of authors over retracted publications in WoS covered journals, 1981-2015

Figure 4 shows us that single authored publications are only a small part of all publications which carry only one address. So if publications that carry only one address are retracted from the literature, this often involves publications that carry several author names. A next observation is that most publications that get retracted have a national rather than an international component. This seems to indicate that publications resulting from international cooperation have a much lower likelihood of getting retracted.

In general, when conducting similar analyses, we find that publications resulting from international cooperation outnumber the outputs resulting from both single address of national cooperation. In this case, we find the opposite, which could be interpreted as indicative of the fact that international cooperation in relation to scientific misconduct has a prohibitive effect: due to the variety of national cultures, the possible availability of certain protocols due to regulations by international funding agencies tighten the situation to such an extent that conducting scientific misconduct might become more difficult.

Figure 5 shows the countries which are most often mentioned in the address by-line of the retracted publications. Papers (co)signed by authors from the USA, PR China, Japan, Germany, and India are in the top five of most retracted publications per country. Particularly the USA and PR China outnumber the other countries with retracted publications, as can be expected for the two largest countries in terms of output. For most other countries the number of retracted publications is much smaller, on a year-by-year basis.

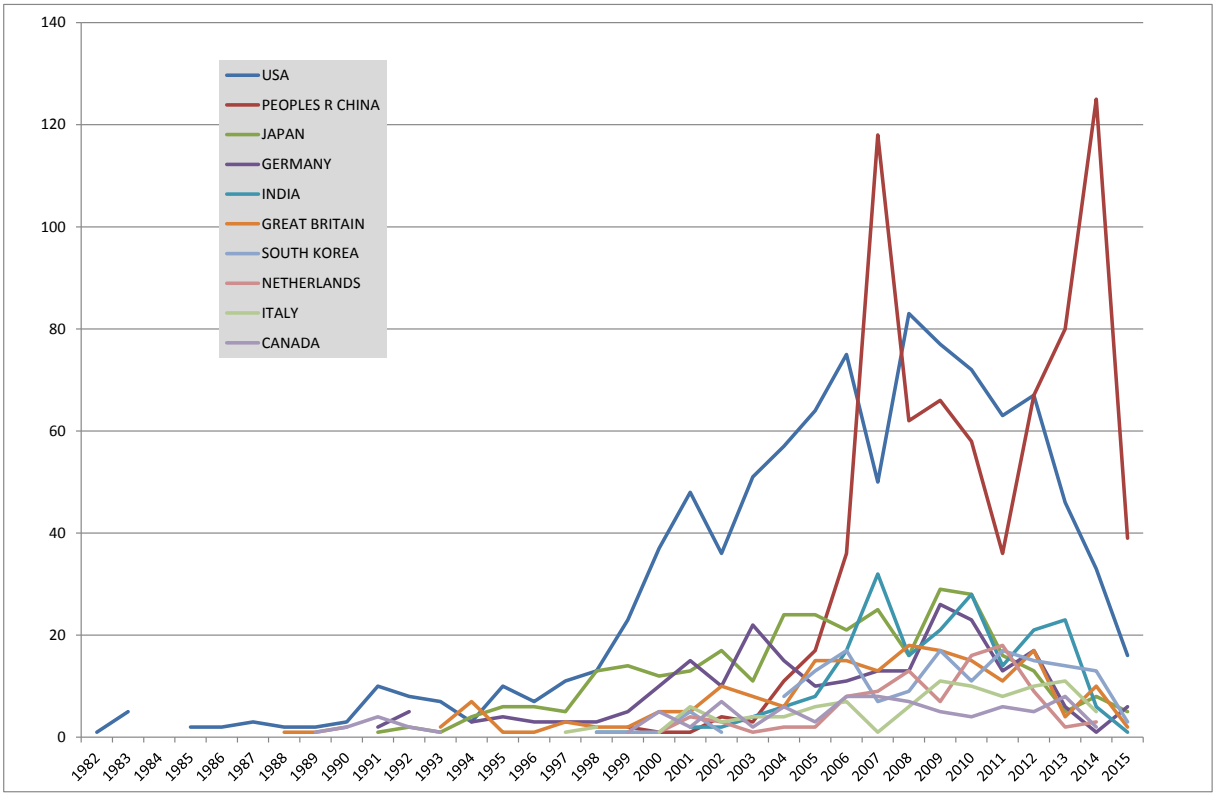

Figure 5: Distribution of number of retracted publications over countries in WoS covered journals, 1981-2015

Table 1 gives the countries which are most often mentioned in the address by-line of the retracted publications. In the table, we show the comparison between the regular ranking of countries based on overall output numbers, with the number of retracted papers, and the way that creates a ranking of countries. Papers (co)signed by authors from especially the PR China, India, the Netherlands, and Iran are retracted more frequently compared to these countries' share in the total

number of publications processed for the WoS. Overall for each country the fraction of retracted papers in its total number of WoS papers is very small, but it varies between countries by a factor of 4. The main reason for these countries to appear more prominent in this rank order coincides with the structuring of academic rewards systems. Chinese academics get rewarded for publishing a certain number of publications in journals with a certain status or reputation (read: those which carry a Journal Impact Factor) before getting into a next phase in their tenure track, while Stapel indicated the 'publish-or-perish' culture in the Dutch academic system as the main reason for his misbehaviour (Stapel, 2012).

	Retracted publications 01-15	Rank	Publications 01-15	Rank	% country/world	% Retracted of country	% Country of retracted
USA	838	1	5161565	1	26%	0,0162%	22%
PEOPLES R CHINA	723	2	1943552	2	10%	0,0372%	19%
JAPAN	255	3	1178425	5	6%	0,0216%	7%
GERMANY	201	4	1320554	4	7%	0,0152%	5%
INDIA	201	5	578237	12	3%	0,0348%	5%
GREAT BRITAIN	166	6	1442576	3	7%	0,0115%	4%
SOUTH KOREA	150	7	543960	13	3%	0,0276%	4%
NETHERLANDS	97	8	450917	14	2%	0,0215%	3%
ITALY	92	9	772082	9	4%	0,0119%	2%
IRAN	87	10	207817	22	1%	0,0419%	2%
FRANCE	78	11	943925	7	5%	0,0083%	2%
CANADA	73	12	791212	8	4%	0,0092%	2%
AUSTRALIA	67	13	589655	11	3%	0,0114%	2%
SPAIN	61	14	639960	10	3%	0,0095%	2%
TURKEY	50	15	298661	19	2%	0,0167%	1%

Table 1. Country of affiliation of the authors of retracted papers for countries between 2001-2015

Next, we shift our focus to the scientific domains which are affected more by retracted literature. For this purpose we linked the journals in which the retracted publications appeared to 35 larger disciplines of science, by coupling the journals and the subject fields as these are linked in WoS. These disciplines were first used in the Dutch Observatory of Science & Technology (NOWT, 2010), and have as such shown their policy relevance. Note that journals are often classified in several subject fields in the WoS, and as such contribute to more than one discipline. For now, we focus on the ten most occurring disciplines when it comes to retractions. The disciplines with most retractions are Clinical Medicine (32%), Basic Life Sciences (21%), and Biomedical Sciences (15%).

Next we find two large disciplines in the natural sciences realm, Physics & Materials Science and Chemistry & Chemical Engineering, with 14% and 8% respectively. The next discipline is the set of journals classified as Multidisciplinary Journals, which contain well known journals such as Nature, Science, and the Proceedings of the national Academy of Sciences of the US. As this is relatively small domain, it is the more remarkable it appears among larger disciplines in this analysis.

	retracted pubs 01-15	Rank	Pubs 01-15	Rank	% field/world	% Retracted of field	% field of retracted
CLINICAL MEDICINE	1193	1	5306622	1	27%	0,0225%	32%
BASIC LIFE SCIENCES	798	2	2446005	4	12%	0,0326%	21%
BIOMEDICAL SCIENCES	561	3	2215493	5	11%	0,0253%	15%
PHYSICS AND MATERIALS SCIENCE	525	4	4008201	2	20%	0,0131%	14%
CHEMISTRY AND CHEMICAL ENGINEERING	305	5	2997733	3	15%	0,0102%	8%
ENVIRONMENTAL SCIENCES AND TECHNOLOGY	159	6	1167362	6	6%	0,0136%	4%
MULTIDISCIPLINARY JOURNALS	142	7	361654	17	2%	0,0393%	4%
BIOLOGICAL SCIENCES	126	8	1099492	7	6%	0,0115%	3%
PSYCHOLOGY	86	9	432642	15	2%	0,0199%	2%
BASIC MEDICAL SCIENCES	83	10	360890	18	2%	0,0230%	2%
MATHEMATICS	82	11	701868	11	4%	0,0117%	2%
AGRICULTURE AND FOOD SCIENCE	81	12	710176	10	4%	0,0114%	2%
MECHANICAL ENGINEERING AND AEROSPACE	75	13	605195	13	3%	0,0124%	2%
ECONOMICS AND BUSINESS	58	14	339777	19	2%	0,0171%	2%
ENERGY SCIENCE AND TECHNOLOGY	54	15	386264	16	2%	0,0140%	1%
ELECTRICAL ENGINEERING AND TELECOMMUNICATION	53	16	866373	8	4%	0,0061%	1%

Table 2. Scientific disciplines related to the journals publishing the retracted papers , 2001/2015

Figure 6 shows the development over time of the ten disciplines with highest intensity of retractions. As we can see in the overall analysis of retracted publications, or by scientific cooperation, or country, the main uptake of publications getting retracted seems to be happening since the Millennium change.

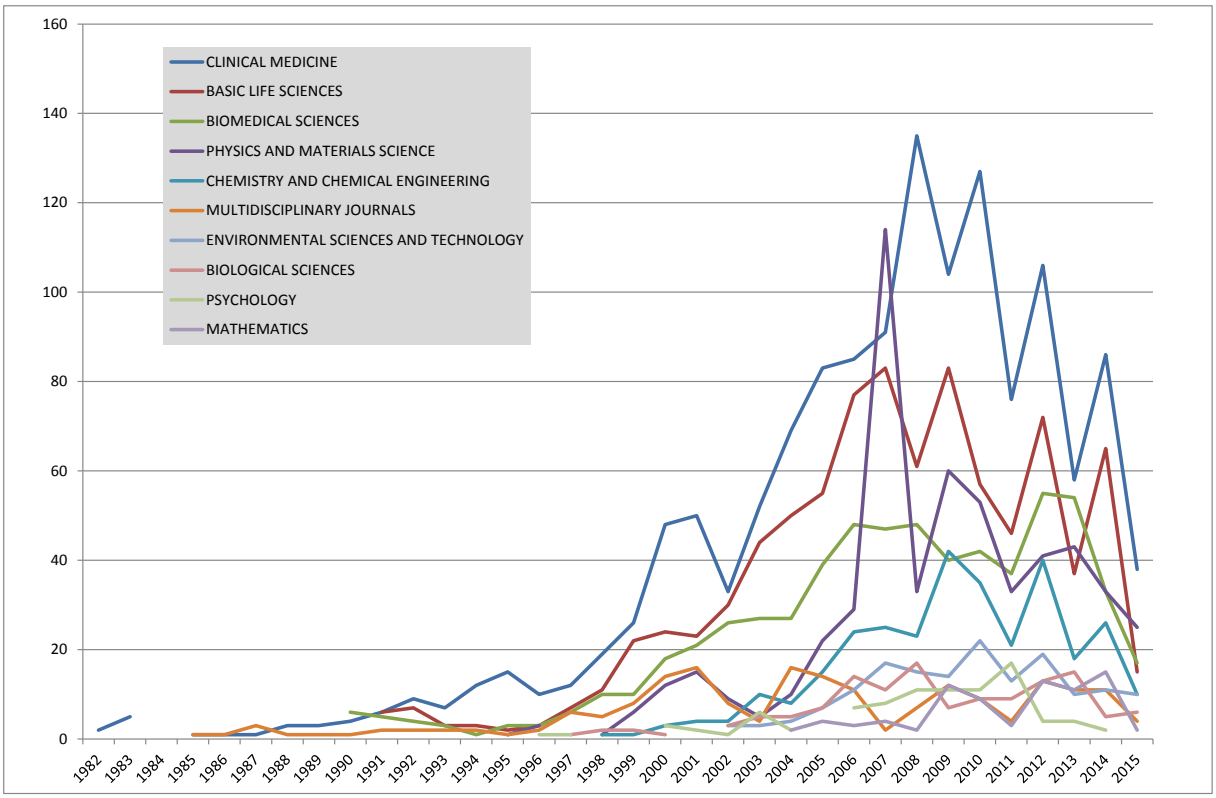

Figure 6: Distribution of number of retracted publications over disciplines in science in WoS covered journals, 1981-2015

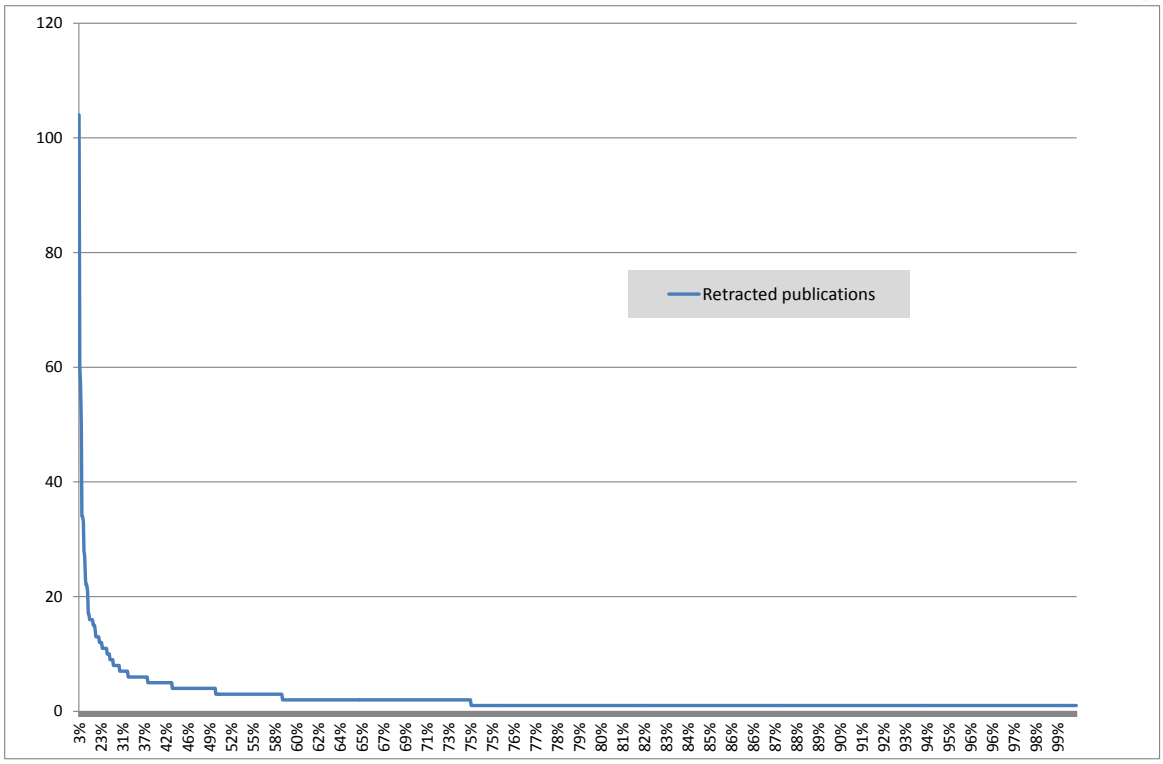

Figure 7: Distribution of retracted publications over journals, 1981-2015

Figure 7 displays a skewed distribution to the left, indicating that most retracted publications appeared in a limited set of journals. This is further illustrated by Table 3, in which the top 15 journals cover 549 retractions, which is roughly 16% of all retracted publications in the study.

	Retracted publications	**Share % of total**	**Cumulative % of total**
ACTA CRYSTALLOGRAPHICA SECTION E-CRYSTALLOGRAPHIC COMMUNICATIONS	104	3%	3%
PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA	60	2%	5%
SCIENCE	57	2%	7%
NATURE	50	1%	8%
ANESTHESIA AND ANALGESIA	34	1%	9%
JOURNAL OF BIOLOGICAL CHEMISTRY	34	1%	10%
JOURNAL OF IMMUNOLOGY	33	1%	11%
MOLECULAR BIOLOGY REPORTS	28	1%	12%
BLOOD	27	1%	13%
BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS	23	1%	13%
CANADIAN JOURNAL OF ANESTHESIA-JOURNAL	22	1%	14%

CANADIEN D ANESTHESIE			
CELL	22	1%	15%
TUMOR BIOLOGY	21	1%	15%
EUROPEAN JOURNAL OF MEDICAL RESEARCH	17	1%	16%
MOLECULAR AND CELLULAR BIOLOGY	17	1%	16%

Table 3. The top of the distribution of the retracted papers over journals, 1981-2015

Next to the journals shown in Table 3, we mention a number of journals that are known for having high Journal Impact Factor (JIF) values, all containing more than 10 retracted publications, such as PLOS One, Journal of the American Chemical Society (JACS), the New England Journal of Medicine (NEJM), FASEB Journal, Applied Physics Letters, Physical Review B, and EMBO Journal. Next to the ones shown in Table 3, the occurrence of journals with high JIF values in the top of the distribution is quite remarkable, which raises questions about the eagerness to publish quickly, and the quality level of the peer review process applied in those journals.

Next, we focus on the classification of the parties responsible for the retraction, that is, the initiators for retraction. Table 4 gives the different parties and the corresponding percentages of retracted papers. A retracted paper has been assigned to combinations of types. We have kept the order in which the initiating parties in the retraction process in the notifications have been mentioned. We printed the most occurring parties involved in retraction in bold print. This means that combinations of the same parties might occur, for example editor(s)/author(s), or author(s)/editor(s). With 27%, 'Author(s)' are the dominant unique retracting party. Next to that, we also notice combinations of authors with other retracting parties. Another 8% of all retractions show combinations of authors with editor(s), and/or publishers. However, within this group it has to be noticed that within 1% of the retractions, one of the authors disagreed ('Author(s) but not all'), and a few papers were retracted by one single author. The editors, and in nearly half of the cases together with the publisher, retracted 35% of the papers. Finally, it is important to remark that in 2% of the retractions, it is unclear from the notifications who initiated the retraction, while in 25% of all retractions, it is simply unknown who retracted, as information was not findable on the reason for retraction.

	Number of retractions	**% of retractions**
author(s)	988	27%
author(s) / editor(s) / publisher	157	4%
author(s) / editor(s)	144	4%
author(s) / editor(s) / publisher / society	12	0%
author(s) / editor(s) / institute	9	0%
author(s) / editor(s) / society	6	0%
author(s) / institute	4	0%

author(s) / Ethical Review Board	2	0%
author(s) / funding agency	1	0%
author(s) / society	1	0%
author(s) but not all	48	1%
author(s) but not all / editor(s) / publisher	7	0%
author(s) but not all / institute	2	0%
editor(s)	671	18%
editor(s) / publisher	559	15%
Editor(s) / author(s)	15	0%
editor(s) / institute	7	0%
editor(s) / publisher / institute	6	0%
editor(s) / publisher / society	4	0%
editor(s) / publisher / COPE	1	0%
Ethical Review Board	1	0%
funding institute	3	0%
institute	22	1%
institute / editor(s) / publisher	6	0%
publisher	5	0%
society	2	0%
unclear from notificiation	82	2%
unknown	950	25%

Table 4. The distribution of the retracted papers over the retracting parties, 1981-2015

For the seven most occurring initiating parties of retraction we plotted the development over time. This is shown in Figure 8. Again, we immediately notice the sharp increase in all initiating parties shortly after the Millennium Change, similar as we observed for other entries on our set of retracted publications. In the previous decade, many retractions are done while the initiating parties are unknown, as can be understood from the various peaks in the period 2001-2012. From 2001 onwards, we also notice a shift in initiating parties, with editor(s) and/or publishers taking a more prominent role in the process. Note that at the end of the period of analysis the numbers decrease, which is not an indication of the problem disappearing, but that in the most recent years, publications with issues have not yet been discovered.

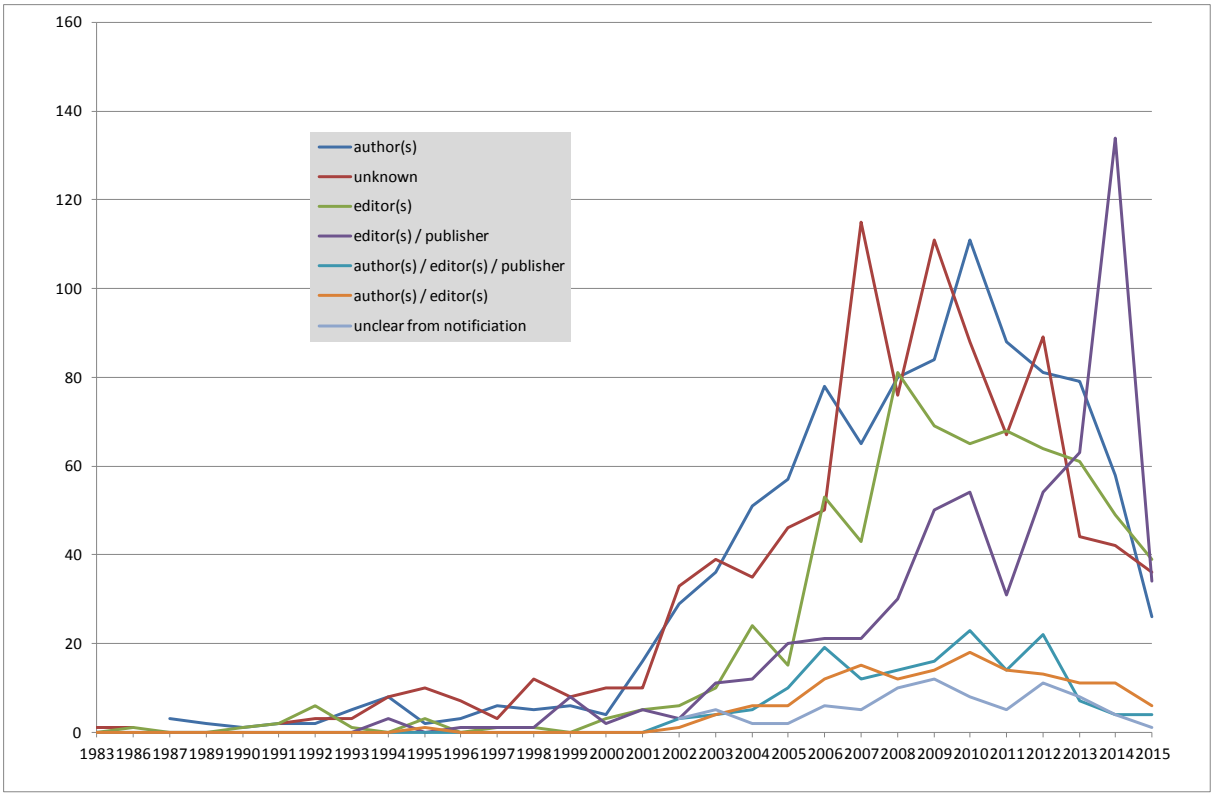

Figure 8: Distribution of initiating parties of retracted publications, 1981-2015

A next step in our analysis is the focus on the reason for retraction. Again, as we did in the analysis of initiating parties, we find in Table 5 that in most cases, one single reason for retraction is given, yet we also find cases in which a combination of reasons is given. It is important to note here that the distinction between various forms of reasons, in particular the distinction between FFP and QRP, are sometimes blurred by the occurrence of reasons for retraction that belong to either of these two large distinctive categories. This occurs in small numbers, but it is still an issue to keep in mind.

	Retracted publications	% retracted publications
authorship issues	108	3%
authorship issues / data incorrect	1	0%
authorship issues / duplicate publishing	1	0%
authorship issues / editorial issues	2	0%
authorship issues / errors / interpretational issues	1	0%
authorship issues / ethical issues	1	0%
authorship issues / interpretational issues	3	0%
authorship issues / plagiarism	1	0%
copyright issues	18	0%
data errors	16	0%
data fabrication	122	3%
data fabrication / data falsification	3	0%

data falsification	258	7%
data falsification / plagiarism	2	0%
data fraud	46	1%
data inaccuracies-inconsistencies-irregularities	110	3%
data integrity	79	2%
data integrity / errors	1	0%
data integrity / ethical issues	1	0%
data integrity / interpretational issues	2	0%
data irreproducible	168	5%
data irreproducible / interpretational issues	2	0%
data unreliability	50	1%
data unreliability / authorship issues	2	0%
data unreliability / ethical issues	1	0%
data validity issues	12	0%
duplicate publishing	386	10%
duplicate publishing / data irregularities	1	0%
duplicate publishing / data issues	2	0%
duplicate publishing / data falsification	2	0%
duplicate publishing / editorial issues	7	0%
duplicate publishing / errors	1	0%
editorial issues	64	2%
editorial issues / authorship issues	2	0%
English & Chinese ?	4	0%
English & German ?	1	0%
Errors	391	10%
errors / duplicate publishing	1	0%
errors / editorial issues	4	0%
errors / interpretational issues	2	0%
errors / plagiarism	1	0%
errors / sloppyness	2	0%
ethical issues	110	3%
ethical issues / data inaccuracies	1	0%
ethical issues / errors	4	0%
ethical issues / study design	1	0%
ghost authorship / incorrect data / data duplication	2	0%
inaccurate methodology	3	0%
incomplete publishing	5	0%
interpretational issues	76	2%
interpretational issues / data falsification	*1*	*0%*
interpretational issues / irreproducible data	*1*	*0%*
invalid results	*8*	*0%*
method inconsistencies	*2*	*0%*
methodological flaws	*1*	*0%*

misconduct	*2*	*0%*
peer review issues	**131**	**4%**
peer review issues / data fabrication	**48**	**1%**
peer review issues / editorial issues	*1*	*0%*
peer review issues / plagiarism	*1*	*0%*
plagiarism	**295**	**8%**
plagiarism / authorship issues	*1*	*0%*
plagiarism / data fabrication	*1*	*0%*
plagiarism / data inaccuracies	*1*	*0%*
plagiarism / data falsification	*1*	*0%*
plagiarism / duplicate publishing	*1*	*0%*
plagiarism / errors	*1*	*0%*
plagiarism / referencing issues	*8*	*0%*
plagiarism / self plagiarism	*11*	*0%*
referencing issues	*17*	*0%*
referencing issues / plagiarism	*1*	*0%*
research integrity issues	**28**	**1%**
self plagiarism	**142**	**4%**
self plagiarism / authorship issues	*1*	*0%*
self plagiarism / duplicate publishing	*4*	*0%*
self plagiarism /data falsification	*11*	*0%*
statistical errorrs	*6*	*0%*
text inaccuracies	*2*	*0%*
unknown	**527**	**14%**
unknown (no full text available / no license agreement with publisher)	**98**	**3%**
unknown (no full text available)	**154**	**4%**
unknown (not found on journal website)	**108**	**3%**
unknown (not/no longer available at publisher)	**29**	**1%**

Table 5. The distribution of the retracted papers over the reasons for retraction, 1981-2015

In Figure 9 we present the distribution of reasons for retraction, in particular the ten most prominent reasons for retraction.

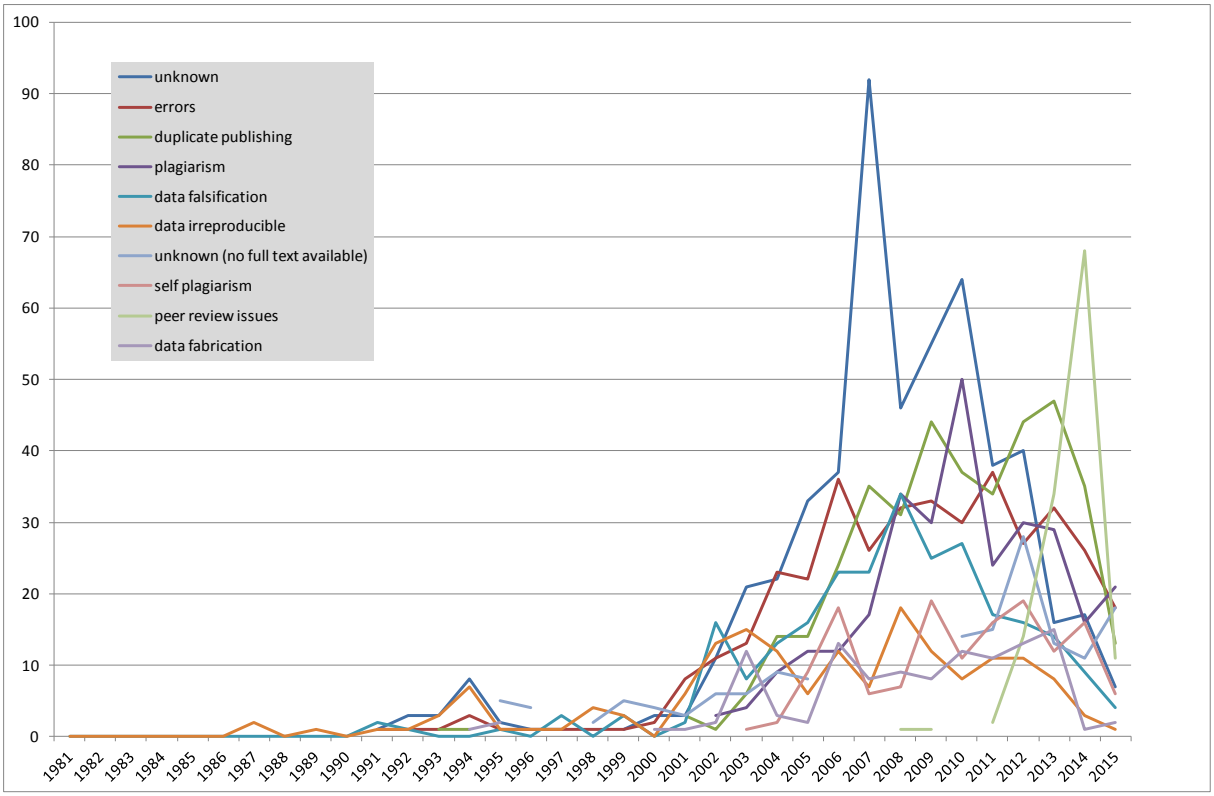

Figure 9: Distribution of reasons for retraction of retracted publications, 1981-2015

Except for the two extreme peaks in the category 'Unknown' in 2007 and 2010, most other reasons for retraction pop up occasionally, with 'Plagiarism' being a main reason for retraction in 2010, and 'Peer review issues' in 2014. The category 'Duplicate publishing' is increasingly visible among the reasons for retraction, which reaches its peak in 2013. Finally, within all the reason, the category 'Errors' is given as reason for retraction in about 30 publications per year form 2005 onwards.

A remarkable similarity is visible in various figures. The peak in retracted material in the geographical break-down, the summing up of Clinical Medicine and Basic Life Sciences, and the peak in retracted material initiated by editor(s)/publisher points at a large collection of publications from PR China, in those two domains, which get retracted by the editor(s)/publisher of the journal(s) in which these Chinese scientists published their work. If we link that to Figure 9, we notice that that combined peaks in the three analyses mentioned before indicate peer review issues (which was in most cases lacking the proper references, which in itself suggests plagiarism issues, but also suspicious behaviour by the referees) as the main reason for retraction of that work.

So we have seen that for most retractions we can identify the reason, although for some 25% we are empty . However as already remarked in other studies (Fang, Steen & Casadevall, 2012), the

retraction notices can be uninformative opaque, hiding the underlying arguments, and secondary sources are often necessary to clarify the real reasons. This is illustrated by the notice published to retract a number of papers published in Science in the well-known 'Schön case' (Reich, 2009):

> *'We are writing as co-authors on the following manuscripts published in Science, which were, in part, the subject of an independent investigation conducted at the behest of Bell Laboratories, Lucent Technologies. The independent committee reviewed concerns related to the validity of data associated with the device measurements described in the papers.*
>
> *….*
>
> *As a result of the committee's findings, we feel obligated to the scientific community to issue a retraction of the above articles. We note that although these papers may contain some legitimate ideas and contributions, we think it best to make a complete retraction.'*
>
> (Boa et al, 2002)

In Table 6a, 6b, and 6c, we combine the reasons for retraction (we selected the 18 most prominent reasons) with the initiators of retraction (here we selected the 8 most prominent categories). Table 5a gives the absolute numbers observed, while Table 5b presents the relative share as compared to the column totals, and table 5c presents the relative shares as compared to the row totals.

When we consider the relatives shares as presented in table 5b (relative shares are based upon the horizontal perspective, calculated over the rows), we note that Authors are the ones most often reporting errors, usually on Data falsification and Data irreproducibility. Editors tend to be most relevant when reporting on Duplicate publishing, Plagiarism, and Data falsification, while editors in combination with publishers alone report most on Peer review issues (21%) as well as Duplicate publishing (13%), and Plagiarism (9%). Editors, in combination with the Publisher and the Author(s) report most on Duplicate publishing (22%), Data fabrication (15%), and Data falsification (9%). Editor(s) in combination with author(s) report on four issues nearly equally, namely Duplicate publishing (15%), Errors (14%), Plagiarism (13%, and Authorship issues (13%). When not all authors retract, the most mentioned reasons are the Irreproducibility of data (17%), Authorship issues (15%), Data falsification (13%), and Errors (10%). When the initiators are unclear, most common reasons for retraction are Duplicate publishing (18%), Plagiarism (13%) and Self-plagiarism (12%).

When we take another perspective, namely through the total number of reasons of retraction (that is, we calculate the relative shares over the columns), in the case of Errors 70% of all retractions are mentioned by Authors. In the case of Duplicate publishing, it is mainly due to the editors of journals that initiate retraction: in 70% of the cases, editors play a role in this reason for retraction. In the case of Plagiarism, again editors in various combinations take the initiative for retraction (in total 74% of the cases). In the case of Data falsification, we observe some sort of a balance between authors on the one hand, and editors in various combinations on the other hand retracting from the literature (38% versus 46% of the retractions). In the case of data

irreproducibility, nearly 80% of all cases are initiated by the authors, in some sort of combination (see Table 3 for the various combinations). In the case of self-plagiarism, it is again mostly the editors in some sort of combination that initiate retraction: in total 76% of the cases, editors are involved. Whenever Peer review issues are at stake, editors are the dominant factor in retracting, as 96% of all cases editors are responsible for retraction. As it could be expected, in the case of data fabrication, authors are not very much involved (in 16% of the cases), it is again the editors in various combinations that play a dominant role here (in over 70% of the cases). Whenever data inaccuracies, inconsistencies, or irregularities are at stake, authors take their responsibilities and retract in 62% of the cases. In Ethical issues, it is again the editor in its various combinations that play an active role in retraction from the literature (66% of the cases). Authorship issues are also for authors an important reason to initiate for a retraction (43%), while editors in various combination take a more or less equal part of all retractions due to authorship issues. In the cases of Data integrity and Interpretational issues, it is mostly authors that take in over 50% of the cases the initiative to retract. Finally, in the case of editorial issues, it is in 69% of the cases the editors that initiate retraction.