Retractions in *Science*

Society for Philosophy of Science in Practice

Gent, Belgium

June 30-July 2, 2018

K. Brad Wray and Line Edslev Andersen

Centre for Science Studies, Aarhus University

Aarhus C, DK8000, Denmark

kbwray@css.au.dk

lea@css.au.dk

Publication plays a pivotal role in the growth and dissemination of scientific knowledge. But the growth of knowledge is neither strictly linear nor unidirectional. Mistakes are made. Retraction is one means by which the scientific record is corrected. In this paper, we examine the retraction practices and prevalence in the journal *Science*. We focus on 35 years of published retractions, from 1983 to 2017. We are not only concerned with determining the scope of the problem, but also the patterns in the data. From a policy perspective, knowledge of any patterns in retractions may be useful in developing targeted responses to deal with the root causes.

Some Background

Retractions are complex, and they can be seen from two very different perspectives, one negative, and one positive. On the one hand, retractions seem to indicate that there is something significantly wrong somewhere in the publication process in science. A retraction may signal a problem with the peer review process, for example, indicating a failure to weed out improperly conducted research. Alternatively, it could indicate a problem in the contexts where research is being conducted, in the laboratories or other sites of research. Perhaps the problem arises even earlier, in the context where the norms of research are being taught.[1] Regardless what the cause of the problem is, in the United States both the popular press and politicians have appealed to retractions as a way to undermine confidence in scientific claims (see Hilgard and Jamieson 2017).

On the other hand, a number of scholars have noted how rare retractions are. By one count, there were "'some 300 retractions among 1.4 million papers published annually'" (Marcus and Oransky 2011, in Hilgrad and Jamieson 2017). Thus, it appears that less than one in 4,500 published papers are retracted. Given that so few papers are retracted, there may be no problem at all. In fact, the low levels of retractions might speak to the good health of the peer review system, the laboratory cultures, and the training of young scientists (see Hilgard and Jamieson 2017).[2]

Perhaps more noteworthy than how infrequently retractions are issued, is the increase in frequency of retractions over time (see van Noorden 2011, 27 and Wager and Williams 2011, 568; but see also Fanelli 2013). And there are concerns that the present culture of scientific publication is apt to make matters worse with respect to retractions. Again, though, an increase in

the frequency of retractions need not necessarily signal that a particular type of problem is becoming more prevalent. It may merely signal a shift in editorial policies, as journal editors try to improve an already well-functioning system, or respond to public perceptions of a problem in science (see, for example, Fanelli 2013).

But some have raised the concern that retractions are an ineffective means to correct the scientific record, as retracted papers continue to be cited for some time afterwards (see Budd et al. 1998; and Marcus and Oransky 2017, 120). Others suggest that the problems created by retractions and papers in need of retraction are quite small compared to the problems related to the crisis surrounding reproducibility (see Marcus and Oransky 2017, 123-124). Our study aims to enrich our understanding of retraction in science. We suspect that a study focussed narrowly on one of the two key general science journals may shed some new light on the issue.

Data and Methods

We report on data on:

the **number of scientists who authored the retracted articles**;

the **number of scientists who signed the retraction**;

how often (i) **all authors**, (ii) **only some authors**, and (iii) the editors of *Science* retracted the article;

the **time between publication and retraction**; and

whether the articles continued to be cited after they were retracted.

Our aim is to identify patterns in the retractions, thus revealing a structure to this aspect of the growth of scientific knowledge.

This study will aid us in understanding the publication process, and the role that corrections play in the growth of science. Retraction is just one form of correction in the scientific literature, but a deeper and systematic understanding of it will provide a fuller picture of the role of publication in advancing scientific knowledge.

Before presenting our findings, it is worth noting that ***retracted articles*** published in *Science* outnumber ***retraction notices*** published in *Science*. This is because a single retraction notice will sometimes announce the retraction of more than one published article. In fact, in our survey of retractions in *Science*, there were four notices retracting two articles, two notices retracting three articles, and one retraction notice retracting eight articles. When we report the data on retractions we are ultimately counting the retracted articles, not the retraction notices. But we date the retraction to the date of the retraction notice.[3]

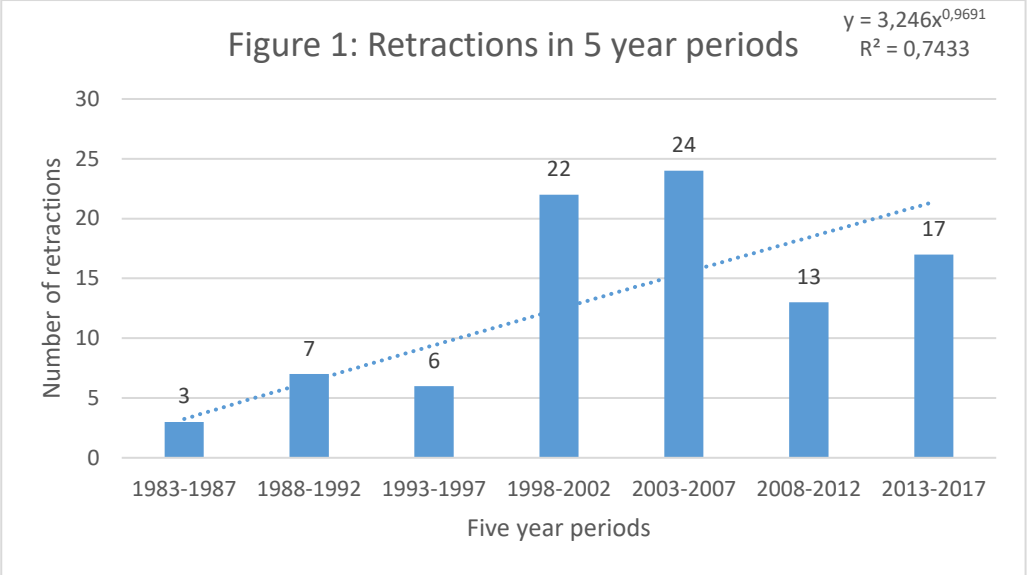
Findings

We collected data on 35 years of publications in the journal *Science*, from 1983 to 2017.[4] In that 35 year period 92 articles were retracted. So on average about 2.6 articles were retracted each year. In some years there were no retractions, whereas in other years there were a significant number. The five years in which no retractions were made are 1984, 1987, 1988, 1993, and 1996. Since 1996, there has not been a year in which no retractions were made. There is a measurable increase in the number of retractions as we approach the present, from 3 retractions in the period

of 1983-1987, to 17 retractions in the period of 2013-2017, though the final period was not the period in which the most retractions occurred. See **Figure 1**.

It is worth noting that the great increase in the number of retractions from the five year period 1993-1997 to the five year period 1998-2002 is due, to a large extent, to eight retractions in 2002 involving a single scientist, Jan Hendrik Schön. Schön did not publish alone, but this is a case in which he alone was singled out as responsible for misconduct in all eight of the retractions (see Kennedy 2002, 495). Schön, incidentally, was also responsible for two additional retractions in the next five year period, 2003-2007.

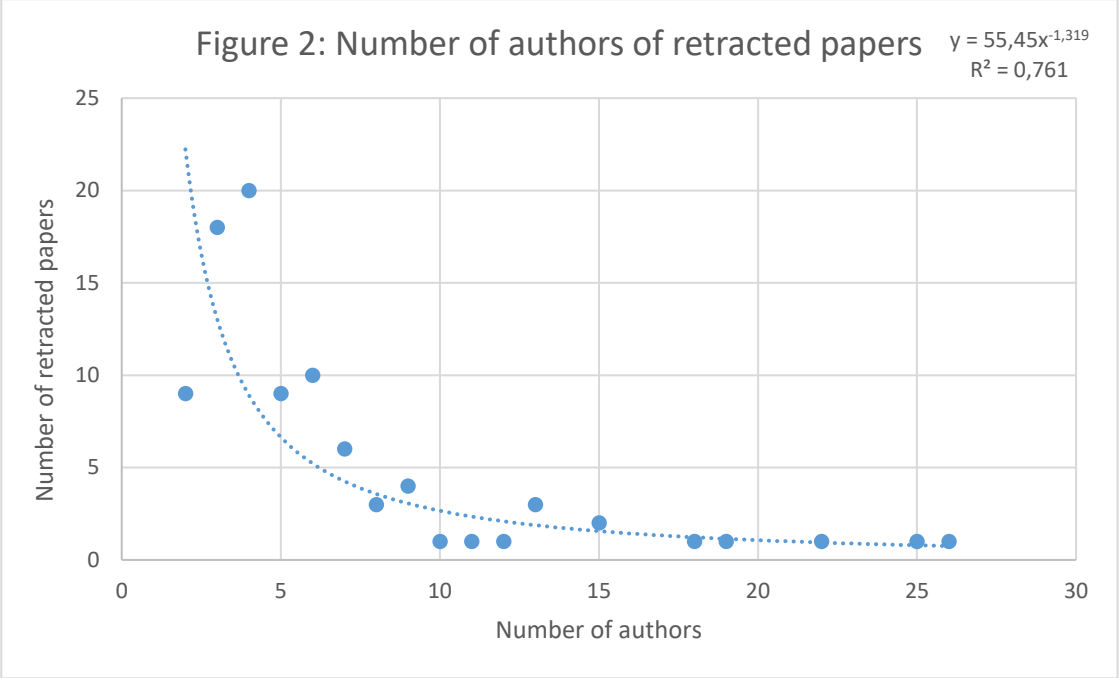
It is worth adding some context to the numbers of retractions. *Science* publishes 51 issues a year. We counted how many Research Articles and Reports were published in 2010, assuming that 2010 was a typical year. There were a total of 762 Research Articles and Reports published in that year. So, given that on average 2.6 articles are retracted each year, approximately 0.34 % of the articles published in *Science* are retracted (or, about 1 in 300).[5] This figure is higher than the figure reported above, by Hilgard and Jamieson (2017). It may reflect the fact that scientists who publish in *Science* are under greater pressure to publish in the most prestigious venues, and as a consequence may be more inclined to compromise their research. Alternatively, it may suggest that other journals are under less scrutiny than *Science*, and articles that should be retracted, and would be retracted if they were published in *Science*, are going undetected. But our finding is in keeping with a previous study which found that there was a correlation between the impact factor of a journal and the rate of retracted articles. That is, journals with higher impact factors have higher rates of retractions (see Fang and Casadevall 2011, 3856, Fig. 1; see also Cokol et al. 2007, 423). But, as Daniele Fanelli notes, "journals with a high impact factor are more likely to have clear policies for scientific misconduct" (see Fanelli 2013, 6).



It is also worth noting how many authors were responsible for the retracted papers, that is, how many scientists authored the original papers that needed to be retracted. There was not one single-authored paper retracted in *Science* during the 35 period we studied. This may seem to suggest that retractions are a consequence of co-authorship and collaboration. But such an inference is unwarranted for the following reason. Given our count of the articles published in *Science* in 2010, we know that very few Research Articles and Reports in *Science* are published by a single author these days. In fact, only 8 of the 762 Research Articles and Reports published in 2010 were single authored (that is about 1 %).

A little over half of the retracted papers were authored by collaborative teams of two to four scientists (47 out of 92 papers; 51 %). The remaining 49 % of the retracted articles were authored by teams ranging from 5 scientists to 26 scientists. See **Figure 2**. The best fit curve is described by a power law, where $y = 55.45x^{-1.319}$ and $R^2 = 0.761$. This distribution is not surprising,

given that earlier research by Derek de Solla Price suggests that there are more publications by smaller research teams than larger research teams in science (see Price 1963, page 88, Figure 19).



Interestingly, although collaborative teams of two to four scientists are responsible for 51 % of the retracted articles, such groups authored only about 29 % of the articles published in *Science*, assuming 2010 is a typical year. See **Figures 3a and 3b**. Thus, these smaller groups are more prone than either lone authors or larger research teams to publish articles that are ultimately retracted. Collaborative teams of five to seven scientists are responsible for roughly their share of the retracted articles. They were responsible for 30 % of the publications, and 27 % of the retracted articles. Larger research teams, teams of eight or more scientists, were responsible for less than their share of the retractions. Though they were responsible for 40 % of the publications in *Science*, they were only responsible for 22 % of the retracted articles.[6]

Future research should explore whether in fact larger teams are more careful than research teams of two to four scientists, or whether some other factors explains this pattern in the data.

Figure 3a: Percentage of Publications, by Size of Authorship Team

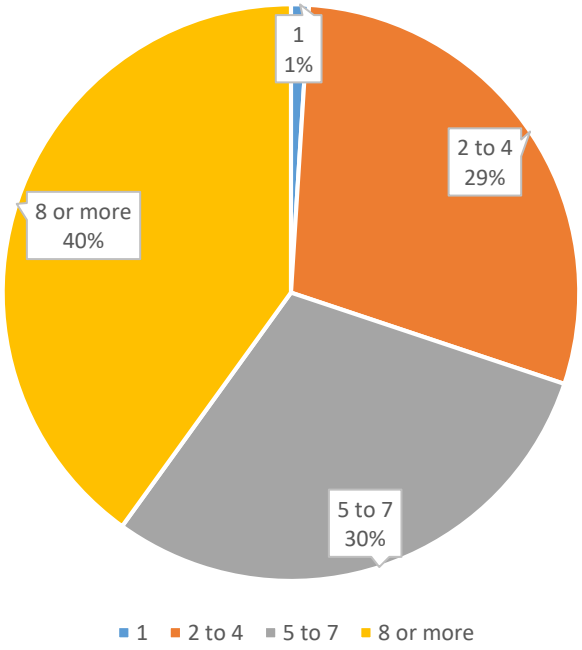
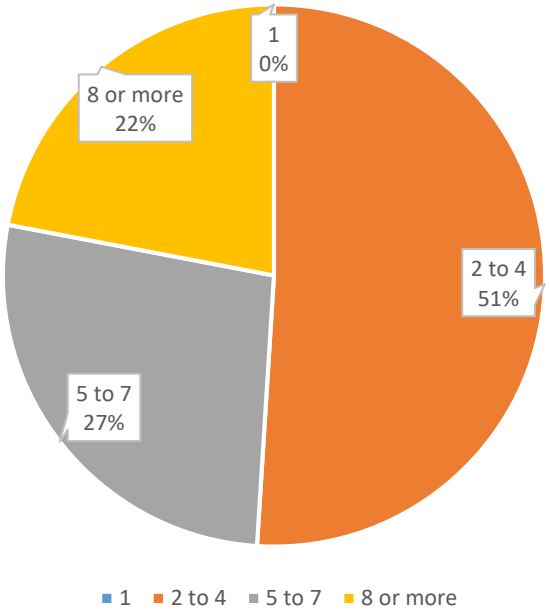


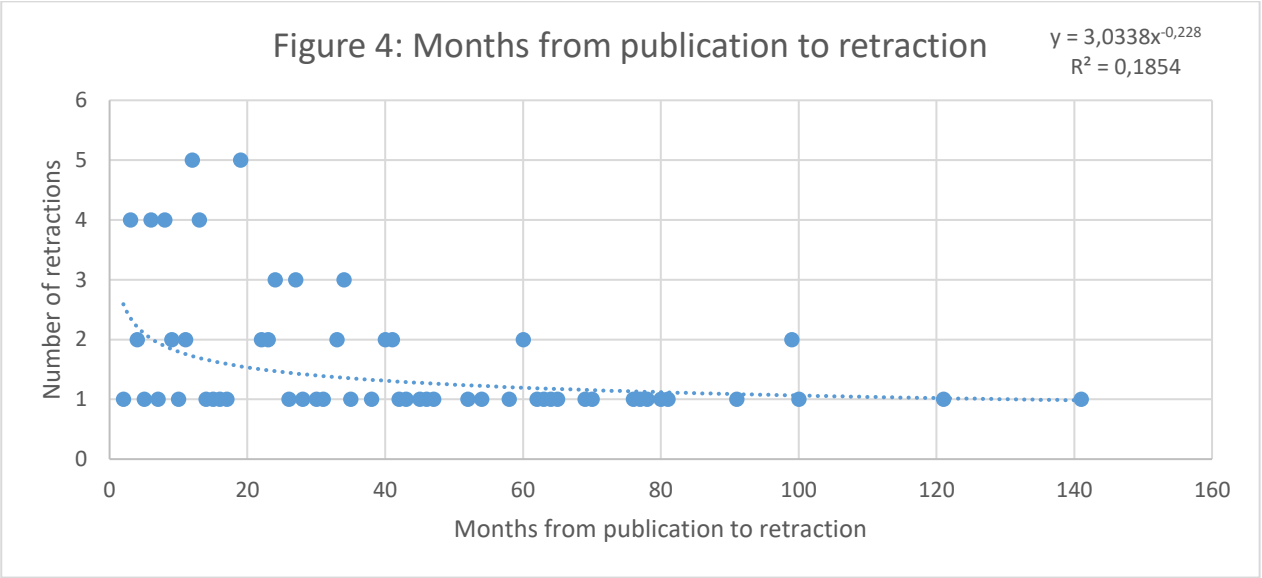
Figure 3b: Percentage of Retractions, by Size of Authorship Team



It is interesting to note the time between the publication of the article and the retraction. The time between publication and retraction will provide some indication of the efficiency of the scientific publication system in responding to problems. We measured the time between publication and retraction in terms of months. In our data, the time between publication and retraction ranged from two months to 141 months, that is, from 2 months to almost 12 years. The median time between the publication of the original article and the publication of the retraction was 24 months, that is, two years. The mean time between publication and retraction is 34 months (3113/92; about 2 years and 10 months). Almost 30 % of the retracted papers were retracted within a year of publication (27/92; 29 %). And 76 % of the retracted papers were retracted within four years of publication (70/92). The remaining 24 % were retracted between 4 years and 12 years of their publication (22/92). See **Figure 4**. The best fit curve is described by a power law, where $y = 3.0338x^{-0.228}$ and $R^2 = 0.1854$. Our figures on the time between publication and retraction are similar to those reported in an earlier study by Fang, Steen and Casadevall (2012, 17030).

The fact that many retractions occur within 12 months of publication speaks to the efficiency of the scientific publication system, at least as it operates in the journal *Science*. A retraction cannot occur until a publication has been scrutinized, and a significant error or concern found with the publication. Then the retraction has to be written, sent to the journal, and vetted by editors or referees. Perhaps it also speaks to the integrity of the system that even papers that were published five years ago (that is, 60 months ago) are retracted. It suggests that scientists do

not take the scientific record as settled after something has been published, not even five years after something is published.



It is also worth noting how many of the authors of the retracted articles sign the retraction. In 15 % of the cases (14/92) none of the authors of the retracted article signed the retraction. These are called "editorial retractions," as it is the editorial team of the journal that issues the retraction.[7] Most often the articles that are retracted by an editorial retraction involve misconduct. Some of the editorial retractions involve honest mistakes but not all the authors agree to the retraction. It is worth noting that in an earlier study of retraction in biomedicine, Budd et al. (2011) found that 29 % of retractions were issued by editors, and an additional 7 % were issued by a combination of authors, editors, and publishers.

Perhaps a more important figure to note is that in 60 % of the retractions all of the authors of the retracted paper signed the published retraction (55/92). In many respects this is a

comforting figure. These findings are similar to those reported in an earlier study, by Wager and Williams (2011, 568).

It is worth noting that in 2002 *Science* had "a standing policy that ***all authors of a paper must agree to its retraction***" (see Kennedy 2002, 495; emphasis added). But policies change. In 2009, "the editors of *Science* informed [a scientist involved in a case where a retraction seemed appropriate] that the journal's editorial practice requires that they ***get signatures directly from all authors wishing to retract a paper***" (see Service 2009, 1611; emphasis added). This is a subtle change, so it is worth highlighting it. The earlier policy was that ALL authors of the retracted paper agree to its retraction. The more recent policy is that ALL AUTHORS WISHING TO RETRACT the paper sign the retraction. Thus, it is important to remember that the publication system in science is dynamic, responding to challenges in different ways at different times, as lessons are learned from past experiences.

In some cases people other than the authors of the paper sign the retraction notice together with all or some authors of the paper. Sometimes these other people assisted in re-evaluating the original data or attempting to replicate the study. At other times, when a retraction notice retracts more than one article, the various retracted articles will not have all the same authors.[8] These cases deserve further study.

We found that almost all of the papers continued to be cited after they were retracted in part or in full, if we disregard the papers that were retracted in 2016 or 2017. These latter papers have had little time to be cited since being retracted. Another paper, from 2010, received no citations at all, neither before nor after it was retracted. This paper was retracted 16 months after it had been published.

The remaining retracted papers got between 1 and 936 post-retraction citations. In an earlier study, conducted before journals became widely available on-line, Budd et al. (1998) found that retracted articles were cited for their reported findings even after they were retracted. There is no evidence that on-line publishing is affecting a change here.

Whether these finding about post-retraction citations are a disconcerting sign depends on the nature of the post-retraction citations. Provided they explicitly acknowledge the problems with the original paper, then there is no problem. Whether the post-retraction citations are a disconcerting sign *also* depends on the nature of the retractions. When a paper is retracted in part or in full, parts of the paper may remain valid. Further study is required to determine the nature of the post-retraction citations. Only then would we be in a position to determine how disconcerting post-retraction citations are. In a subsequent study we hope to analyze the nature of post-retraction citations.

Discussion

We have only reported on the most straightforward descriptive statistics here in our study of retraction in *Science*. We have a wealth of additional data that will support additional studies in the near future. For example, we would like to examine the implications that our retraction data have for understanding the notion of collective authorship in science. We would also like to conduct a detailed qualitative study of the causes of retraction. A few qualitative remarks are in order.

We believe that the notion of retraction can give the false impression that this is a categorical concept. The varieties of retractions may be better conceived of as laying on a

continuum. A full retraction signed by all the authors of the original paper, for example, is quite different from a retraction of some claims in a paper, which is in turn different from a retraction of a particular analysis or interpretation. Second, at the one end of the retraction continuum, a retraction can differ little from a correction. Clearly, there is a significant difference between a full retraction and a correction. But we want to insist on the importance of recognizing more distinctions.

It is perhaps natural to think of retractions as a sign of the toxic effects of the pressure to publish. But there are other considerations that need to be acknowledged. Some retractions are made after a research team has conducted further research. Again, one's initial impulse may be that these retractions could and should have been avoided by not rushing to publish in the first place. We believe that the situation may be more complex than this explanation suggests. It is likely that some important research that is published in articles that never need to be retracted may never have been conducted in the first place if some research that was published in a retracted paper was not published first. Not all mistakes in science can be anticipated or identified in the short term, before they are made. This is a normal part of learning from our mistakes.

References

Berg, J. 2017. "Editorial Expression of Concern," *Science*, 357: 6357 (22 September 2017), 1248.

Budd, J. M., M. Sievert, and T. R. Schultz. 1998. "Phenomena of Retraction: Reasons for Retraction and Citations to the Publications," *Journal of the American Medical Association*, 280: 3, 296-297.

Budd, J. M., Z. C. Coble, and K. M. Anderson. 2011. "Retracted Publications in Biomedicine: Cause for Concern," *Association of College and Research Libraries National Conference Proceedings*, 390-395.

Cokol, M., I. Iossifor, R. Rodriguez-Esteban, and A. Rzhetsky. 2007. "How Many Scientific Papers Should be Retracted?," *EMBO Reports*, 8: 5, 422-423.

Culliton, B. J. 1983. "Coping with Fraud: The Darsee Case," *Science*, 220: 4592 (April 1, 1983), 31-35.

Fanelli, D. 2013. "Why Growing Retractions Are (Mostly) a Good Sign," *PLOS Medicine*, 10: 12, 1-6. (e1001563)

Fang, F. C., and A. Casadevall. 2011. "Editorial: Retracted Science and the Retraction Index," *Infection and Immunity*, 27: 10, 3855-3859.

Fang, F. C., R. G. Steen, and A. Casadevall. 2012. "Misconduct Accounts for the Majority of Retracted Scientific Publications," *Proceedings of the National Academy of Sciences*, 109: 42, 17028-17033.

Hilgard, J., and K. H. Jamieson. 2017. "Science as 'Broken' Versus Science as 'Self-Correcting': How Retractions and Peer-Review Problems Are Exploited to Attack Science," in K. H. Jamieson, D.

Kahan, and D. A. Scheufele, (eds.), *The Oxford Handbook of the Science of Science Communication*. Oxford: Oxford University Press, pages 85-92.

Hull, D. L. 2001. "Why Scientists Behave Scientifically," in D. L. Hull, (ed.), *Science and Selection: Essays on Biological Evolution and the Philosophy of Science*. Cambridge: Cambridge University Press, pages 135-138.

Kennedy, D. 2002. "Next Steps in the Schön Affair," *Science*, 298 (18 October 2002), 495.

Marcus, A., and I. Oransky. 2017. "Is There a Retraction Problem? And, If So, What Can We Do About It?," in K. H. Jamieson, D. Kahan, and D. A. Scheufele, (eds.), *The Oxford Handbook of the Science of Science Communication*. Oxford: Oxford University Press, pages 119-1126.

Price, D. de Solla. 1963. *Little Science, Big Science*. New York: Columbia University Press.

Service, R. F. 2009. "A Dark Tale Behind Two Retractions," *Science*, 326 (18 December 2009), 1610-1611.

Van Noorden, R. 2011. "The Trouble with Retractions," *Nature*, 478 (6 October 2011) 26-28.

Wager, E., and P. Williams. 2011. "Why and How do Journals Retract Articles?: An Analysis of Medline Retractions 1988-2008," *Journal of Medical Ethics*, 37: 567-570.

Notes

[1] One high profile case from the USA suggests that the propensity to cheat in scientific research may originate quite early in one's training. In the case of John R. Darsee, a research fellow at Harvard who was caught fabricating data, an inquiry traced his irresponsible behavior back to his undergraduate days at Notre Dame (see Culliton 1983, 34).

[2] David Hull (2001) argues that, compared to politicians, scientists seem to be far less inclined to act dishonestly. Hull attributes this to the fact that the interests of individual scientists line up with the interests of science, the institution.

[3] In one case, a single article was retracted in two notices. In this particular case, two months after the initial retracton a more elaborate retraction was issued. In the data we report below, we only count this retracted paper once, and we focus on the initial retraction. In our database, we refer to a retraction notice by the date of its publication. **[2016 12 16]**, for example, refers to the retraction notice published in *Science* on December 16, 2016.

[4] The most recent retraction notice published in *Science* prior to 1983 was published in 1968.

[5] We have included one retracted piece in our study that is neither a Research Article nor a Report, but rather a Review Article.

[6] Our findings deviate somewhat from findings reported in an earlier study of articles in PubMed. In that study, 9 % of the retracted articles were single authored, and 46 % were authored by teams of two to four scientists. The remaining 45 % were authored by teams of five or more scientists (see Wager and Williams 2011, 568-569). It is noteworthy that PubMed is an archive of the scholarly literature in biomedicine and the life sciences only.

[7] Five of the 14 editorial retractions were not titled as such in *Science*, though it was the editors of *Science* who issued the retractions. Perhaps a lower level of warning is the "Editorial Expression of Concern." An example is Berg (2017).

[8] So, for example, a notice may announce the retraction of two articles, one authored by authors A, B and C, and the other authored by A, B, D, and E. The retraction may then be authored by A, B, C, D and E. C, D, and E are then retracting a paper that they did not co-author together with a paper they did co-author.