

---

*The Integrity of Science II*

---

## Two Cheers for the Retraction Boom

*Ivan Oransky and Adam Marcus*

Between 2000 and 2010, the number of published papers in the sciences rose by 40 percent, from about 1 million per year to about 1.4 million. Over that same period, the number of retracted articles—the ultimate in academic take-backs—grew tenfold, from about 40 per year to about 400. The figure is now somewhere close to 700 papers retracted annually. Although retractions represent a small sliver of the total literature, accounting for roughly 0.05 percent of all articles, the remarkable increase in the retraction rate has been seen by many as a symptom of sickness in the body scientific.

It is tempting to look at the growing rate of retractions as an indicator that scientists increasingly don't know what they are doing or, worse, are becoming less honest about their work. After all, two-thirds of retractions are due to research misconduct, rather than honest error. The kinds of actions that count as misconduct are plagiarism, fabrication of data, and falsification—a category that includes the deliberate manipulation of data or research protocols that leads to the misrepresentation of results. Our work and the scholarship of others, however, suggest a more hopeful view of the rise in retractions: Not only are the vast majority of researchers playing by the rules, but the practice of science itself has never been healthier.

To understand this point, it helps to know why retractions have been on the rise in recent years. The reasons are manifold, but they start with technology. Technology has provided ever-clever fraudsters new ways to pull the wool over the eyes of editors and reviewers. Some cheats use Photoshop to splice together parts of images to misrepresent the results of an experiment. Others take advantage of the fact that many journals ask authors to provide suggestions on who they would like to conduct peer review of their submissions. While this practice may have problems even when it works as designed, it can also allow for outright fraud when authors give journals

---

*Ivan Oransky is vice president and global editorial director of MedPage Today and distinguished writer in residence at New York University's Arthur Carter Journalism Institute. Adam Marcus is managing editor of Gastroenterology & Endoscopy News and teaches science journalism at Johns Hopkins University. They are co-founders of Retraction Watch.*

SPRING/SUMMER 2016 ~ 41

made-up e-mail addresses (that the authors themselves have access to) for either real or invented peer reviewers, and proceed to conduct peer review for their own papers.

But technology has also allowed readers to become better at catching problems. For example, most publishers now use some version of plagiarism-detection software to identify manuscripts with plagiarized text early in the process of submission. While the market for plagiarism-detection tools supports some sophisticated and expensive software such as iThenticate, freeware like DeJaVu—and for that matter, even Google—are pretty effective first-pass methods for catching would-be cheats. With these tools available there is no good reason that in 2016 any journal would publish a paper that plagiarizes a previously published article. Detecting doctored images is not as easy as catching plagiarists, because the programs needed to scan and compare pictures are more data intensive and harder to code for than the programs used to compare text. But computer scientists have developed tools for detecting image manipulation, and some of them are freely available to journal editors.

Technological developments don't just provide tools for either committing or detecting fraud, but also make it easier for scientists to communicate and monitor one another's work. While the Internet has made committing plagiarism and image manipulation easier in many ways, it has also made it easier to detect such misconduct—for the simple reason that more eyes allow for greater scrutiny on publications.

### **What Retractions Mean (and What They Don't)**

This brings us to a question that we are often asked: If readers are finding problems in papers after they are published, why are peer reviewers not catching them beforehand? Does the failure of reviewers to identify misconduct or honest error prior to publication mean that peer review is broken? It certainly means that we are more aware of its flaws now that all scientists can easily be post-publication peer reviewers. But anyone who thought peer review was a *Good Housekeeping* seal of approval, even before the Internet, was sold—perhaps willingly—a bill of goods. Why should we expect that a few experts, who may not really be experts at all in the techniques used in a given study, would be able to spot every error? Post-publication peer review is not just a way to fix a supposedly broken system of pre-publication peer review, but a necessary adjunct to that system.

The effectiveness of post-publication peer review is evident in the higher retraction rates of prestigious journals such as the *New England*

---

*Journal of Medicine, Cell, Nature, and Science.* One explanation for these higher retraction rates is that scientists push the envelope to earn the career-making brass ring that a paper in one of these publications can offer. But high-prestige journals are highly read, and so are subject to higher scrutiny, which may be a more likely explanation for why fraud and error are detected more frequently in their pages.

Some of the scrutiny of scientific papers comes from post-publication review by fellow scientists, and efforts to encourage and facilitate such scientific-community-wide review, like the website PubPeer, where scientists can anonymously comment on published papers, and PubMed Commons, where they must use their real names, have been on the rise in recent years. But there are also older and more institutionalized sources of combating scientific misconduct; for example, in the United States, the creation in 1989 of what has now become the Office of Research Integrity established a formal infrastructure for policing federally funded science. The ORI is far from perfect: It has no real prosecutorial authority, cannot launch inquiries without an invitation from an institution, and is not adequately staffed for the scope of the problem, launching inquiries for roughly 30 to 40 of the 300 to 400 cases that come its way each year. Nonetheless, when the ORI does publish its findings of research misconduct—about a dozen a year (and it only publishes case summaries when it actually discovers misconduct)—this draws the attention of scientists and publishers to a particular researcher’s misdeeds, and so serves as a signal to fraudsters and would-be fraudsters that getting caught carries consequences. But we would argue that the typical recent penalties—a few years of supervision on publicly funded research projects and a temporary inability to receive government grants directly, and so forth—are too light. It seems odd, given today’s funding environment, that these bans don’t last longer. Criminal sanctions may also be appropriate in some cases, and a recent survey suggests that the vast majority of Americans agree.

The work of both concerned scientists and agencies like the ORI have led some observers to argue that the rising rates of retractions are more a sign of improvements in the scientific community’s vigilance than of increasing corruption and vice. For instance, Daniele Fanelli, a researcher who studies scientific misconduct and bias, published an article in 2013 titled “Why Growing Retractions Are (Mostly) a Good Sign.” Consider the field of anesthesiology, which has the dubious distinction of having the top two record holders in retractions: the two scientists with the most retractions on their CVs, Yoshitaka Fujii and Joachim Boldt, are both anesthesiologists, and anesthesiologist Scott Reuben spent six

months in federal prison for health care fraud. Combined, just these three researchers have lost more than 300 papers. But rather than showing the corruption of the field of anesthesiology, this exceptionally high number of retractions reflects a commendable vigor and dedication on the part of a handful of journal editors who root out fraud in their discipline at the price of a few moments of bad press. The long-term benefits—greater trust in the integrity of papers in these journals—have been well worth the brief embarrassment. Indeed, people should be far more suspicious of a discipline whose journals never or rarely retract, even for honest error. A lack of retractions in a scientific field or in a particular journal could well be a sign that only pristine work is being published, which hardly seems likely, or a sign that the fraud is simply not being retracted and perhaps never even detected.

Low retraction rates in a given field or journal are probably a sign of how difficult it is to convince authors and editors to retract, a phenomenon that we have certainly witnessed in our own work running the website *Retraction Watch* since 2010, and that others have commented on as well. This situation may be changing. As Grant Steen, Arturo Casadevall, and Ferric Fang, who have also studied misconduct, note in a 2013 paper, editors appear to be retracting more quickly, and to be reaching back further into their archives to scrutinize all studies by an author found to have committed misconduct. So we may well expect that “the overall rate of retraction may decrease in the future as editors continue to process a glut of articles requiring retraction.”

But just as knowing where speed traps are may encourage some drivers to slow down only to break the limit where they don’t fear scrutiny, the increased attention to retractions could drive some researchers to commit acts of “almost misconduct,” also known as “questionable research practices.” Even in 2005, three researchers who study scientific fraud were warning that science needs to look beyond the traditional “fabrication, falsification, and plagiarism” definition of misconduct.

And that’s why retractions are not a very useful proxy for what is really happening in much of science. Perhaps it is surprising to see the co-founders of a website that focuses almost exclusively on retractions say that. But it has always been clear to us that dividing publications neatly into “retraction-worthy” and “trustworthy” was as misleading as suggesting that the world is composed only of complete villains and absolute heroes.

As Alison Abris, a researcher for our site, wrote in her Ph.D. thesis, “research misconduct occurs at a greater rate than retractions for miscon-

duct are published, and retraction and correction notices cannot be relied upon to convey the presence of fraudulent data within the publication.” They don’t happen to be very good proxies for honest errors that cause serious reliability problems, either. Researchers have published hundreds of papers with what turn out to be contaminated cell lines, and, as Steen, Casadevall, and Fang report, the vast majority of those studies are not even marked with a warning about the problem.

### **Incentives and Virtues**

So what should change? It is our belief that it would be easier to correct the scientific literature if our academic reward systems did not treat the published paper as so much of a sacred object. It is understandable that scientists who know that their futures depend on their publication record would be loath to have that record marred with retractions. We need to replace these incentives with ones that reward open data sharing, post-publication peer review, and similar activities that reflect how we want science to work, encouraging honest efforts both to produce the best results and to correct one’s own mistakes and those of others.

And it is also understandable that some researchers are concerned that retractions collectively are a mark against science that will be used by politicians to cut its funding. But those scientists who now silently curse the wasted effort and resources they expend when they find that they have been trying to build on flawed work might want to raise their voices, to remind us all that self-correction is one of the defining virtues of science.