nature neuroscience

The science of retraction

Recent findings of misconduct by researchers at both Bell Labs and Lawrence Berkeley National Laboratory have once again brought the issue of fraud and scientific trust into the public spotlight. Although damaging to science's public image, cases like this are often cited as evidence that science is self-correcting, that faulty research will eventually be uncovered and if necessary retracted from the published literature. This may be true in the long run, but to conclude from a few high-profile cases that the system always works well is too complacent.

Formal retraction is the appropriate way to make others aware that a published study is not to be trusted. However, only a tiny proportion of published papers are ever retracted. A survey[1] of the Medline database from 1966 to 1997 revealed that of the nine million or so articles indexed during this period, only 235 were retracted—in other words, less than 0.01%. (As of this writing, the numbers have risen to over twelve million articles and 464 retractions.) Of these retractions, some reflected failure to replicate the original results, some reflected scientific misconduct, and others provided no explanation. It seems likely, however, that many retractions are the result of misconduct investigations. For example, over the last ten years the NIH Office of Research Integrity has found 135 cases of scientific misconduct, and has required 81 papers to be retracted or corrected as a result.

Unfortunately, however, retraction alone does not appear sufficient as a way to clean up the scientific literature, because articles often continue to be cited after they have been retracted. For example, the 235 retracted papers in the survey above received more than 2000 citations, of which the great majority presented the conclusions as if they were still valid[1]. Presumably this is because many authors are sloppy about checking their references, and if citations are simply copied from other papers or from reviews, then errors can be propagated almost indefinitely.

It is impossible to know how much of the scientific literature is unreliable, but formal retractions surely represent only the tip of the iceberg; simple statistical considerations suggest that if (say) results are reported with p-values of 0.01, then on average 1% of findings should prove to be artifactual. Whether for statistical or other reasons, every field has papers whose conclusions are widely disbelieved by experts, yet that remain in the literature without any formal notification to the community that their scientific validity has been questioned. This is regrettable but not surprising. Researchers are often hesitant to challenge their peers in writing, for reasons that may include loyalty, fear of retribution, or reluctance to devote time and energy to public disputes. There is also a perception that journals are unreceptive to refutations that might reflect poorly on the journal's editorial reputation. (The policy of *Nature Neuroscience* is to publish refutations of our own papers only in cases where there is compelling evidence that the main conclusion of the original paper was incorrect.) Authors themselves might reasonably be expected to take responsibility for retracting or correcting claims that they themselves no longer believe, but in reality they have little incentive to do so, and unless confronted with clear evidence of misconduct, the line of least resistance is often to ignore the issue until it is forgotten.

How then can the community develop better ways for work to be debated and questioned, and to ensure that this information is accessible to everyone? An obvious first step is to ensure that retracted papers are clearly marked as such. The PubMed interface for the Medline database includes retraction information, but many journal web sites do not. Publishers need to be aware of this problem, and authors need to be conscientious about going back to original sources to confirm their validity.

Proper labeling of retractions, however, does not deal with the larger problem of questionable studies that are never retracted. One idea might be to create a database to which scientists could submit refutations, 'failures to replicate', or other findings that are not easily published in a traditional journal format. Of course, many logistical problems would need to be addressed before such a system could be made to work. For example, to maintain scientific credibility it would need editors and peer reviewers, who would need to commit themselves to the somewhat thankless task of cleaning up other people's errors. Such a database would need to be easily accessible and citable, with appropriate links to and from the rest of the scientific literature. It would also need a secure financial model, and because the commercial potential is unclear, this would probably mean public funding. Whether such an archive would justify the cost and effort required is a matter of debate, but unless or until some solution is devised, then negative results (which after all represent a high proportion of all results and a huge investment of research funds) will continue to be discarded.

Science will never be completely foolproof, and erroneous work will continue to enter the literature—whether through honest mistake, low scientific standards or clever fraud. Peer review can (and often does) catch at least some of these problems before a study reaches publication, but when bad science makes it into the public domain, it is important that its influence be limited. Science is undoubtedly self-correcting over the long term, but no one can plausibly claim that the process is prompt and efficient in every case.

1. Budd, J. M., Sievert, M. E. & Schultz, T. R., *JAMA* **280**, 296– 297 (1998).