Scientific Misconduct

Charles Gross

Department of Psychology and Neuroscience Institute, Princeton University, Princeton, New Jersey 08544; email: cggross@princeton.edu

Annu. Rev. Psychol. 2016. 67:693–711

First published online as a Review in Advance on August 13, 2015

The *Annual Review of Psychology* is online at psych.annualreviews.org

This article's doi: 10.1146/annurev-psych-122414-033437

Keywords

scientific misconduct, fabrication, falsification, plagiarism

Abstract

Scientific misconduct has been defined as fabrication, falsification, and plagiarism. Scientific misconduct has occurred throughout the history of science. The US government began to take systematic interest in such misconduct in the 1980s. Since then, a number of studies have examined how frequently individual scientists have observed scientific misconduct or were involved in it. Although the studies vary considerably in their methodology and in the nature and size of their samples, in most studies at least 10% of the scientists sampled reported having observed scientific misconduct. In addition to studies of the incidence of scientific misconduct, this review considers the recent increase in paper retractions, the role of social media in scientific ethics, several instructional examples of egregious scientific misconduct, and potential methods to reduce research misconduct.

Contents

INTRODUCTION

Over the past few decades, there has been an apparent outbreak in scientists behaving very badly. One such case is that of Dutch social psychologist Diederik Stapel, who fabricated more than 50 influential studies, usually "finding" things that academic liberals wanted to believe, including that dirty environments encouraged racism, that eating meat made people selfish, and that power had a negative effect on morality. His studies elevated his status to that of a superstar and dean before some of his students finally turned him in (Bhattacharjee 2013). Then there is the case of Eric Poehlman, a University of Vermont medical scientist who published 10 reports containing fabricated patients and data to support his claim of metabolic changes in menopausal women. Poehlman defrauded the government of $2.9 million in grants and became the first US scientist jailed for scientific fraud (Dalton 2005).

In this review, we consider the incidence of scientific misconduct and what might be done about it. We use the latest National Institutes of Health (NIH) and National Science Foundation (NSF) definitions of scientific misconduct as consisting of fabrication, falsification, and plagiarism. Fabrication is making up data or results. Falsification is manipulating research materials, equipment, or processes, or changing data or results. Plagiarism is the appropriation of another person's ideas, processes, results, or words without giving appropriate credit (NSF 1996, ORI 2011).

Research misconduct does not include honest errors or differences of opinion. Under the federal (NIH and NSF) definitions, it also does not include more general offenses, such as misuse of funds, sexual abuse, or discrimination. Furthermore, other dubious behaviors of a scientist that are more difficult to assess, such as in the determination of authorship or in the responsible mentoring of students, are excluded.

SCIENTIFIC MISCONDUCT IN THE PAST

Scientific misconduct is not a recent phenomenon simply tied to some decline of morality or increased competition for tenure and research funds. Rather, accusations of scientific misconduct,

sometimes well supported, pepper the history of science, from the Greek natural philosophers onward.

Ptolemy of Alexandria (90–168), the greatest astronomer of antiquity, has been accused of using (without attribution) observations of his predecessor Hipparchus of Rhodes (~162–127 BCE), who in turn used much earlier Babylonian observations as if they were his own (Newton 1977; cf. Neugebauer 1875). Isaac Newton used "fudge factors" to better fit data to his theories (Kohn 1986, Westfall 1973). Gregor Mendel, in his work with crossing pea plants, reported near-perfect ratios, and therefore statistically very unlikely ones. The high unlikelihood of getting exact ratios was first pointed out by Ronald A. Fisher, the founder of modern statistics and one of the founders of population genetics, when he was still an undergraduate at Cambridge University in 1911 (Franklin et al. 2008). Though Charles Darwin has been cleared of accusations of stealing the idea of natural selection from Alfred Russell Wallace, he seems to have only reluctantly credited some of his predecessors (Eisley 1979, Gross 2010). Robert A. Millikan, in his measurement of the charge of an electron, which led to his Nobel Prize in Physics in 1923, failed to report unfavorable data (Kohn 1986; cf. Goodstein 2010). Incidentally, Millikan also failed to give coauthorship to his student Harvey Fletcher, whose work was crucial to the discovery (Kohn 1986), but in those days—as still today—this was not scientific misconduct as defined by the NIH and NSF. William T. Summerlin at Memorial Sloan-Kettering faked a skin transplantation experiment in 1974 with the help of a black marker pen (Hixson 1976), giving rise to the term "paint the mouse" as a synonym for scientific misconduct.

The first formal discussion of scientific misconduct is Charles Babbage's *Reflections on the Decline of Science in England, and on Some of Its Causes* (1970/1830). Babbage held Newton's chair at Cambridge and made major contributions to the development of computers ("difference machines," "analytical engines") and to astronomy, mathematics, and many other fields. He distinguished "several species of impositions that have been practiced in science . . . hoaxing, forging, trimming and cooking" (p. 174). An example of hoaxing would be the Piltdown Man (discovered in 1912 and exposed in 1953): Parts of an ape skull and a human skull were put together, supposedly to represent a "missing link" (Weiner 1955). Hoaxes are intended to expose naïveté and credulousness and to mock pseudowisdom. A more recent brilliant hoax that drew a flurry of attention in the academic world was Alan Sokal's article in *Social Text* (1996), in which he purported to argue that "physical reality . . . is at bottom a social and linguistic construct" (p. 217).

Unlike hoaxes, the other impositions distinguished by Babbage are carried out to advance the perpetrator's scientific career. Forging, which Babbage thought rare, is making up results, today called fabrication. Trimming consists of getting rid of outliers to decrease the variance and make the results look better. Cooking is the selection of data. Trimming and cooking fall under the modern rubric of falsification.

ORIGINS OF UNITED STATES GOVERNMENT INVOLVEMENT

Around 1980, several cases of egregious scientific misconduct were extensively publicized. One was the case of John Long. While working at Massachusetts General Hospital, Long apparently made a major advance in studying Hodgkin's disease by developing four lines of tissue culture cells from Hodgkin's disease tumors, enabling the disease to be studied in detail. However, Long's colleagues eventually discovered that three of the lines were from a healthy monkey and the fourth from a normal human (Wade 1981). A second case was that of Vijay R. Soman, a promising Yale University assistant professor of medicine who working on anorexia nervosa. Twelve of Soman's papers eventually had to be retracted, ending his research career and seriously sullying that of his mentor and protector. (This case is sufficiently instructive that I return to it in detail

below.) These and other cases spurred US Representative Al Gore, who was then chairman of the Investigations and Oversight Subcommittee of the House Science and Technology Committee, to hold hearings on "Fraud in Biomedical Research" on March 31 and April 1, 1981 (US House Rep. 1981). This marked the beginning of the US federal government's involvement in scientific misconduct.

Gore began the hearings by noting, "We need to discover whether recent incidents [of data falsification] are merely episodes that will drift into the history of science as footnotes, or whether we are creating situations and incentives . . . that make such cases as these the tip of the iceberg" (US House Rep. 1981, p. 1). The first witness was Philip Handler, president of the National Academy of Sciences, who made it clear that he found "little pleasure and satisfaction in testifying" because in his view the problem had been "grossly exaggerated in the full, and occasionally fulsome treatment that it has had in the press" (US House Rep. 1981, p. 10). He contended, "The matter of falsification of data . . . need not be a matter of general societal concern. It is, rather, given the size of the total research effort, a relatively small matter which . . . is normally effectively managed by . . . the scientific community itself" (pp. 10–11). Handler added, "None of us know the real magnitude of the problem that concerns you. Nor can we" (p. 11). Yet he went on to say, "The number of untoward incidents involving fraud or falsification of data is remarkable, not so much for the fact that they occur, but for the fact that they occur so very rarely" (p. 13). As to what should be done about the problem of falsification in science, Handler said, "We would much prefer that the institutions within which science is conducted be responsible for their own affairs, that is, the universities, research institutes, etcetera" (p. 50).

The next witness was Donald S. Fredrickson, director of the NIH. He opined, "I think cases of downright fraud in science have always been quite rare . . . I don't know whether scientific fraud in less spectacular forms occurs more frequently today than it has in the long history of science" (US House Rep. 1981, p. 26). Gore found it quite "perplexing" that the witnesses claimed "that the problem of fraud is essentially a private one that should be dealt with by the informal codes of the scientific community and that fraud was not an important ethical problem that should worry those of us charged with the public trust" (p. 111). Another member of the committee was "taken somewhat aback by the chastisement of the committee for being presumptuous enough to even have the hearing at all" (p. 154).

On both days of the hearings, witnesses made it clear—somewhat to the committee's surprise—that neither the government nor most institutions that conduct scientific research had any formal and public procedures for dealing with scientific misconduct. In the years after the hearings, cases of misconduct continued to surface, and Congress continued to be interested. In 1985, Congress passed the Health Research Extension Act, which required the Secretary of Health and Human Services to issue a regulation requiring applicant or awardee institutions to establish "an administrative process to review reports of scientific fraud" and "report to the Secretary any investigation of alleged scientific fraud which appears substantial"; the Act also required the NIH director to establish a process for receiving and responding to reports from institutions (Goodstein 2010, p. 59). Four years later, the Office of Scientific Integrity (OSI) within the NIH was established. The OSI's mission was to protect the integrity of scientific research by investigating scientific fraud (Goodstein 2010). In 1992, the OSI moved out of the NIH and became the Office of Research Integrity (ORI) within the Public Health Service. The NSF created a similar Office of the Inspector General.

The OSI and Office of the Inspector General definitions of misconduct gradually became more specific and uniform until 2000, when the Office of Science and Technology Policy in the Office of the President set out the federal government definitions given in the second paragraph of the present review.

The NIH and NSF now require all institutions that apply for research support to have a set of procedures for dealing with allegations of scientific misconduct. These are usually available on each institution's website. Very briefly, the typical drill is that after an allegation is made to a university departmental chair or dean, an inquiry is set up to determine if a formal investigation is warranted. If it is, a small committee of faculty members from other departments carries out the investigation. During both phases, the accused scientist is given opportunities to respond, and all proceedings are supposed to be completely confidential. The investigating committee is given full access to all computer disks, unpublished data, and notes from research supported by the federal government, as all such material legally is the property of the federal government, which vests its control in the university. Specific time limits are usually set for each stage of the process.

If the investigation finds misconduct, the university can take a variety of possible actions, ranging from removal of the scientist from the particular project and withdrawal of published papers to firing. The ORI or an equivalent federal agency then conducts it own investigation. The agency has the power to deny, for a specific period or forever, future research funds and study section memberships. Federal prosecution for misuse of research funds can occur as well. Partial or total secrecy is usually attempted until after the federal investigation is completed, although earlier leaks to newspapers sometimes occur.

Using the representative University of Maryland procedures, Shamoo & Resnik (2003) estimate that if all deadlines are met and no "hitches" occur, the internal university process should take approximately 10 months, excluding the time for possible appeal at the university level and then the ORI deliberations and possible appeal from their decisions. This seems like an underestimate. In the recent case of Harvard Professor Marc Hauser's misconduct in connection with a small number of studies in comparative cognition, the investigation took three years and the ORI deliberations another two. Yet there do not appear to have been any hitches, delays, or appeals, and the Harvard committee is reported to have met diligently and often (Gross 2012). Sometimes this process of investigating scientific conduct is even more prolonged: In the well-publicized and complex case of Margot O'Toole, Thereza Imanishi-Kari, and David Baltimore, the dispute was not resolved until 10 years later (by the end of which Baltimore was president of Caltech, where today he is the Robert A. Millikan Professor of Biology) (e.g., Judson 2004, Kevles 1998, Sarason 1993).

INCIDENCE OF MISCONDUCT TODAY

A very low incidence of scientific fraud was claimed not only by über establishment figures Handler and Fredrickson but also by the usually acute founder of the modern sociology of science, Robert Merton, who thought scientific fraud to be "extremely infrequent" (Merton 1957, p. 651). In the decades since Gore asked whether the well-publicized examples of scientific misconduct starting in the 1970s were just a few bad apples or the tip of the iceberg, a number of attempts have been made to answer this question. Studies began asking scientists at every level in a variety of fields, and under the cover of anonymity, whether they themselves had engaged in fabrication, falsification, or plagiarism or had direct evidence of such scientific misconduct by others. Although the results were variable and involved different survey response rates and methodologies, the overall picture is disturbing.

In a large and pioneering survey of science graduate students and faculty at 99 universities, ethicist and biology historian Judith Swazey and colleagues found that 44% of students and 50% of faculty had knowledge of two or more types of misconduct, broadly defined; approximately 10% had observed or had direct knowledge of fabrications of data (Swazey et al. 1993). The International Society of Clinical Biostatistics surveyed its membership and found that 51% of the respondents knew of at least one fraudulent project in the previous 10 years (Ranstam et al.

2000). Of 549 biomedical research students at the University of California, San Diego, 11% said they had firsthand knowledge of scientists intentionally altering or fabricating data for purposes of publication (Kalichman & Friedman 1992). In a similar survey, 9% of biological and medical postdoctoral fellows at the University of California, San Francisco, said they had observed scientists altering data, and 3% admitted to having fabricated data (Eastwood et al. 1996). The American Association for the Advancement of Science surveyed a random sample of its members, and 27% of the respondents said they believed they had encountered or witnessed fabricated and falsified or plagiarized research over the previous 10 years, with an average of 2.5 examples (Titus et al. 2008). In a study carried out by the Gallup Organization for the ORI, of 2,212 researchers receiving NIH research grants, 210 reported instances of federally defined misconduct over a three-year period, of which 60% were fabrication or falsification and 36% plagiarism. Noting that 155,000 personnel receive research support, the authors suggest that under the most conservative assumptions, a minimum of 2,325 possible acts of research misconduct occur in a year (Wells 2008). Finally, in a meta-analysis of 18 studies, 2% of scientists admitted to fabricating or falsifying data, and 14% had observed other scientists doing the same (Fanelli 2009).

Retractions of Published Papers

Further insights into the incidence of scientific misconduct come from studies of retractions of published scientific articles. A particularly strong examination of this phenomenon is that of Fang et al. (2012). Unlike previous studies of retraction (e.g., Grieneisen & Zhang 2012, Nath et al. 2006, Steen 2011, Wagner & Williams 2011), Fang and colleagues explored the reason for the retraction if none was mentioned in the retraction notice, as was often the case. Among their sources for the reason for a retraction were reports from ORI, blogs such as Retraction Watch (**http://retractionwatch.com**; see below), news media, and other pubic records. Their review of 2,047 retractions of biomedical articles indicated that only 23% were due to error, whereas 67% were attributed to some type of misconduct such as fraud or suspected fraud (43%), duplicate publication (14%), and plagiarism (10%) (see figure 1 in Fang et al. 2012). The most common nonmisconduct errors that led to retractions were laboratory errors, analytical errors, and irreproducible results (Casadevall et al. 2014).

Fang et al. (2012) found a tenfold increase in retractions for fraud or suspected fraud since 1975. The rates of retraction varied across academic disciplines, with cell biology and oncology being on the high end and sociology and political science on the low end (Margraf 2015). Presumably, this reflects that criteria are more objective and the medical consequences of error more serious in cell biology and oncology than in sociology and political science. The incidence of retractions in psychology was intermediate (Margraf 2015).

Considerable variation exists in geographic origins, with the United States, Germany, Japan, and China responsible for three-quarters of the cases of fraud or suspected fraud; in comparison with other countries, China and India accounted for more cases of duplicate publication and plagiarism (Fang et al. 2012). The authors note several cases of fraudulent articles that were never retracted and suggest this indicates that the number of retractions is an underestimate of the incidence of fraud. Fang et al. (2012) and others (e.g., Steen 2011) have noted a correlation between the impact factor of a journal and the number of retractions for fraud or suspected fraud.

There are a several possible reasons why the marked increase in the incidence of retractions in the past 10 years may not have been primarily due to an increase in scientific misconduct (Fanelli 2013, Steen et al. 2013). First, although the number of retracted journal articles has grown, the number per retracting journal has not. Second, the frequency of finding of misconduct by OSI has not increased. Third, retractions for plagiarism and duplicate publication are

a relatively new practice and thereby enhance the contemporary rate of retraction. Finally, and most likely to be the primary reason for the increase, is the greater scrutiny by editors and reviewers, who also have more of a tendency to look into the past papers of an author found guilty of misconduct.

Retracted papers continue to be cited even after their retractions (see, e.g., Budd et al. 1998, Grieneisen & Zhang 2012). Eighteen percent of authors of retracted papers continue to cite them, and of these, less than 5% mention that the papers were retracted (Madlock-Brown & Eichmann 2015). In addition to self-citation, a reason for the continued citing of retracted papers may be that the majority of citations in the literature seem to derive from secondary sources, not from the original paper (Broadus 1983, Siskin & Roychowdhury 2006), and therefore the citer often may not know of the retraction.

Role of Social Media

Social media has entered the world of scientific ethics, and there are now blogs that deal with possible scientific misconduct. Among the early blogs was Retraction Watch (**http://retractionwatch.com**), founded in August 2010 by science writers and editors Ivan Oransky and Adam Marcus to publicize and comment on retractions of scientific papers. They note that many retractions and the reasons for them remain buried in obscurity and say they view tracking retractions as a window into the scientific process.

Another new postpublication review is PRE: Peer Review Evaluation, whose stated purpose is to support publishers in promoting trust and transparency related to peer review. They say they hope to provide information on a journal's peer review practices, such as the number of rounds of review an article went through and what roles the editors played.

One blog devoted to scientific misconduct is Aubrey Blumsohn's Scientific Misconduct Blog (**http://scientific-misconduct.blogspot.com**), which began in July 2006. It includes links to a number of other blogs that include discussions of scientific misconduct as well as many topics related to academic and big pharma life and ethics. Some of these blogs specialize in the postpublication discussion of papers, something like a very critical lab meeting on problematic research practices or obvious misconducts.

Another example of postpublication peer review is PubPeer (**http://www.blog.pubpeer.com**). Recently, PubPeer was sued by a cancer scientist at Wayne State University, who claimed he lost a prospective tenured position at the University of Mississippi as a result of anonymous comments in PubPeer that suggested misconduct in his research (Servick 2014).

One blog devoted to scientific misconduct (sciencefraud.org, which is no longer active) had the unusual opportunity to examine the effects of Internet publicity (Brookes 2014). From July to December 2012, this blog site received and published 274 anonymous emails claiming data integrity problems, most of which were in biomedical journals. In January 2013, legal threats forced the closure of this site, but an additional 233 anonymous emails claiming integrity problems were submitted and could not be published. Brookes (2014) found no differences in the characteristics of the 274 "public" cases and the 223 unpublished "private" cases submitted between November 2012 and January 2013. He then examined whether papers in the two groups (*a*) were retracted, (*b*) had errata published, or (*c*) had no action taken. He found that in comparison with the private set, the public set had a rate of retractions that was 6.5 times higher and a rate of corrections that was 7.7 times higher. Overall, some kind of corrective action was taken on 23% of the publicly discussed papers as opposed to 3.1% of the private, nondiscussed papers. These findings suggest the value of Internet publicity for greater scientific integrity.

Similar to the growth in postpublication review blogs has been the increased use of laboratory websites for the critique of published papers. This use of the Internet is analogous to posting the results of a lab meeting critique of a published paper.

WHO ARE THE MISCREANTS?

Scientists guilty of misconduct have been found in many fields and at different levels in universities and research institutions; their social and educational backgrounds vary. There appear to be no systematic empirical studies of the characteristics of perpetrators of scientific misconduct and no good evidence for any common characteristics. However, there does seem to be a modal scientist whose misdeeds in science are well publicized in *The New York Times*, *Science*, and *Nature* and in books on the subject (e.g., Broad & Wade 1982, Goodstein 2010, Judson 2004). This scientist is a bright and ambitious young man working in an elite institution in a rapidly moving and highly competitive branch of modern biology or medicine, where results have important theoretical, clinical, or financial implications. He has been mentored and supported by a senior and respected establishment figure who is often the coauthor of many of his papers but may have not been closely involved in the research.

The following sections give examples of two well-publicized cases that fit this stereotype. They are probably not typical of the entire of sample of scientists engaged in misconduct, but they are instructive because they involve senior US scientists and distinguished US institutions; therefore, they may indicate some of the problems that government and research institutions are struggling with in maintaining scientific integrity. These cases have also, presumably, spurred the recent attempts of federal government agencies and research institutions to reduce scientific misconduct.

The Case of Vijay Soman

The story of Vijay Soman began on November 9, 1978, when a young Brazilian physician, Helena Wachslicht-Rodbard, who was working in the laboratory of Jesse Roth, chief of the diabetes branch of the National Institute of Arthritis, Metabolism and Digestive Diseases at the NIH, submitted a paper on insulin binding in patients with anorexia nervosa to the prestigious *New England Journal of Medicine* (*NEJM*). The paper, with Roth as coauthor, showed that insulin receptors in blood cells of patients with anorexia nervosa bound more insulin than normal. Editor Arnold Relman sent it to two referees, one of whom recommended rejection, so a third referee was called in. After two and a half months, the editor wrote Rodbard that two of the referees were favorable and one recommended rejection. He also wrote that the paper could be acceptable after minor revisions.

Unknown to Rodbard, the negative reviewer had been Philip Felig, an endowed professor at Yale School of Medicine, vice chairman of the Department of Medicine, and chief of endocrinology research. Felig gave Rodbard's paper to an assistant professor in his laboratory, Vijay Soman, who had been supposedly working on the identical subject as that of the paper. Shortly after receiving Rodbard's paper, Felig sent off a paper on the same subject, with Soman as coauthor, to the *American Journal of Medicine*.

Ironically, the Soman-Felig manuscript was send to Roth for review, and Roth passed it on to Robard. To quote Broad & Wade (1982, p. 165), "She was aghast. Here was her paper, complete with verbatim passages . . . and even a formula she had devised." Comparing the content and typeface of the submitted paper with that of the negative review, she realized that Felig had written the negative review. She immediately wrote to *NEJM* Editor Relman, accusing Felig and Soman of "plagiarism, of conflict of interest in reviewing her paper and of trying to slow down the acceptance of her work" (Broad & Wade 1982, p. 166).

In late February of the next year, Relman called Felig to discuss the conflict of interest charges, and Felig claimed (incorrectly) that their work had been completed before the Rodbard paper arrived. In any case, Relman then published the Rodbard-Roth paper in the *NEJM*. A week later, Felig and Roth (grade-school buddies from Brooklyn) got together at the NIH and compared the two manuscripts. Felig agreed that there was a minor plagiarism problem and said he would (*a*) refer in the Rodbard-Roth paper in the Soman-Felig manuscript, (*b*) delay publishing theirs so Rodbard would have priority, and (*c*) not publish it at all if there were any "legitimate" questions about the independence of Soman's work. Returning to New Haven, Felig confronted Soman, who confessed to have used Rodman's paper in writing his own, but Felig saw enough of Soman's notes to satisfy himself that Soman had done the work on their paper.

Meanwhile, Rodbard had come to believe that beyond the plagiarism, the Soman-Felig study had been entirely fabricated on the basis of her study. Roth didn't accept this view, told Felig so, and pressured Rodbard not to pursue the issue. Rodbard then wrote to Robert Berliner, dean of Yale School of Medicine, and demanded an investigation to establish the authenticity of the data in the Soman-Felig paper, giving a number of cogent reasons for this in addition to the plagiarism. For example, neither the hospital nor the collaborating psychotherapists were identified, and the author seemed unfamiliar with the methodology. If such an investigation was not undertaken, Rodbard threatened to denounce the Soman-Felig paper at an upcoming meeting of the American Federation of Clinical Research. The requested investigation, or audit, is unusual in the world of academic research. A mutually agreed auditor, Joseph E. Rall, Roth's boss at the NIH, was chosen as the auditor to examine Soman's data.

In the interim, two major developments occurred in the career trajectories of Rodbard and Felig in July 1979. First, Rodbard quit research at the NIH to take a residency in internal medicine. Second, Felig was offered and accepted the prestigious Bard professorship and chairmanship of Columbia University's Department of Medicine, a post he planned to take up in June 1980. In January he took Soman to Columbia and recommended he be made an assistant professor. In the same month, the Soman-Felig paper, which Felig had agreed to hold back until the audit was complete, was published in the *American Journal of Medicine*. Rall, the prospective auditor, had been too busy, so the agreed-upon audit had never been carried out.

Rodbard continued to demand that Roth get another auditor. In February, the new auditor, Jeffrey Flier, an assistant professor at Harvard Medical School (now dean), examined the data that formed the bases of Soman's paper. Soman soon confessed to "fudging" and fabricating data but said that Felig didn't know about it. Soman said, "He'd been under great pressure to publish as soon as possible to obtain priority" (Broad & Wade, p. 173). Upon hearing Soman's confession, the Yale Department of Medicine chairman told Soman that "his best choice was to resign and give up research" (Broad & Wade 1982, p. 174). At that point, the only excuse Soman gave for his misconduct was that it was his "fate" (Broad & Wade, p. 174). He disappeared into India a few weeks later.

Felig thereupon retracted the disputed article with Soman, which he had published in violation of the agreement with Rodbard. Because of Rodbard's continued charges about data fabrication by Felig's laboratory, Yale called in another auditor, Jerrold Olefsky, then of University of Colorado, to examine all of Soman's 14 publications dealing with insulin receptors. Of them, Olefsky could find data supporting only two. Felig had coauthored 10 of the remaining papers. Yale retracted the 12 papers by the end of May. In August 1980, Columbia asked Felig to resign from his new post as Bard Professor and chair of the Department of Medicine.

Yale rehired Felig as a tenured full professor but not to his original endowed chair. He was found not guilty of fraud by an NIH investigating committee, and his major NIH grant was renewed. A few years later he left Yale to work briefly for Sandoz, a major pharmaceutical company, and is

now in private practice in New York. Apparently, he has disappeared from the biomedical scene and no longer publishes or attends the major meetings he had attended in the past.

Rodbard is now in private practice as an endocrinologist in the Washington, DC area and is active in clinical research. She is past president of the American College of Endocrinology, past president of the American Association of Clinical Endocrinologists, a Master of the American College of Endocrinology (an average of only two are so honored every year), and a Fellow of the American College of Physicians. She has been consistently listed in Best Doctors in America and Best Doctors in Washington, DC. Under the name H.W. Rodbard, she is listed in Pub Med as having authored or coauthored 49 clinical endocrinology papers since she left the NIH to enter private practice. [This (abbreviated) account of the story of Vijay Soman is based on Altman (1980), Broad (1980), Broad & Wade (1982), Hunt (1981), and H.W. Rodbard (personal communication).]

The Case of John R. Darsee

The next case is again about the fall of a promising young man protected by a senior figure, this time at Harvard Medical School. In 1979, John R. Darsee, after a brilliant five-year research and clinical career at Emory University, joined Eugene Braunwald's large cardiac research laboratory at Harvard. Braunwald held the hoary chair of Professor of Theory and Practice of Physic at Harvard and was chairman of Medicine and physician-in–chief at both the Brigham and Women's and Beth Israel hospitals. In 1979, Braunwald had over 800 papers listed in PubMed and more than $3.3 million in NIH funding.

Darsee went to Harvard as an NIH postdoc and was slated for a faculty appointment at Harvard Medical School. In two years at Harvard, he published over 100 abstracts and papers (Broad 1983a,b). Braunwald often called Darsee "one of the most outstanding of the 130 research fellows he has trained, more than 40 of whom are now full professors, department chiefs, or directors of academic cardiology divisions" (Knox 1983, p. 1799).

Several of Darsee's lab colleagues had been suspicious of his prodigious output. For example, he claimed "to have done some complicated long-term dog experiments . . . spanning up to six weeks and involving indwelling intracoronary pressure transducers, ultrasonic crystals, and meticulous sterile techniques [which] no one in the laboratory could remember seeing" (Knox 1983, p. 1802). Finally, in May 1981, some of his coworkers went to Robert Kloner, their immediate supervisor, and suggested that Darsee had faked the data for an abstract that he was about to send off. When Kloner confronted Darsee and asked to see the raw data, Darsee started taking a "hemodynamic reading off a dog in an experiment and as the chart paper came out, he marked it day 1, day 2 and so forth, making it look as though the data had been taken over the course of several days" as several colleagues "watched in awe" (Broad 1982a, p. 479). When later confronted by Kloner, Darsee admitted his misconduct but called it a single foolish act performed because he had thrown away the real raw data (Culliton 1983b). Kloner reported this to Braunwald, who initiated an investigation of Darsee's work. As a result of this inquiry, the abstract was never sent out and Darsee's NIH postdoc and his Harvard appointments were terminated, but the NIH was not informed of his fabrications.

Yet Darsee continued his active cardiac research in the lab. He was paid from private funds controlled by Braunwald and continued to publish many abstracts and papers, most of them coauthored by Kloner or Braunwald. Braunwald apparently trusted Darsee in spite of his recently demonstrated misconduct (Braunwald 1992). Darsee wrote Braunwald that his single misdeed "should be considered in the context of the 'laboratory environment,' which he characterized as bristling with envy, spite, spying, and even sabotage directed toward him . . . as such a hard worker, meticulous scientist, and lone wolf" (Knox 1983, p. 1803).

At this time Braunwald's laboratory in Brigham and Women's Hospital was participating in an NIH-sponsored multi-institutional study of drug treatment for heart attacks. When the data from the different institutions were compared, the Harvard data, collected by Darsee, were different from those of the other labs. When challenged, Darsee could not support his claims. Harvard and the NIH set up formal committees to investigate these discrepancies.

In January 1981, the Harvard committee revealed multiple fabrications by Darsee. In March, the NIH committee reported "extensive irregularities" in five papers authored by Darsee, Kloner, and Braunwald and suggested one cause was insufficient supervision by Braunwald (Broad 1982b). The NIH required Brigham and Women's Hospital to return $122,371 in research funds designated for their part of the study, which apparently marked the first time that money had to be returned to the NIH because of fraud.

In response to the implication of the NIH report that his lax supervision had contributed to Darsee's behavior, Braunwald defended his practices and declared, "I had gotten a bum rap" (Culliton 1983b, p. 31). In an apparent effort to counter the implication that he was responsible for Darsee's misconduct at Harvard, Braunwald then reviewed some of Darsee's papers from Emory and reported his concerns about their integrity to both the NIH and Emory. An Emory investigation concluded that only 2 of the 10 papers published by Darsee from Emory were valid, and of 45 abstracts, only 2 stood up to scrutiny. The investigating committee viewed many of the publications as fiction and noted that some mentioned collaborators who did not exist.

Darsee left Harvard and obtained a clinical position at a Schenectady hospital. He issued the following statement: "I am asking for forgiveness for whatever I have done wrong and want to contribute to the medical system" (Culliton 1983b, p. 35). He excused his behavior in a letter: "I had too much to do, too little time to do it . . . I had not taken a vacation, sick day or even a day off from work for six years. I had put myself on a track that I hoped would allow me to have a wonderful academic job" (Knox 1983, p. 1805). [This account is based on Broad (1982a,b, 1983a,b); Broad & Wade (1982); Culliton (1983a,b,c); Knox (1983); Moran (1985).] In 1984, New York State revoked Darsee's license to practice medicine. He appears to be currently practicing internal medicine in Indianapolis (**https://www.doximity.com/pub/john-darsee-md**).

In spite of the considerable publicity (e.g., the references cited in the previous paragraph) about Darsee's fraudulent publications, these publications continued to be cited in a positive fashion with no mention of fraud (Kochan & Budd 1992, Moran 1985). From 1982 to 1990, only 5.7% of 256 citations to Darsee acknowledged fraud. The continued citation of retracted articles is a more general phenomenon (Madlock-Brown & Eichmann 2015).

Walter Stewart and Ned Feder, two NIH scientists, used the Darsee issue to inquire into the roles of coauthors, editors, and referees in Darsee's publications. They examined 109 of Darsee's papers, abstracts, and chapters, which had a total of 47 coauthors at Emory and Harvard, approximately half junior and half senior, the latter including departmental chairs and full professors (Stewart & Feder 1987). They were not concerned with Darsee's science or his fraud. They were only concerned with the papers themselves—with "lapses in generally accepted standards" (p. 207) for publication, such as internal errors (major and minor), inconsistencies in numerical values within the papers and with previous publication, lack of controls, inadequately presented data, unsupportable claims, honorary authorship, repeated publication of the same data without acknowledgment, and similar flaws that "should have been detected by any competent scientist who read the papers carefully" (Boffey 1986), let alone their coauthors (Boffey 1986, Mervis 1986, Stewart & Feder 1987).

They unsuccessfully submitted a draft of their paper to *Nature* in 1983 and to *Cell* in 1985, "and the two journals were bombarded with more than 20 letters and memorandums, some 150 typed pages in all, from lawyers for the scientists accused of misdeeds"; most of the correspondence

originated from the lawyers for Braunwald, probably the most eminent of the scientists involved (Boffey 1986; N. Feder, personal communication). They also submitted their paper to another 14 journals, none of which were interested in publishing it (Boffey 1986).

In letters to some of Darsee's Emory and Harvard coauthors, Stewart and Feder inquired about apparent flaws in their papers with Darsee. In September 1983, Lester Salans, chief of the National Institute of Infectious Diseases, wrote Stewart that he and Feder did "not have the authority to act on behalf of NIH or to utilize official time and/or with the aid of NIH resources, including NIH stationery" (N. Feder, personal communication). According to Feder, "This was the first of many episodes in which NIH officials challenged the legitimacy of our examination of the practices of Darsee's coauthors, as well as the legitimacy of our subsequent public comments, spoken and in print, on these practices" (N. Feder, personal communication).

Eventually, *The New York Times*, under a Freedom of Information Act request, obtained a copy of the article, along with its voluminous legal history, from the NIH and published an account (Boffey 1986). Soon afterward *Nature* published a toned-down version of the paper, i.e., one in which the editors had made changes without the authors' permission (Stewart & Feder 1987). In the same issue of *Nature*, Braunwald (1987) answered Stewart and Feder. In his long reply in *Nature*, Braunwald concentrated on defending his lab and characterized the admitted errors in the various papers he coauthored with Darsee as either Darsee's fault or trivial.

Their work on Darsee was only the beginning of Stewart and Feder's long, distinguished, and sometimes controversial career of exposing scientific misconduct. The other issues they were active on included the O'Toole, Imanishi-Kari, and Baltimore controversy discussed below, the Gallo case on priority for the discovery of HIV, and the phenomenon of NIH grantees receiving huge fees from drug companies. For these and similar activities, they were subjected to many years of harassment and threats to stop their investigations by their NIH bosses. Near the end of their government career, Stewart and Feder invented a plagiarism-detecting computer program and used it to file a 1,400-page report accusing the best-selling Lincoln historian Stephen Oates of massive plagiarism. This led the NIH to close their lab, seize their files, and transfer them to positions that were incompatible with misconduct investigations (Hilts 1993a). Stewart went on a 33-day hunger strike in protest, but to no avail. Many scientists defended them, including Margot O'Toole, who said, "But when all the critics are silenced, who will be left to dissent?" (O'Toole 1993).

In 2007, Feder left the NIH to continue his work on scientific misconduct at the private Project on Government Oversight (POGO) (Feder 2012). Accounts of Stewart and Feder's campaign against scientific misconduct can be found in Boffey (1986); Feder (2008, 2012, personal communication); Grossman (1993); Hoke (1993, 1995a); Monastersky (2008); Powledge (1987); Sarasohn (1993); Wade (1988); Wadman (2008).

PREVENTION OF SCIENTIFIC MISCONDUCT

In the previous sections, we presented data on the incidence and characteristics of scientific misconduct. We also recounted two case histories in two leading biomedical research facilities; the case histories suggest that some institutional structures may foment scientific misconduct. In this final section, we consider ways that scientific misconduct might be reduced.

Whistle-Blowing

Fabrication and falsification of scientific data can be uncovered in a variety of ways. One way is through referees or readers identifying internal contradictions and inconsistencies with previous

papers from the same lab. Another way is through the detection of previously published or apparently doctored figures. An unusual way, as in the case of Helena Wachslicht-Rodbard, is by discovering plagiarism of one's own work and detecting the similarity of fonts in a referee report with those in a submitted paper. However, most detection of scientific misconduct appears to come from laboratory colleagues: from students, peers, technicians, and supervisors of the miscreant (Shamoo & Resnik 2003).

The individual making the accusation of possible misconduct is often termed a whistle-blower, especially if he or she is not the senior member of the research unit. Even if the whistle-blowing turns out to be justified, the consequences for the whistle-blowers are often disastrous in terms of their income, research, personal relations at work, and future in science, as has been repeatedly related (e.g., Alford 2002, D'Angelo 2012, Hilts 1993b, Lock et al. 2001, Martin 2013, Penslar 1995, Rivlin 2004).

The plight of the whistle-blowing junior scientist was widely publicized in the case of Margot O'Toole versus Thereza Imanishi-Kari. The case, which began in 1986, continued for over ten years and, *inter alia*, involved Congress, the Secret Service, fraud-busters Feder and Stewart, and the resignation of Nobel Laureate David Baltimore as president of Rockefeller University (see, e.g., Kevles 1998; Judson 2004; O'Toole 1991, 1993; Sarasohn 1993). O'Toole, a postdoc working in the laboratory of Imanishi-Kari, questioned the latter's data in a paper about genes that are involved with the immune system of mice; the paper was coauthored by Imanishi-Kari and Baltimore. Independent of whether O'Toole's accusations were valid and whether Imanishi-Kari's treatment of data was sloppy, fraudulent, or neither, it was clear that Baltimore, the senior scientist, treated O'Toole's stubborn skepticism with inappropriate arrogance, intolerance, and brutality. Baltimore eventually apologized to O'Toole: "I commend Dr. O'Toole for her courage . . . and I regret and apologize for my failure to act vigorously enough in my investigation of her doubts" (Hilts 1991).

Perhaps partially as a result of this incident, the ORI began to systematically investigate the consequences of whistle-blowing for the whistle-blower. In 1995, the ORI reported that "over 60% of whistleblowers suffered at least one negative consequence, such as being pressured to withdraw their allegation, being ostracized by colleagues, suffering a reduction in research support, or being threatened with a lawsuit. Over 10% noted significant negative consequences, such as being fired or losing support" (Lubalin et al. 1995, p. 51). The ORI then issued to institutions receiving NIH funds a detailed set of guidelines on protecting whistle-blowers against retaliation (ORI 1995). These guidelines provide specific and detailed procedures for investigating any claims by whistle-blowers of retaliation for their actions.

However, if retaliation of whistle-blowing is validated, the recommendations as to what to do next are rather vague: The deciding official shall determine what remedies are appropriate to satisfy the institution's regulatory obligation to protect whistle-blowers. The deciding official shall, in consultation with the whistle-blower, take measures to protect or restore the whistle-blower's position and reputation, including making any public or private statements, as appropriate. In addition, the deciding official may provide protection against further retaliation by monitoring or disciplining the retaliator (ORI 1995).

However, in the typical case of whistle-blowing in a research laboratory, these elaborate protections against retaliation may be of little value to whistle-blowers. For example, if the whistle-blowers are graduate students or postdocs and if and their evidence eventually is accepted, their supervisor is likely to leave and their lab likely to be closed, leaving the whistle-blowers without facilities, financial support, mentoring, or even a project; in addition, whistle-blowers sometimes are forced to endure the obloquy of other members of the department (e.g., Couzin 2006).

Because whistle-blowing by members of a laboratory is a major weapon against research misconduct, it would seem important to encourage such behavior rather than punishing it. For example, if the whistle-blower is a graduate student in good standing, he or she might be guaranteed financial support, laboratory facilities, and a mentor until he or she receives a degree. Postdocs or technicians might be insured a paid position for a year or two, until they obtain other support. More generally, the institution should commit to preventing whistle-blowing from resulting in career destruction, whatever the level of the whistle-blower.

Of course, premature, inadequately justified, unjustifiable, and/or inappropriately carried out whistle-blowing can be a disaster for all involved. There are a number of good and cautious guides to whistle-blowing (see, e.g., Gunsalus 1998 and other articles in the same issue; Hoke 1995b, Rennie & Gunsalus 2008).

Responsible Conduct of Research Courses

As a result of the scientific misconduct scandals and hearings in the late 1980s, the federal government has been increasingly involved in requiring education in the responsible conduct of research (RCR). As of 2011, the NIH has required that all trainees, fellows, participants, and scholars receiving support through any NIH training, career development award, research education grant, and dissertation research grant must receive instruction in responsible conduct of research (NIH 2009).

Similarly, the NSF requires each institution that applies for financial assistance to describe in its grant proposal a plan to provide appropriate training and oversight in the responsible and ethical conduct of research to undergraduate students, graduate students, and postdoctoral researchers participating in the proposed research project (NSF 2009).

The NIH and NSF provide detailed guidelines for the development of institutional RCR courses that have been developed over the past 20 years. These courses cover:

1. conflict of interest—personal, professional, and financial;
2. policies regarding human subjects, live vertebrate animal subjects in research, and safe laboratory practices;
3. mentor/mentee responsibilities and relationships;
4. collaborative research, including collaborations with industry;
5. peer review;
6. data acquisition and laboratory tools—management, sharing, and ownership;
7. research misconduct and policies for handling misconduct;
8. responsible authorship and publication; and
9. the scientist as a responsible member of society, contemporary ethical issues in biomedical research, and the environmental and societal impacts of scientific research.

The strongly recommended format for RCR courses is small-group discussion of case studies. A semester-long series of meetings (or at least eight contact hours) is recommended over a more compressed schedule. Online courses are not normally an acceptable substitute. Faculty and mentors are encouraged to participate as "discussion leaders, speakers, lecturers, and/or course director" on a rotational schedule to insure "full faculty participation" (NIH 2009).

Participation should occur at least once at every career level from undergraduate to faculty, at a frequency of no less than every four years. Detailed instructions for proposing and reporting on RCR implementation are available for each type of instructional and individual award.

Excellent print and online materials are available for such RCR courses [e.g., Macrina (2005), Natl. Acad. Sci. (2009), Penslar (1995), Shamoo & Resnik (2003), Steneck (2007)]. The ORI has

produced a virtual experience interactive learning simulation program entitled *The Lab: Avoiding Research Misconduct*. In this sophisticated simulation, class participants assume different roles, such as that of graduate student or principal investigator, and are required to decide how to handle possible research misconduct. In my experience, students quickly get emotionally and intellectually involved in dealing with this case of research misconduct. Training is also available for RCR course leaders, such as workshops offered at the Poynter Center for the Study of Ethics and American Institutions at Indiana University in Bloomington.

To determine whether US research institutions are meeting federal government requirements for RCR instruction, Resnik & Dinse (2012) conducted a national survey of 200 such institutions. Of the 144 institutions that responded to the survey, all had an RCR program, which was required either of federally mandated individuals (48%) or of everyone involved in research (52%).

RCR courses usually end with student evaluations. A meta-analysis of 26 evaluations suggested that case-based discussion was preferred over lectures, and stand-alone courses were preferred over embedded ones (Antes et al. 2009). An ORI panel suggested that most of the usual topics in an RCR course were favorably judged; among the ones that were the least so were (*a*) detailed philosophical or historical issues in ethics, (*b*) financial matters, and (*c*) health and safety (**http://www.ori.hhs.gov/rcr-objectives-introduction-0**).

I have taught an RCR course a number of times since 2006. The courses have met for six three-hour weekly sessions and have generally followed the NIH guidelines. Discussion has worked best with 15 participants, usually graduate students along with a few other participants. One very successful variation was to invite two different faculty members to every session. The differences between the faculty members, and between the faculty members and the students, on such topics as authorship, mentoring, and personal use of laboratory facilities were surprising and, I think, instructive to all.

There does not seem to be any good evidence that RCR courses make scientific misconduct less likely to occur or even that the courses change attitudes toward fabrication, falsification, and plagiarism. However, it is my strong feeling, and that of the faculty and student participants, that the vigorous class discussions are valuable in raising consciousness; that is, in increasing sensitivity, empathy, understanding, and awareness on a number of major issues in the life of a research scientist. These issues involve mentoring, authorship, and reviewing as well as the ethical implications of research, conflict of interest, and ownership of data (for a review of studies of the effect of RCR courses, see Kligyte et al. 2008).

The Role of Publication Number

A common response to incidents of scientific misconduct is a call to de-emphasize the quantity of publications—as opposed to their quality—as an important criterion for appointments, promotions, tenure, prizes, and membership in honorific organizations. A number of elite institutions have made gestures in this direction by considering as few as two to five papers (of the applicant's choice) in tenure decisions. Nominations to the National Academy of Sciences require a maximum of ten papers. The emphasis placed on the impact factor of the journal in which an article is published has also been criticized.

Data Recording and Data Sharing

In the past decade, there has been a new development in the handling of data that has been termed "a marriage of word processing and software command scripts" (Clearbout & Karrenback 1992). In this marriage, all data from an experiment, all experimental parameters, and all programs used

to analyze the data are stored in an organized fashion so that, for example, pushing a button attached to a figure will yield access to the data and the analysis and programs that are the basis of the published figure. This idea of "reproducible research" derives from the computational and computer science community, and it is applicable to any branch of science (Wandell et al. 2015; **http://reproducibleresearch.net/bibliography/**). For example, in primate field biology, perhaps on the other end of the dimension from computational neuroscience, Tomasello & Call (2011) have advocated for the recording, in an accessible form, of videos in all primate field studies.

The widespread adoption of reproducible research tools would be valuable for many purposes. For the individual investigator, it would provide a permanent, accessible history of his or her own data for further analysis, communication, and publication. For collaborators, it would provide a common core of information. For members of the laboratory, it would provide the opportunity to find out what their colleagues have been doing as well as a platform for internal laboratory discussion and critique before the presentation and publication stages. For journals, it would provide tools that would enable all authors to submit all data, all experimental parameters, and all programs used to analyze the data for examination by the journal referees and, if the paper is accepted, by the entire scientific community. With modern technology, this process represents a small fraction of the cost of acquiring and analyzing the original data; the cost is comparable to the publication fee or the incremental cost of printing a color figure. Twenty years ago, 132 of 850 journals surveyed had already required deposition of sequence or structure data in a data bank and deposition or sharing of research results upon request (McCain 1995).

Reproducible research practices cannot eliminate outright fabrication and falsification of data. But widespread adoption of these practices would certainly radically reduce the type of scientific misconduct described in this review, and would promote the sharing of methods and data across large and distributed research communities.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

Thanks for generous help from Drs. Ned Feder and Helena Wachslicht-Rodbard and, on data sharing, from Dr. Brian Wandell. Some of the material in this article appeared in Gross (2012).

LITERATURE CITED

Alford CF. 2002. *Whistleblowers: Broken Lives and Organizational Power*. Ithaca, NY: Cornell Univ. Press

Altman LK. 1980. Columbia's medical chief resigns; ex associate's data fraud at issue. *New York Times*, Aug. 9

Antes AL, Murphy ST, Waples EP, Mumford MD, Brown RP, et al. 2009. A meta-analysis of ethics instruction effectiveness in the sciences. *Ethics Behav.* 19:379–402

Babbage C. 1970 (1830). *Reflections on the Decline of Science in England*. New York: Kelley

Bhattacharjee Y. 2013. The mind of a con man. *New York Times Magazine*, April 26

Boffey P. 1986. Major study points to faulty research at two universities. *New York Times*, April 22

Braunwald E. 1987 On analyzing scientific fraud. *Nature* 325:215–16

Braunwald E. 1992. Cardiology: the John Darsee experience. In *Research Fraud in the Behavioral and Biomedical Sciences*, ed. DJ Miller, M Henson, pp. 55–79. New York: Wiley

Broad WJ. 1980a. Imbroglio at Yale (I): emergence of a fraud. *Science* 210(4465):38–41

Broad WJ. 1980b. Imbroglio at Yale (II): a top job lost. *Science* 210(4466):171–73

Broad WJ. 1982a. Harvard delay in reporting fraud. *Science* 215:478–82
Broad WJ. 1982b. Report absolves Harvard in case of fakery. *Science* 215:874–76
Broad WJ. 1983a. Fraud in science taints the high and mighty. *New York Times*, March 20
Broad WJ. 1983b. Notorious Darsee case shakes assumptions about science. *New York Times*, June 14
Broad WJ, Wade N. 1982. *Betrayers of the Truth*. New York: Simon & Schuster
Broadus RN. 1983. An investigation of the validity of bibliographic citations. *J. Assoc. Inf. Sci. Technol.* 34:132–35
Brookes PS. 2014. Internet publicity of data problems in the bioscience literature correlates with enhanced corrective action. *Peer J.* 2:e313
Budd JM, Sievert M, Schultz TR. 1998. Phenomena of retraction: reasons for retraction and citations to the publications. *JAMA* 280:296–97
Casadevall A, Steen RG, Fang FC. 2014. Sources of error in the retracted scientific literature. *FASEB J.* 28:3847–55
Clearbout J, Karrenbach M. 1992. *Electronic Documents Give Reproducible Research a New Meaning*. Stanford, CA: Stanford Explor. Proj. **http://sepwww.stanford.edu/doku.php?id=sep:research:reproducible:seg92**
Couzin J. 2006. Truth and consequences. *Science* 313:1222–26
Culliton BJ. 1983a. Fraud inquiry spreads blame. *Science* 219:937
Culliton BJ. 1983b. Coping with fraud: the Darsee case. *Science* 220:31–35
Culliton BJ. 1983c. NIH to review Emory in Darsee case. *Science* 220:1029
D'Angelo J. 2012. *Ethics in Science*. Boca Raton, FL: CRC Press
Dalton R. 2005. Obesity expert owns up to million-dollar crime. *Nature* 434:424
Eastwood S, Derish P, Leash E, Ordway S. 1996. Ethical issues in biomedical research: perception and practices of postdoctoral research fellows responding to a survey. *Sci. Eng. Ethics* 2:89–114
Eisley L. 1979. *Darwin and the Mysterious Mr. X.* New York: Dutton
Fanelli D. 2009. How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. *PLOS ONE* 4:e5738
Fanelli D. 2013. Why growing retractions are (mostly) a good sign. *PLOS Med.* 10:e1001563
Fang FC, Steen RG, Casadevall A. 2012. Misconduct accounts for the majority of retracted scientific publications. *PNAS* 109:17028–33
Feder N. 2008. *Ethical Problems at NIH—The Struggle Continues*. Washington, DC: Proj. Gov. Overs. **http://www.pogo.org/about/press-room/releases/2008/ph-iis-20081024.html**
Feder N. 2012. *Project on Government Oversight*. Washington, DC: Proj. Gov. Overs. **http://www.pogo.org/about/board-staff/staff-profiles/ned-feder.html**
Franklin A, Edwards AWF, Fairbanks DJ, Hartl DL, Seidenfeld T. 2008. *Ending the Mendel-Fisher Controversy.* Pittsburgh, PA: Univ. Pittsburgh Press
Goodstein DL. 2010. *On Fact and Fraud: Cautionary Tales from the Front Lines of Science*. Princeton, NJ: Princeton Univ. Press
Grieneisen ML, Zhang M. 2012. A comprehensive survey of retracted articles from the scholarly literature. *PLOS ONE* 7:e68397
Gross CG. 2010. Alfred Russell Wallace and the evolution of the human mind. *Neuroscientist* 16:496–507
Gross CG. 2012. Disgrace: on Marc Hauser. *The Nation*, Jan. 9–16
Grossman R. 1993. Machine designed to find plagiarism winds up wrecking inventors' careers. *Chicago Tribune*, May 13
Gunsalus CK. 1998. How to blow the whistle and still have a career afterwards. *Sci. Eng. Ethics* 4:51–63
Hilts PJ. 1991. Nobelist apologizes for defending research paper with faulty data. *New York Times*, May 4
Hilts PJ. 1993a. Institutes of Health close fraud investigation unit. *New York Times*, May 5
Hilts PJ. 1993b. Ideas & Trends: why whistle-blowers can seem a little crazy. *New York Times*, June 13
Hixson J. 1976. *The Patchwork Mouse*. New York: Doubleday
Hoke F. 1993. HHS pressed to reverse whistleblowers' reassignment. *The Scientist*, June 28
Hoke F. 1995a. On their own: Stewart and Feder persist with misconduct inquiries. *The Scientist*, Feb. 6
Hoke F. 1995b. Veteran whistleblowers advise other would-be "ethical resisters" to carefully weigh personal consequences before taking action. *The Scientist*, May 5

Hunt M. 1981. A fraud that shook the world of science. *New York Times Magazine*, Nov. 1
Judson HF. 2004. *The Great Betrayal: Fraud in Science*. Orlando, FL: Harcourt
Kalichman MW, Friedman PJ. 1992. A pilot study of biomedical trainees' perceptions concerning research ethics. *Acad. Med.* 67:769–75
Kevles DJ. 1998. *The Baltimore Case: A Trial of Politics, Science, and Character*. New York: Norton
Kligyte V, Marcy RT, Sevier ST, Godfrey ES, Mumford MD. 2008. A qualitative approach to Responsible Conduct of Research (RCR) training development: identification of metacognitive strategies. *Sci. Eng. Ethics* 14:3–31
Knox R. 1983. The Harvard fraud case: Where does the problem lie? *JAMA* 249:1797–99, 1802–7
Kochan CA, Budd JM. 1992. The persistence of fraud in the literature: the Darsee case. *J. Am. Soc. Inf. Sci.* 43:488–93
Kohn A. 1986. *False Prophets*. Oxford, UK: Blackwell
Lock S, Wells F, Farthing F. 2001. *Fraud and Misconduct in Biomedical Research*. London: BMJ Books. 3rd ed.
Lubalin JS, Ardini ME, Matheson JL. 1995. *Consequences of Whistleblowing for the Whistleblower in Misconduct in Science Cases*. Washington, DC: Res. Triangle Inst. **https://ori.hhs.gov/sites/default/files/final.pdf**
Macrina FL. 2005. *Scientific Integrity*. Herndon, VA: ASM Press
Madlock-Brown CR, Eichmann D. 2015. The (lack of) impact of retraction on citation networks. *Sci. Eng. Ethics* 21:127–37
Margraf J. 2015. Zur Lage der Psychologie. *Psychol. Rundsch.* 66:1–30
Martin B. 2013. *Whistleblowing: A Practical Guide*. Sparsnäs, Sweden: Irene Publ.
McCain K. 1995. Mandating sharing: journal policies in the natural sciences. *Sci. Comm.* 16:403–31
Merton RK. 1957. Priorities in scientific discovery: a chapter in the sociology of science. *Soc. Rev.* 22:635–59
Mervis J. 1986. Study sharpens debate on role of co-authors. *The Scientist*, Nov. 17
Monastersky R. 2008. NIH punished scientist who had called for open records. *The Chronicle of Higher Education*, Oct. 22
Moran NC. 1985. Report of ad hoc committee to evaluate research of Dr. John R. Darsee at Emory University. *Minerva* 23:276–305
Nath SB, Marcus SC, Druss BG. 2006. Retractions in the research literature: misconduct or mistakes? *Med. J. Aust.* 185:152–54
Natl. Acad. Sci. 2009. *On Being a Scientist: A Guide to Responsible Conduct in Research*. Washington, DC: Natl. Acad. Press
Neugebauer O. 1875. *A History of Mathematical Astronomy*. Berlin: Springer
Newton RR. 1977. *The Crime of Claudius Ptolemy*. Baltimore, MD: Johns Hopkins Univ. Press
NIH (Natl. Inst. Health). 2009. *Update on the Requirement for Instruction in the Responsible Conduct of Research*. Bethesda, MD: NIH. **http://grants.nih.gov/grants/guide/notice-files/NOT-OD-10-019.html**
NSF (Natl. Sci. Found.). 1996. *Research Misconduct*. Arlington, VA: NSF. **https://www.nsf.gov/oig/resmisreg.pdf**
NSF (Natl. Sci. Found.). 2009. *Responsible Conduct of Research* (*RCR*). Arlington, VA: NSF. **http://www.nsf.gov/bfa/dias/policy/rcr.jsp**
O'Toole M. 1991. The whistle-blower and the train wreck. *New York Times*, April 12
O'Toole M. 1993. Silencing the dissenters: an unwise, ignominious move. *The Scientist*, May 17
ORI (Off. Res. Integr.). 1995. *ORI Guidelines for Institutions and Whistleblowers: Responding to Possible Retaliation Against Whistleblowers in Extramural Research*. Rockville, MD: US Dep. Health Human Serv. **http://ori.hhs.gov/guidelines-whistleblowers**
ORI (Off. Res. Integr.). 2011. *Definition of Research Misconduct*. Rockville, MD: US Dep. Health Human Serv. **http://ori.hhs.gov/definition-misconduct**
Penslar RL. 1995. *Research Ethics*. Bloomington: Indiana Univ. Press
Powledge T. 1987. Stewart-Feder (finally) in print. *The Scientist*, Feb. 9
Ranstam J, Buyse M, George SL, Evans S, Geller NL, et al. 2000. Fraud in medical research: an international survey of biostatisticians. *Control. Clin. Trials* 21:415–27
Rennie D, Gunsalus CK. 2008. What is research misconduct? In *Fraud and Misconduct in Biomedical Research*, ed. S Lock, F Wells, F Farthing, pp. 29–52. London: BMJ Books. 4th ed.

Resnik DB, Dinse GE. 2012. Do U.S. research institutions meet or exceed federal mandates for instruction in responsible conduct of research? A national survey. *Acad. Med.* 87:1237–42
Rivlin S. 2004. *Scientific Misconduct and Its Cover-Up*. Boca Raton, FL: BrownWalker
Sarasohn J. 1993. *Science on Trial*. New York: St. Martin's Press
Servick K. 2014. Researcher files lawsuit over anonymous PubPeer comments. *Science*. **http://news.sciencemag.org/scientific-community/2014/10/researcher-files-lawsuit-over-anonymous-pubpeer-comments**
Shamoo AE, Resnik DB. 2003. *Responsible Conduct of Research*. New York: Oxford Univ. Press
Siskin M, Roychowdhury V. 2006. Do you sincerely want to be cited? Or: Read before you cite. *Significance* 3:179–81
Sokal AD. 1996. Transgressing the boundaries: toward a transformative hermeneutics of quantum gravity. *Soc. Text* 46/:217–52
Steen RG. 2011. Retractions in the scientific literature: Do authors deliberately commit research fraud? *J. Med. Ethics* 37:113–17
Steen RG, Casadevall A, Fang FC. 2013. Why has the number of scientific retractions increased? *PLOS ONE* 8:e68397
Steneck NH. 2007. *ORI Introduction to the Responsible Conduct of Research*. Washington, DC: US Gov. Print. Off.
Stewart WW, Feder N. 1987. The integrity of the scientific literature. *Nature* 325:207–14
Swazey JP, Anderson MS, Louis KS. 1993. Ethical problems in academic research. *Am. Sci.* 81:542–53
Titus SL, Wells JA, Rhoades LJ. 2008. Repairing research integrity. *Nature* 19:453:980–82
Tomasello M, Call J. 2011. Methodological challenges in the study of primate cognition. *Science* 334:1227–28
US House Rep. 1981. Fraud in biomedical research. Hearings before Committee on Science and Technology. *Subcomm. Investig. Overs., 97th Congr.*, 1st sess. Washington, DC: US Gov. Print. Off.
Wade N. 1981. A diversion of the quest for truth. *Science* 211:1022–25
Wade N. 1988. Ideas & Trends: looking hard at science's self-scrutiny. *New York Times*, Aug. 21
Wadman M. 2008. Department of beams in the eye. *Nature News Blog*, Oct. 23. **http://blogs.nature.com/news/2008/10/department_of_beams_in_the_eye.html**
Wagner E, Williams P. 2011. Why and how do journals retract articles? An analysis of Medline retractions 1988–2008. *J. Med. Ethics* 37:567–70
Wandell BA, Rokem A, Perry LM, Schaefer G, Dougherty RF. 2015. Data management to support reproducible research. arXiv:1502.06900
Weiner JS. 1955. *The Piltdown Forgery*. London: Oxford
Wells JA. 2008. *Final Report: Observing and Reporting Suspected Misconduct in Biomedical Research*. Washington, DC: Gallup Org. **http://ori.hhs.gov/sites/default/files/gallup_finalreport.pdf**
Westfall RS. 1973. Newton and the fudge factor. *Science* 179:751–58

Annual Review of Psychology

Volume 67, 2016

Contents