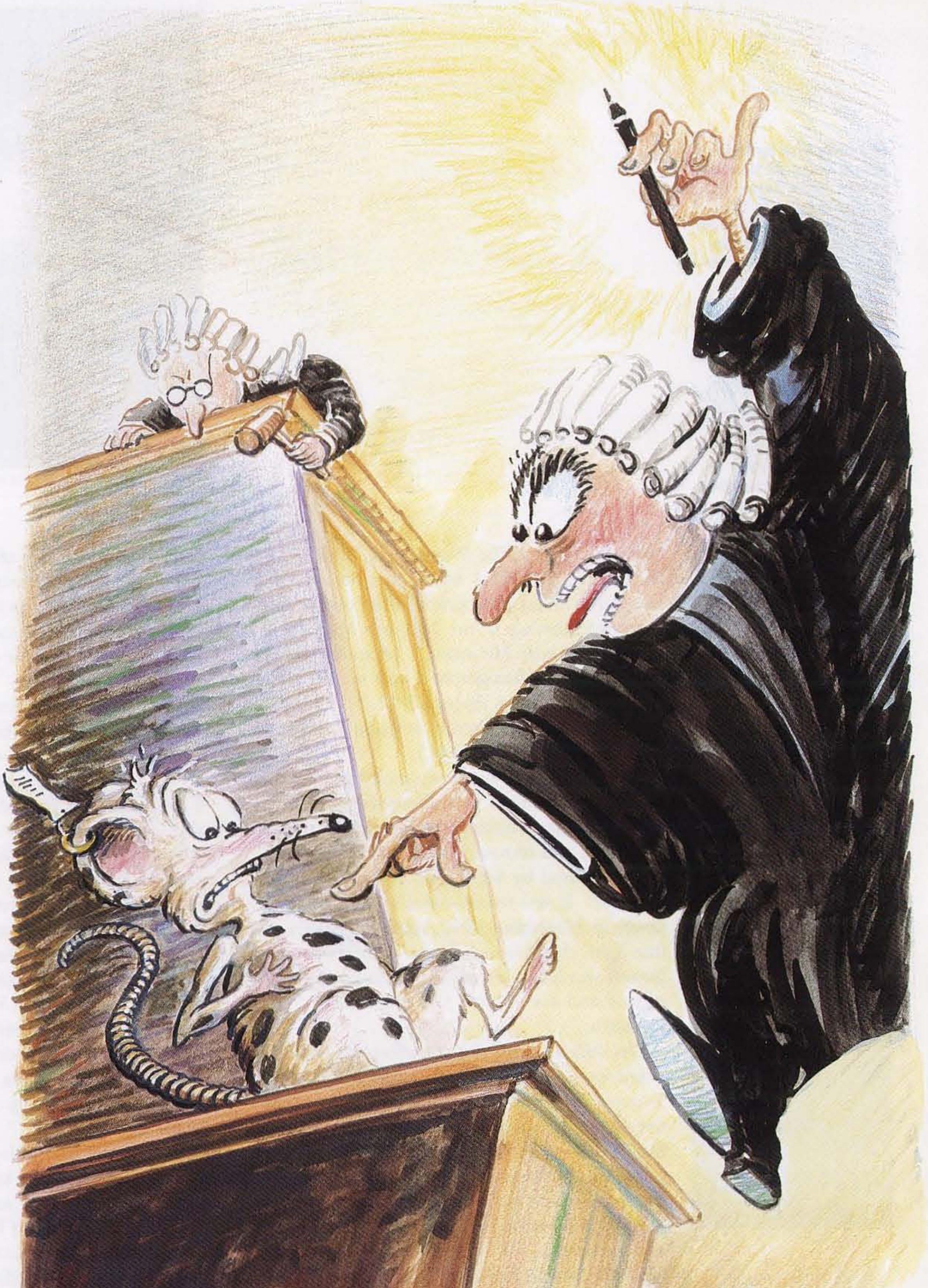

©CATHY HILL 1991

Scientific Fraud

One of the reasons that nobody knows the exact extent of scientific fraud is that nobody knows exactly what scientific fraud is.

by David Goodstein

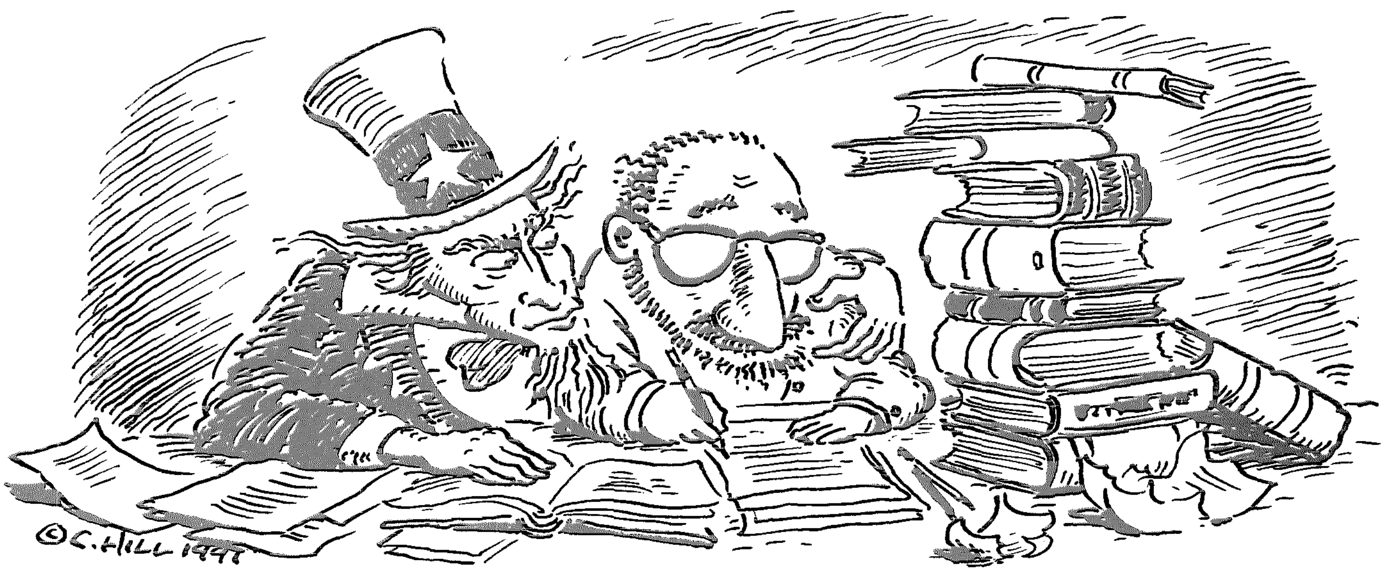
Most scientists have traditionally believed scientific fraud to be rare or nonexistent. Nevertheless, it has recently become a very hot topic. And there certainly have been some well-documented cases in the past. Perhaps the most famous incident of scientific fraud in this century was the case of Piltdown man—a human cranium and ape jaw that were found in a gravel pit in England in 1908 and 1912. Substantial academic reputations were made by discerning human characteristics in the jaw and ape characteristics in the cranium. However, this missing link was exposed as a fake in 1954. Another famous case was that of Sir Cyril Burt, a psychologist who worked on the heritability of intelligence by studying identical twins who were separated at birth and brought up in different environments. Unfortunately, there were very few cases of such convenient subjects for research, so Burt obligingly invented 33 more and further helped matters along by inventing two assistants to help him study them. Burt died in 1971, but his hoax was not discovered until 1974.

In 1974 William Summerlin was doing research at the Sloan-Kettering Institute in Minnesota that required nature to produce for him some rats with black patches on their skin. Since nature was not sufficiently cooperative, he helped her along with a black, felt-tip pen and was caught in the act. In another case, John Darsee, a brilliant young cardiologist at Harvard Medical School, was producing approximately a hundred papers a year. Until he was caught red-handed fabricating data in 1981, it didn't occur to anyone that with that rate of production maybe he didn't have time to do the actual experiments. In yet another case, Stephen Breuning made headlines in 1987, when it was revealed that he had fabricated data in his research at the University of Pittsburgh on the effects of psychoactive drugs in children.

The most recent notorious case involves a paper in which Nobel-prizewinning biologist David Baltimore, now president of Rockefeller University, was one of the authors. A postdoc in the group, Margot O'Toole, without accusing anybody of fraud, claimed that the evidence did not support the conclusions in the paper. The particular work under criticism was actually done by one of Baltimore's collaborators, Thereza Imanishi-Kari, but because of his name on the paper, the case attracted only slightly less journalistic attention than the Persian Gulf situation.

I started to become personally more involved with fraud about three years ago when, as the new vice provost, I had to dig through the avalanche of paper on my desk that reported on what was going on in Washington. As I read some of that material, it started to become obvious that Caltech was going to be forced to have a set of formal regulations on what to do in the unthinkable event of scientific fraud. So, in order to prepare myself, I started to collect information on fraud. I now have a file that fills a whole file drawer, and Caltech now has regulations on scientific fraud. The file tells, among other things, the political history of this issue.

The first serious congressional attention to the problem seems to have been in 1981, when the investigations subcommittee of the House Com-



This was something beyond their understanding and they should keep their grubby hands out of it.

mittee on Science and Technology was prompted to look into a Harvard Medical School case. Albert Gore, then representative and now Democratic senator from Tennessee, was chairman of the committee. Philip Handler, then president of the National Academy of Sciences, made a presentation to the committee in which he told them pretty much what most scientists today would say to most congressmen—that this was something beyond their understanding and they should keep their grubby hands out of it. This was not exactly well received by Congress, which felt that the scientists, after all, were being supported by the public and ought to accept congressional oversight. Nevertheless, these hearings did not lead to any congressional action. During the early eighties, Orrin Hatch, Republican senator from Utah, started poking into the National Cancer Institute, also without permanent effect. But more recently, and with greater publicity, two Democratic representatives, John Dingell of Michigan and Ted Weiss of New York, tried to get into the act of investigating the Baltimore case (and conceivably benefiting from it politically) by holding hearings in their respective subcommittees.

Dingell had succeeded Gore as chairman of the investigations subcommittee of the Science and Technology Committee, and Weiss was head of the subcommittee on human resources and intergovernmental relations of the Government Operations Committee. In April 1988 these two competed in a somewhat unseemly race to be the first one to hold hearings. Dingell's hearings were to lead to a much-discussed report that has not yet appeared at this writing. Just last October Weiss's committee issued a booklet containing an analysis of 10 cases of scientific fraud, entitled "Are Scientific Misconduct and Conflict of Interest Hazardous to Our Health?" The title says a lot about the slant of the booklet, which is especially critical of the universities for their handling of these cases. The committee report was not well received in the press, which pointed out that it was based largely on an analysis of cases that had occurred in the early 1980s. Much has happened since then, and the universities have improved in their handling of fraud cases, so that the report is by now largely irrelevant. It seems to have dropped out of sight.

Meanwhile, at the National Institutes of Health, a couple of biologists named Ned Feder and Walter Stewart have set themselves up as a kind of self-appointed truth squad. According to their critics they had not been very productive biologists and were trying to find a way of holding on to their laboratory and office space. They hit upon the fraud issue and were particularly visible in the Baltimore case. In many other cases too, they have become the lightning rod for whistle blowers. Anyone can call to report an instance of scientific fraud. These two now have official permission from their superiors to spend a certain percentage of their time pursuing wrongdoers.

In 1988 and 1989 the National Institutes of Health (NIH) and the National Science Foundation (NSF) each published in the *Federal Register* formal sets of regulations regarding scientific

In tort law, proving fraud is quite a different matter from what we regard as sufficient indication of fraud in science.

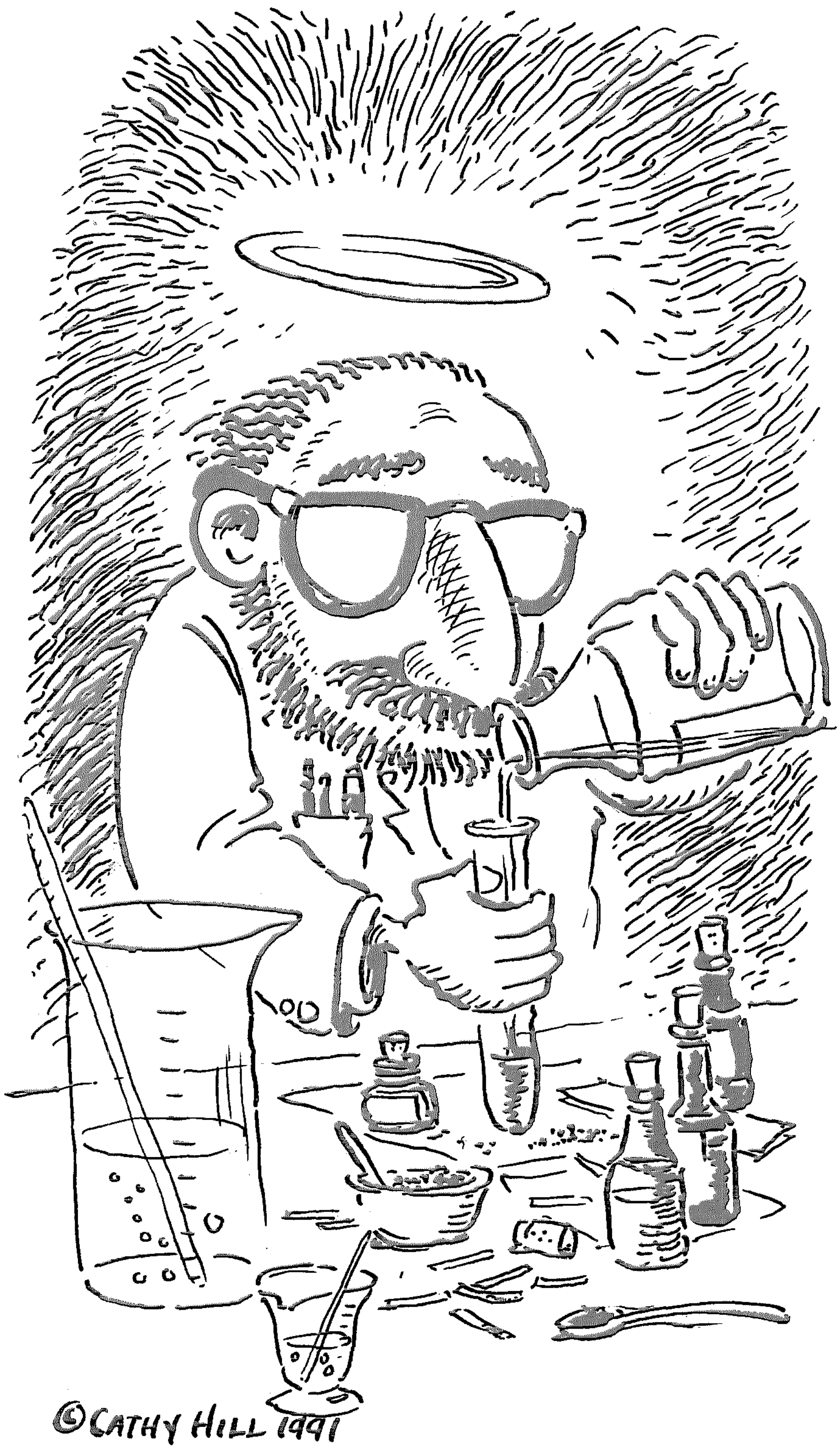
fraud. These two sets of regulations, many pages long, are virtually identical. Both of them call on the university (if the fraud has been committed at a university) to investigate the situation first and only later to hand it over to the agency. A rule was declared in late 1989 by the Public Health Service, the parent organization of the NIH, stating that after January 1990 no research proposal would be accepted from any university that did not certify that it had in place a formal set of regulations on how to handle research fraud. That was the point at which it became necessary for Caltech to have such regulations. An Office of Scientific Integrity has been established within the NIH. The very name calls up images of "1984." (1984 is now a date in the past, but it was once a date in the future.) The NSF doesn't yet have such an office, but it has an Inspector General who seems to serve much the same function. These entities are concerned with fraud, misconduct, and conflict of interest—three types of misbehavior that may not always be so easily distinguishable.

One question any thoughtful person must ask is: How common is scientific fraud? How often does it happen? Is it something that's so rare we shouldn't worry about it? Or is it really quite common and a major threat to the scientific enterprise? One of the reasons that nobody knows the exact extent of scientific fraud is that nobody knows exactly what scientific fraud is. What do we mean by the phrase? For an answer, we turn first to the most authoritative possible source—the Caltech regulations. They define science fraud or research fraud as "serious misconduct with intent to deceive, for example, faking data, plagiarism, or misappropriation of ideas." That's a clear definition. Barbara Mishkin, a Washington lawyer often quoted in this context, has listed three types of scientific misconduct: 1) knowing misrepresentation of data, procedures, or analysis; 2) plagiarism and other authorship misdeeds, such as guest authorship and the like (guest authorship means putting the boss's name on the paper even though he didn't really do any research); and 3) outright violation of laws, such as laws regarding human subjects, recombinant DNA, and so on. The Caltech regulations address the first two of these but explicitly rule out the third as not coming under their jurisdiction. If you violate a law—for example on the handling of human subjects—there are means and procedures already in place for dealing with that.

Personally, I don't think these definitions cover the whole map. In my 25 years as a working scientist, by far the most serious instances of misconduct that I have seen at first hand in my own field have come in the arena of anonymous reviews of journal articles and research proposals. This type of thing is never mentioned at all by anyone who deals with the subject of scientific misconduct—the lawyers, the philosophers, or the sociologists. But they're not the scientists in the trenches. Seen from my own narrow trench in physics, that's where you find the misconduct.

In tort law, proving fraud is quite a different matter from what we regard as sufficient indication of fraud in science. First of all, the law envisions a plaintiff and a defendant; someone has to bring the case to court. In order to prevail, the plaintiff must prove five points: 1) that a false representation was made—in other words, that the defendant cheated; 2) that the defendant knew it was false (or recklessly disregarded whether it was); 3) that there was intent to induce belief in this misrepresentation; 4) that there was reasonable belief on the part of the plaintiff; and 5) that there was resulting damage.

In science fraud nobody pays attention to the fourth and fifth points—that there was reasonable belief and actual damage. Nobody pretends that we have to prove that in order to, in effect, convict someone of research fraud. The Caltech regulations, which spell out "serious misconduct with intent to deceive" and so on, seem to encompass the first three: false representation, knowledge that it's false, and intent to induce belief. But I think that a clever lawyer taking on a real case of fraud, such as some of the recent examples I've cited, could argue that there



The Noble Scientist is somehow supposed to be more virtuous and upright than ordinary people and therefore can be expected not to misbehave even in the smallest way.

was not false representation in the ordinary sense. In most cases of science fraud, the person committing the fraud was not trying to perpetuate an untruth, but rather was trying to help along what he believed to be the truth. I'll get back to this point later on, when I discuss some specific historical cases.

In any case, the barrier against proving fraud is much higher in legal precedent than it is in the standards we apply to scientists. Now, we scientists of course tend to be arrogant; we think we know what is right. But the law has a great deal of experience with the behavior of real people, and since science is a very human activity, perhaps we have something to learn from lawyers about the standards for proving such serious allegations as fraud. The five legal points necessary to prove fraud are based on long experience with the way people really behave, whereas the idea of what constitutes science fraud is based on what I call the Myth of the Noble Scientist. The Noble Scientist is somehow supposed to be more virtuous and upright than ordinary people and therefore can be expected not to misbehave even in the smallest way. This myth only makes us more vulnerable to misunderstanding what we do and what actually constitutes fraud. The effects of this can be seen in an analysis of journalistic accounts of fraud in science.

Betrayers of the Truth, published in 1982 by Simon and Schuster, was written by William Broad and Nicholas Wade. Both were reporters for *Science* magazine, and Wade is now on the editorial board of *The New York Times*—hardly schlock journalists. Rather than try to analyze

Newton's theory was so good he was able to calculate the speed of sound and then compare it with measurements. When he did, they disagreed by about 10 percent.

Among the other scientists they accuse of being "Known or Suspected Cases of Fraud" are Galileo, Newton, Dalton, Mendel, Millikan, and quite a number of others.

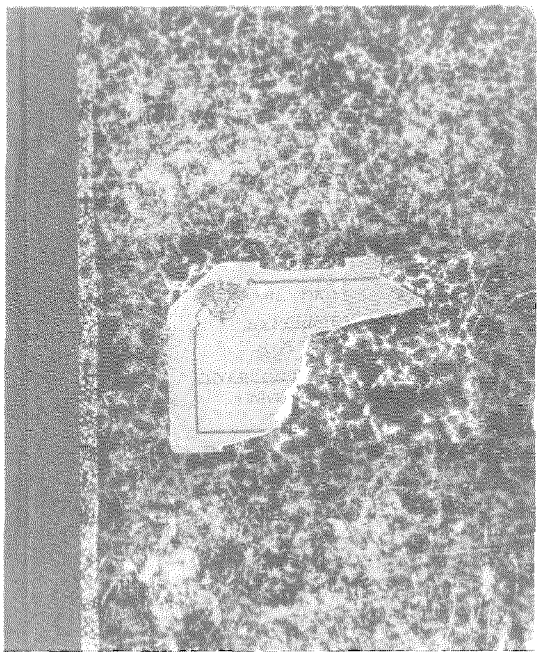
my own file drawer full of newspaper clippings, it makes sense to take this book as an example of a serious study of science fraud by the best journalists, since they understand science better than others of their profession and are probably more dependable in what they write.

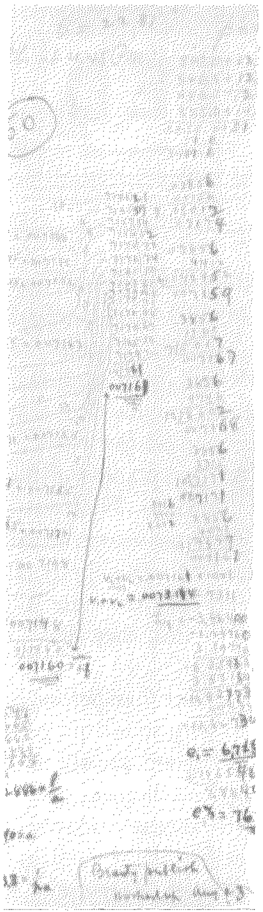
The book has an appendix entitled "Known or Suspected Cases of Scientific Fraud," which includes the case of Claudius Ptolemy, the Alexandrian astronomer of the second century A.D. who wrote the *Almagest*, upon which all of astronomy was based until the time of Copernicus. Broad and Wade claim that Ptolemy committed fraud because he could not possibly have made the astronomical observations he claimed he made. By techniques of archaeoastronomy—using knowledge of how the sky works to run it backwards to see what the sky looked like at a particular time in the past—researchers have found that the observations Ptolemy reports were not made in Alexandria in the second century A.D., but rather, at the latitude of Rhodes in the second century B.C. So they concluded that the actual readings were taken by Hipparchus of Rhodes.

Another person on the list of "Known or Suspected Cases of Scientific Fraud" is Hipparchus of Rhodes, whose observations, Broad and Wade say, were actually made by the ancient Babylonians. The authors make no comment about this impossible contradiction. Both accusations cannot be correct. But obviously they hold themselves to a less stringent standard than they apply to scientists.

Among the other scientists they accuse of being "Known or Suspected Cases of Fraud" are Galileo, Newton, Dalton, Mendel, Millikan, and quite a number of others. I'm not personally familiar with the case of Mendel, who studied the genetics of peas and came up with data that some people have said are too good to be true, but I am familiar at first hand with some of the others—for example, Isaac Newton. Newton explained the propagation of sound waves in air. Newton's theory was so good he was able to calculate the speed of sound and then compare it with measurements. When he did, they disagreed by about 10 percent.

Now, you have to understand that before this, there was no idea at all why sound propagates in air, and to have calculated the speed within 10 percent was a huge intellectual triumph. Nevertheless, the 10-percent discrepancy bothered Newton, and so he set out to explain the difference. The real explanation for the difference has turned out to be that sound is adiabatic, and Newton had assumed implicitly that it was, instead, isothermal. In other words, in a sound wave there's heating and cooling that pushes the sound along a little faster than it would otherwise go. Newton didn't take account of that effect, so he calculated the speed that sound would have if it were all at one temperature. That subtle difference would not be understood for another 200 years, so you certainly can't blame Newton for not knowing it. But because he was disturbed that his theory didn't quite correspond to the observation, he tried to cook up some explanation for the discrepancy. He came up with all kinds of things





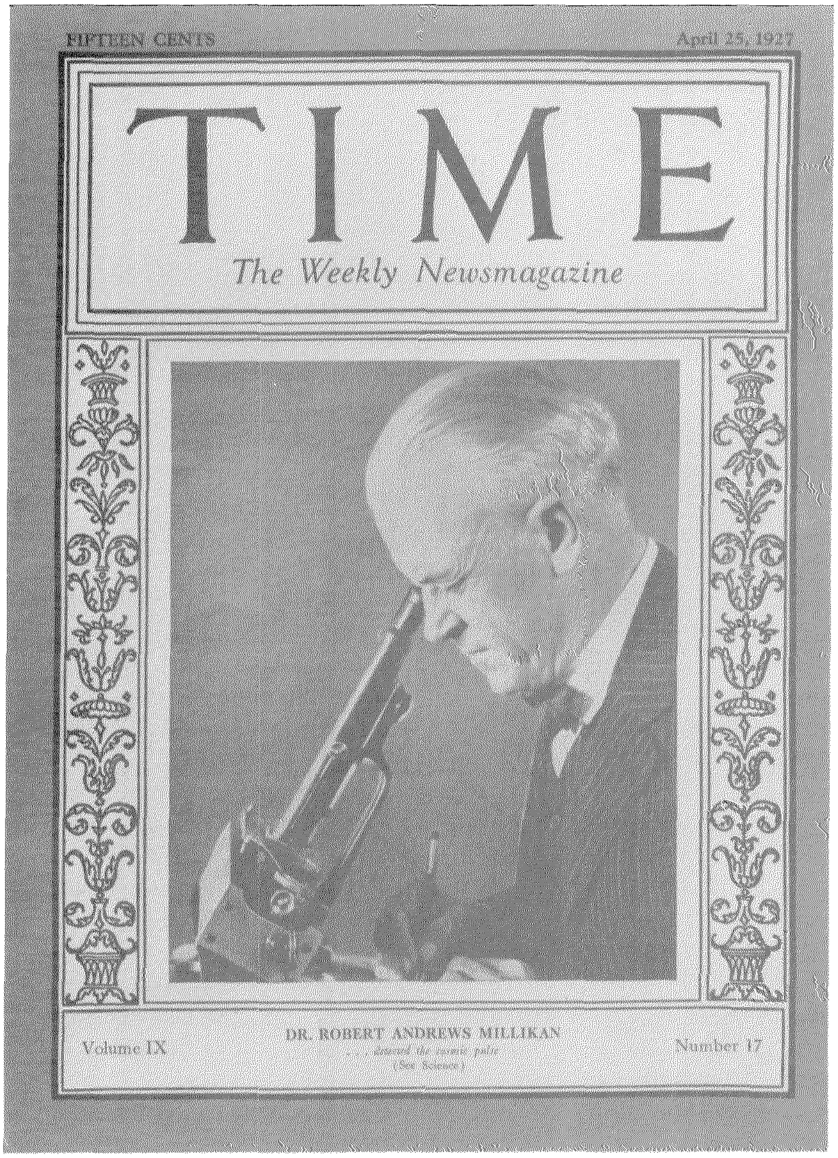
Millikan's lab notebooks for his Nobel prizewinning oil drop experiment contain several examples of his comments about how publishable the data are. The page portion at left also notes that he decided it was a "beauty" on August 23; the initial observations were made on March 30, 1912.

that sound hilarious to us now: the water vapor in the air didn't participate, he had ignored the space taken up by the molecules of air, and other things like that. He made little fixes until he finally got the theory in agreement with the experiment. It's the sort of thing that every theorist does today; if you have a theory that doesn't quite agree with the experiment, you speculate on what might cause the small discrepancy. That's exactly what Newton was doing. This is an example of what these two journalists regard as fraud. In hindsight Newton's fixes are funny; it's the way people really act. But fraud? No, it's not fraud.

Another example Broad and Wade give is our own sainted Robert Andrews Millikan. The accusation is based on notations in his laboratory notebooks, which we have in our Archives at Caltech, and which I have read and shown to my students. In fact, I show these notations every year precisely because they are instructive; they tell you something about the real world. Millikan was measuring the electric charges of oil drops; he wanted to prove that the electron charge came in definite units—that it was quantized—and then he wanted to measure what that unit was. He had actually already made his preliminary measurements, and he knew very nearly what the answer was. Millikan had a rival, Felix Ehrenhaft, who believed that electric charge was a continuous quantity rather than quantized. Ehrenhaft criticized Millikan's results, so Millikan went back to the laboratory to get better data to have ammunition against Ehrenhaft. Later on he published a paper in *Physical Review* in which he says (roughly): "I've published every piece of scientific data I got without bias; I have looked at 60 drops and here are all 60 drops," or something like that.

But when you look through his notebooks, it appears a bit different. Each page has notations on one drop. Millikan would spend a whole evening watching one drop go up and down in his electric field, measuring its speed, taking down data, making calculations, getting the result for the charge. He knew, of course, what result he expected. So in some cases he would write in red (everything else is black), "Beauty—Publish," or "One of the best I've ever had—Publish." And then on one page he wrote, "Very low—something wrong." And you know that that one did not get published, in spite of the fact that he said he published everything.

What's happening is that he has some idea of what he expects, and when he gets the wrong result it's a clue that something is wrong. But he doesn't just throw it out because he doesn't like it; he examines his experiment to figure out what mistake he's made, and when he finds the mistake, it is duly noted on the page ("distance wrong" he wrote on that particular page). People make mistakes; experiments are always pushing the limits of the possible. Scientists are always at the state of the art, and we make mistakes in the laboratory all the time. If everybody were obliged to publish every mistake, the scientific literature would be so full of garbage that you wouldn't be able to read it. It's bad enough as it is. What Millikan was doing was perfectly legitimate: he would examine the



Caltech's own Noble (and Nobel) Scientist makes TIME's cover in 1927.

If everybody were obliged to publish every mistake, the scientific literature would be so full of garbage that you wouldn't be able to read it. It's bad enough as it is.

"wrong" result and would find that he had made a mistake and conclude that that result had to be tossed out. Of course, he did not try quite so hard to find some reason for throwing away the results that were "right." That's really the point where bias enters his result. This kind of bias is built into all scientific research. Even though we take elaborate precautions—such as double-blind tests—to try to avoid this kind of unconscious bias, it still creeps into scientific results. But to call it fraud, as Broad and Wade do, is absolutely irresponsible. Millikan is merely another casualty here of the Myth of the Noble Scientist, which ignores the dynamics of the way real scientists work.

Millikan and Newton weren't guilty of fraud, but clearly some others are. Who are they? Patricia Woolf, a Princeton University sociologist, did a study of 26 cases of serious scientific misconduct that surfaced in one way or another between 1980 and 1987. It turned out that, of these 26 cases, two were in chemistry and biochemistry, one was in physiology, two were in psychology, and 21 were in biomedical sciences. Furthermore, of the 26, some 17 were committed by MDs rather than PhDs. So the conclusion is inescapable that scientific fraud is essentially biomedical fraud, at least in recent times. The $64,000 question is: Why is that true?

One reason some have suggested is that there is more money in biomedical sciences, and money corrupts. Fraud in the form of plagiarism, however, is not unusual in fields such as history, where there is very little money to be found. So it seems to me that money is not the principal motivating force. I believe that career pressure is more important. In every case of science fraud I've looked at, somebody was advancing a career rather than seeking money. Other people have suggested that since the large majority of fraud perpetrators are MDs rather than PhDs, perhaps it's because medical doctors have a different sort of ethic from scientists—doctors care about the health of the patient rather than pure scientific truth. Being brought up in this ethic might give you a different attitude toward what's permissible and what's not. This is a subtle argument, and I don't know whether there's anything in it or not.

I used to have a theory that had to do with the reproducibility of results. In physics, and in other fields where there is little fraud, people believe that experiments are precisely reproducible, in the sense that if somebody else goes into the laboratory and does the same experiment, they'll get the same results. Now, every experimentalist knows that this is not true. Real experiments are too hard for that to be the case, but the whole field is pervaded by the idea that things are causally related in a relatively straightforward way, and therefore reproducible. So it would be foolish for me to fake a data point, because somebody else will repeat the experiment and find the data point in a different place.

Going back to the rivalry between Millikan and Ehrenhaft, this is what kept Millikan from being too cavalier and just keeping good results. He knew that if he got it wrong, his rival would bite his head off without any hesitation at all. So perhaps physicists are less likely to fake than scientists doing experiments in biology or biomedicine, where "truth" is more statistical, rather than causal or precise. I might feel in those fields that if I cheat a little bit, nobody's ever going to find out because my cheating will be within the range of uncertainty of the data.

This is what I thought before I started looking at some of these cases of fraud. What I found instead was that in every single case the person who perpetrated the fraud thought he knew the answer. That's quite different from feeling that you're in an imprecise field where things are not very reproducible. These scientists really thought that they knew what the answer was, and that by faking the data all they were doing was helping things along a little bit. They weren't perpetrating a false result; they were just taking a bit of a shortcut—leaving out some steps that weren't really necessary because they knew what the answer was. You can see that in the case of Cyril Burt, the psychologist who faked data on identical twins. He *knew* that intelligence was inherited, and to go out and find 33 more sets of identical twins that had been separated at birth would be impossibly difficult. And it was really unnecessary because he knew what the answer would be if he went through all that work. So why go through all that work, right?

And everybody knew that God is an Englishman. If they had been discovered in those other places, there had *to be prehistoric human remains in England.*

You can see it even in the case of Piltdown man. By 1912 prehistoric human remains had been discovered in France and Germany, and there was some indication that there might *even* be some in Africa. And everybody knew that God is an Englishman. If they had been discovered in those other places, there *had* to be prehistoric human remains in England. It was only a matter of helping things along a little bit.

There are many mysteries about this whole business, and perhaps all of these factors play a role. It seems clear that scientists are most vulnerable to cheating or cutting corners when: 1) they are under career pressure to produce something; 2) they think they know what the answer is and feel that actually going to the trouble of taking the data just slows down the inevitable process; and 3) they think they are somewhat protected by "soft reproducibility."

There is no human activity that can stand up to the glare of relentless, absolute honesty. We all build little hypocrisies into what we do in order to make life a little bit easier to live. Because science is a very human activity, there are hypocrisies and misrepresentations built into the way we do it. For example, every scientific paper is written as if that particular investigation were a triumphant procession from one truth to

another. All of us who actually work in the trenches, however, know that every scientific experiment is chaotic—like war. You never know what's going on; you can't understand what the data mean. But in the end you figure out what it was all about and then, with hindsight, you write it up describing it as one clear and certain step after the other. This is a kind of hypocrisy, but it's deeply embedded in the way we do science. We're so accustomed to it that we don't even regard it as a misrepresentation anymore.

The wry phrase "typical best case," for example, describes the routine procedure of saying that the data are typical, but presenting the best set of data that were produced. Everybody does this, and everybody recognizes that it's what everybody does. It's regarded as acceptable behavior; it's not considered fraud. There's an important distinction here: if I present my best case as typical, that's acceptable. But if I take those data and move one data point to make it look a little bit prettier, that's fraud. Scientists do recognize the difference. There's something sacrosanct about data; there's a hard line there that cannot be crossed.

'Owing to difficulties in sample handling' really means something like 'we dropped it on the floor.'

Glossaries explaining the real meanings of terms found in scientific papers occasionally make the rounds of the trenches. For example, "owing to difficulties in sample handling" really means something like "we dropped it on the floor." This only recognizes that scientific papers may disguise what really happened, even though they are supposed to present things in a rigorously honest way. We don't hold classes in the rules of misrepresentation in scientific papers, but the apprenticeship that one goes through to become a scientist does involve learning them. That same apprenticeship, however, also inculcates a deep respect for the inviolability of scientific data. It teaches how one distinguishes the indelible line that separates harmless fudging from real fraud.

I believe that scientists are basically honest, even if they don't quite live up to the Myth of the Noble Scientist. Cases such as those of Summerlin and Darsee shocked every scientist I know. Although I've said here that we might have a thing or two to learn from the lawyers, I don't mean that we should go the whole legal route and insist on proving those five elements to demonstrate fraud. If someone has cheated on scientific data, we should regard that as fraud without having to prove anything else. Nevertheless, I think that the Myth of the Noble Scientist does not serve us well. Scientists are fallible human beings. So are congressmen and journalists. We could all benefit from just a little more understanding and honesty about what we really do, and how and why we do it. □

David Goodstein discussed scientific fraud at the November 28, 1990 session of a regular series of informal seminars on Science, Ethics, and Public Policy. As vice provost since 1987, Goodstein was chiefly responsible for drafting Caltech's regulations on scientific fraud, a document he is proud of. And if the above article, which is based on his seminar remarks, seems to tilt just a bit toward the physicist's point of view, it's because Goodstein is also professor of physics and applied physics (and creator of the prizewinning TV physics course, The Mechanical Universe*). Goodstein has been a member of the Caltech faculty since 1966. He earned his BS from Brooklyn College and PhD from the University of Washington.*