

Difficulties in Defining Errors in Case Against Harvard Researcher

By Nicholas Wade

Oct. 25, 2010

The still unresolved case of Marc Hauser, the researcher accused by Harvard of scientific misconduct, points to the painful slowness of the government-university procedure for resolving such charges. It also underscores the difficulty of defining error in a field like animal cognition where inconsistent results are common.

The case is unusual because Dr. Hauser is such a prominent researcher in his field, and is known to a wider audience through his writings on morality. There seemed little doubt of the seriousness of the case when Harvard announced on Aug. 20 that he had been found solely responsible for eight counts of scientific misconduct.

But last month two former colleagues, Bert Vaux and Jeffrey Watumull, both now at the University of Cambridge in England, wrote in the Harvard Crimson of Dr. Hauser's "unimpeachable scientific integrity" and charged that his critics were "scholars known to be virulently opposed to his research program."

Also last month his principal accuser outside of Harvard, Gerry Altmann, allowed that he may have spoken too hastily. Dr. Altmann is the editor of *Cognition*, a psychology journal in which Dr. Hauser published an article said by Harvard to show scientific misconduct.

When first shown evidence by Harvard for this conclusion, Dr. Altmann publicly accused Dr. Hauser of fabricating data. But he now says an innocent explanation, based on laboratory error, not fraud, is possible. People should step back, he writes, and "allow due process to conclude."

Due process, in this case, includes an independent inquiry by the Office of Research Integrity, a government agency that investigates scientific misconduct. Its inquiries take seven months on average, ranging up to eight years, says John Dahlberg, director of the agency's investigations unit.

Under Harvard's faculty policy, the university cannot make known its evidence against Dr. Hauser, nor can he defend himself, until the government's report is ready. That leaves both in difficult positions. Harvard has accused a prominent professor of serious failings yet has merely put him on book leave. Dr. Hauser, for his part, cannot act publicly to prevent the derailment, at least for the moment, of his rising scientific career.

Harvard's investigation has been "lawyer-driven," says a faculty member who spoke on condition of anonymity, and has stuck so closely to the letter of government-approved rules for investigating misconduct that the process has become unduly protracted — it lasted three years — and procedurally unfair to the accused.

"I think it legitimate to ask why the Harvard brass did not push back against their lawyers," this member said. "At Harvard we now have the Un-Larry administration — no risk-taking, no thinking outside the box, no commitment to principles that challenge standard university practice," he said, referring to Harvard's previous president, the economist Larry Summers.

Dr. Hauser's difficulties began in 2007 when university officials went into his lab one afternoon when he was out of the country and publicly confiscated his records, an action based on accusations by some of his students.

For the next 18 months he had no idea what he was accused of. A troika of Harvard department heads then delivered a secret report. Dr. Hauser has amassed substantial legal debts in defending himself, his friends say. Harvard presumably has substantial evidence against Dr. Hauser.

He was investigated by a committee of fellow professors, and their findings were endorsed by the dean of the faculty of arts and sciences, Dr. Michael D. Smith. But from what is on the record so far, at least, Harvard's charges may or may not meet the government's definition of scientific misconduct, which is reserved for ethical offenses, like fabrication, falsification or plagiarism, that directly undermine the research process.

Two of Harvard's eight charges of scientific misconduct involve published papers for which some of the original raw data is missing. But Dr. Dahlberg, of the Office of Research Integrity, said: "Missing data is not scientific misconduct. The whole purpose of O.R.I. is to go after serious fraud and not the peccadilloes one might find in many labs."

Dr. Hauser and a colleague have redone the experiments and notified the two journals involved that they got the same results as reported. A third charge, apparently the most serious, concerns the article in *Cognition*.

The article, published in 2002, reported that rhesus monkeys can distinguish a novel string of sounds from a control sequence, an issue which has important bearing on their capacity for language. The novel and control sound sequences must be alternated so as to keep background conditions as similar as possible. But the video of the experiment contains only novel sequences.

Critics like Dr. Altmann at first charged that the controls had never been done, and that since control conditions are reported in the paper, they must have been concocted. But Dr. Altmann, a psychologist at the University of York in England, now says his earlier accusation was "heavily dependent on the knowledge that Harvard found Professor Hauser guilty of misconduct." When he gave the issue further thought, he saw an alternative explanation.

In the experimental setup, the monkey is in a soundproof box. The researchers can see the computer is playing a sound but cannot hear it. What could have happened is that the computer, through a programming error, substituted a second test sound for the control sounds, and the researchers, unaware of the problem, wrote up their report assuming the control sounds had been played as planned. Even so, it is far from clear how the data on the video led to the reported results. This would be a devastating error, but not fraud. "It is conceivable that the data were not fabricated, but rather that the experiment was set up wrong, and that nobody realized this until after it was published," Dr. Altmann wrote.

Mr. Watumull, a linguistics student, said that when he worked in Dr. Hauser's lab in 2007 he performed a similar experiment. It is "perfectly possible" that such an error could occur, he said, because the experimenters are blinded to the conditions of the experiment.

Harvard's five other charges of scientific misconduct involve disagreements between Dr. Hauser and his students, all of which were corrected before any articles were published. E-mails in one of these cases, leaked to *The Chronicle of Higher Education*, concerned the same kind of experiment as the *Cognition* paper: researchers scored how often a monkey turned its head to the loudspeaker, meaning it heard the sound as novel.

In analyzing the experiment, Dr. Hauser scored the head turnings as significant, but a graduate student and a research assistant both found the monkey did nothing. The e-mails show Dr. Hauser telling his students that "we need to resolve this because I'm not sure why we are going in circles."

The research assistant later wrote to the Harvard authorities, "The most disconcerting part of the whole experience to me was the feeling that Marc was using his position of authority to force us to accept sloppy (at best) science." It was this complaint that prompted the inquiry.

In at least one previous disagreement with students, Dr. Hauser backed off when challenged. A former student who worked in Dr. Hauser's lab before 2007 said Dr. Hauser had required the use of a statistical test that provided a publishable result.

The student, who spoke on condition of anonymity, felt the test was inappropriate and objected. After discussion, Dr. Hauser agreed and the result was not published. "I worked with Marc for years on dozens of experiments, and I never saw any problems with the handling of data that were this serious," this student said, referring to the Harvard committee's charges.

A more recent student, Mr. Watumull, said he never saw Dr. Hauser putting improper pressure on people to reach a conclusion. "He's truly one of the greatest teachers I had as an undergraduate," Mr. Watumull said. "He's very well known in the department for being solicitous of students and inviting them to offer their own opinions."

One of the few people to have seen any documents from the Harvard inquiry is Bennett Galef, an expert on animal behavior at McMaster University in Ontario. Because of his interest in research ethics, he was asked by Dr. Hauser to review the charges relating to the three published papers. Dr. Galef said he concluded, based on what he was shown, that there was no clear evidence that Dr. Hauser had acted unethically.

Dr. Galef referred to the tensions that can arise in a large laboratory where some students are more successful than others. "Marc should have supervised more closely," he said. Dr. Galef also questioned whether those conducting the inquiry fully understood the culture of an animal behavior laboratory. "As I understand it, the investigating committee were all physical scientists, and they have a very different approach to research and data-keeping than behavioral researchers do," he said.

In an interview, Dr. Hauser declined to discuss the eight charges against him. But he did talk about another of his experiments cited by critics, a mirror recognition test, which is not part of Harvard's investigation.

In 1995 he published a finding, which he later wrote that he could not repeat, that cotton-top tamarin monkeys could recognize themselves in a mirror. This contradicted a well-known finding by the psychologist Gordon G. Gallup that only humans, chimps and orangutans can recognize themselves.

Dr. Gallup asked for a tape of the experiment, which Dr. Hauser provided. But Dr. Gallup could see no evidence, he has said, that the monkeys were reacting as Dr. Hauser had reported. To critics, this seemed an example of Dr. Hauser rushing to unsustainable conclusions.

In Dr. Hauser's view, his article correctly reported the cotton-tops' reactions. One of the difficulties of the animal cognition field is that experimenters have to recognize often subtle changes in an animal's head movements, and judge whether this is a response to the test sound. Scoring an animal's responses is quite subjective. It can take months to train someone to score rhesus monkeys, and a person skilled at scoring rhesus may fail with tamarins.

Dr. Hauser's 1995 article was written with two colleagues trained in scoring cotton-top tamarins reliably. Dr. Gallup may not have spotted the reactions because he is not trained in scoring cotton-tops, Dr. Hauser said.

Why, then, could Dr. Hauser not repeat the experiment? The reason, he believes, has to do with another unresolved problem in the animal cognition field, that of how to deal with the variability in individuals.

Just as with people, some animals are gifted, others less so. Alex was the wonderfully intelligent gray parrot studied by Irene Pepperberg; no other parrots have equaled his abilities to distinguish colors and numbers. A collie dog called Rico was reported in *Science* in 2004 to possess a 200-word vocabulary, but has never been heard of since, suggesting that for whatever reason the experiment cannot be repeated.

The prodigy problem can interfere with less spectacular experiments. Dr. Hauser says that the first group of tamarins he tested for self-recognition may have included a few very adept individuals but that later groups were more average. He was unable to get the same result, and published his failure to do so.

Disagreements over the appropriate method are quite common in the animal cognition field, as is evident in the fact that some of the most spectacular experiments cannot be repeated. Disagreements over method also seem to have been involved in at least some of the five cases involving differences between Dr. Hauser and his students.

The e-mails leaked to *The Chronicle of Higher Education* were portrayed as an instance of Dr. Hauser pressuring his students to reach conclusions they thought unjustified. But they could also have involved a technical difference of opinion about how to score rhesus monkey behavior, a matter in which Dr. Hauser is trained and the two students were not.

Dr. Hauser has already acknowledged making "significant mistakes," but has admitted to nothing worse. It remains to be seen whether or not the Office of Research Integrity will see these mistakes as serious enough to count as scientific misconduct.

"Maybe down the line there'll be some forgiveness and a way to re-enter," Dr. Hauser said. "I feel I have a lot more to contribute. But it's been brutal."