

The Washington Post

SUNDAY, DECEMBER 23, 2012

FBI lab's woes cast growing shadow

DOUBTS ABOUT STATE, LOCAL HAIR MATCHES

Federal training linked to suspect court testimony

BY SPENCER S. HSU

Thousands of criminal cases at the state and local level may have relied on exaggerated testimony or false forensic evidence to convict defendants of murder, rape and other felonies.

The forensic experts in these cases were trained by the same elite FBI team whose members gave misleading court testimony about hair matches and later taught the local examiners to follow the same suspect practices, according to interviews and documents.

In July, the Justice Department announced a nationwide review of all cases handled by the FBI Laboratory's hair and fibers unit before 2000 — at least 21,000 cases — to determine whether improper lab reports or testimony might have contributed to wrongful convictions.

But about three dozen FBI agents trained 600 to 1,000 state and local examiners to apply the same standards that have proved problematic.

None of the local cases is included in the federal review. As a result, legal experts say, although the federal inquiry is laudable, the number of flawed cases at the state and local levels could be even higher, and those are going uncorrected.

The FBI review was prompted by a series of articles in The Washington Post about errors at the bureau's renowned crime lab involving microscopic hair comparisons. The articles highlighted the cases of two District men who each spent more than 20 years in prison based on false hair matches by FBI experts. Since The Post's articles, the men

have been declared innocent by D.C. Superior Court judges.

Two high-profile local-level cases illustrate how far the FBI training problems spread.

In 2004, former Montana crime lab director Arnold Melnikoff was fired and more than 700 cases questioned because of what reviewers called egregious scientific errors involving the accuracy of hair matches dating to the 1970s. His defense was that he was taught by the FBI and that many FBI-trained colleagues testified in similar ways, according to previously undisclosed court records.

In 2001, Oklahoma City police crime lab supervisor Joyce Gilchrist lost her job and more than 1,400 of her cases were questioned after an FBI reviewer found that she made claims about her matches that were "beyond the acceptable limits of science." Court filings show that Gilchrist received her only in-depth instruction in hair comparison from the FBI in 1981 and that she, like many practitioners, went largely unsupervised.

Federal officials, asked about state and local problems, said the FBI has committed significant resources to speed the federal review but that state and local police and prosecutors would have to decide whether to undertake comparable efforts.

FBI spokeswoman Ann Todd defended the training of local examiners as "continuing education" intended to supplement formal training provided by other labs. The FBI did not qualify examiners, a responsibility shared by individual labs and certification

Improper forensic testimony

Since at least the early 1970s, the FBI Laboratory's hair and fiber unit trained hundreds of state and local examiners to testify about the rareness of coincidental hair matches, citing their case experience. However, there was no accepted scientific research behind the claim. The FBI stopped declaring matches without DNA confirmation in 1996, but it took until this year for authorities to announce a review, limited to cases involving FBI examiners.

When the FBI was aware of problems with hair analysis:

At an FBI symposium, defense experts and British and German police criticized FBI testimony practices as misleading and unsupported by data.

A federal judge in Oklahoma excludes hair evidence under a 1993 U.S. Supreme Court ruling aimed at junk science, stating that the inability to explain the rareness of a coincidental match made results vague and prejudicial.

A Justice Department review of 13 discredited FBI lab agents eventually focuses on a hair examiner who exaggerated the significance of hair matches.

DNA tests exclude hair matches declared by FBI experts more than 11 percent of the time. Earlier, in 1996, the Justice Department reported that more than 20 percent of the first 28 DNA exonerations included hair evidence.

How the FBI responded:

From 1985 to 1989, the FBI Laboratory's hair and fiber unit begins requiring that any hair match be confirmed by a second agent.

The unit stops declaring matches based on visual comparison alone, requiring confirmation by DNA testing.

Between 1997 and 2001, the unit begins requiring written protocols for tests, note-keeping and testimony review.

In 2012, the FBI and Justice Department announce a review of lab reports and testimony in all hair matches declared by the unit before Dec. 31, 1999.

Source: Staff reports

THE WASHINGTON POST

bodies, she said.

Michael Wright, president of the National District Attorneys Association, said local prosecutors cannot simply order labs to audit all or even a sample of cases handled by FBI-trained examiners, because such an undertaking might be time- and cost-prohibitive for smaller agencies.

The chairman of the laboratory accreditation board of the American Society of Crime Laboratory Directors said it is gathering information to guide members.

"It is something we take seriously, and we are going to address it accordingly," said Pamela Bordner, the chairman.

The announcement in July of the Justice Department review of federal cases marked a turnabout from the mid-1990s, when an inquiry looked at a limited number of cases and, in the area of hair comparison, focused on the work of one examiner at the FBI lab.

In its April investigation, The Post found that Justice Department officials failed to tell many defendants or their attorneys of questionable evidence and that the results of the review remained largely secret.

In addition, Justice Department officials have for years blamed errors on isolated failures by rogue examiners, careless prosecutors or inept defense lawyers.

But former chiefs of the FBI lab's hair

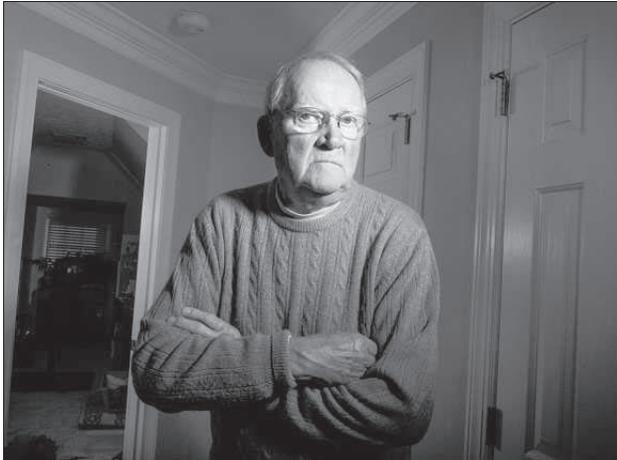
and fiber unit now acknowledge that the problems were more widespread. Some federal examiners, testifying in cases across the country, overstated the importance of hair evidence and responded to questions about the scientific accuracy of hair matches by citing amorphous statistics drawn from their experience.

Moreover, they said, examiners should have been trained to accurately portray their findings in court. When local lab examiners went to the FBI for training, they received the same inadequate instruction.

Myron T. "Mike" Scholberg, hair unit chief from 1978 to 1985, and Alan T. "Al" Robillard, chief from 1988 to 1990, said that in hindsight, they were not properly trained to answer a crucial question for jurors: How often might the hairs of different people appear to match? The truth is that there was no scientific way to know.

Instead of simply acknowledging the uncertainty, agents at times drew statistics from their cases without explaining why that was an incomplete or even misleading answer, Scholberg and Robillard said.

Harold A. "Hal" Deadman Jr., a top hair unit scientist who trained more than 600 examiners from 1972 to 1987, said he always explained to jurors why his case experience gave an incomplete picture of the accuracy



PHOTOS BY RICKY CARIOTI/THE WASHINGTON POST

Myron T. Scholberg, left, a former head of the FBI hair unit, acknowledged weaknesses in how examiners were trained to give testimony. Max Houck, left, the D.C. forensics chief, was the FBI unit's first civilian examiner. He said he changed the way he testified in the late 1990s.

of hair comparisons.

But Deadman said DNA testing should be done in all convictions that were based mainly on visual hair comparison, because of weaknesses in trial testimony and examiner results.

Interviews with the former unit chiefs, as well as more than 20 practitioners, scientists and legal experts, and a review of court records, training notes and transcripts of meetings indicate that some FBI lab examiners tried to skirt the limitations of their scientific findings in testimony and that they were encouraged to do so by their trainers.

As warnings about the problems mounted — through DNA exonerations, whistleblower complaints, court rulings — bureau managers implemented stronger protocols, but they limited disclosure of the problems they found. More forthcoming disclosure could have jeopardized convictions.

“If the FBI is going to be a role model, we need to see this federal audit lead to wider audits of labs across the country,” said Myrna S. Raeder, a Southwestern University law professor who is leading an American Bar Association effort to improve forensic evidence. “If you had even the elite FBI analysts out there crossing the line and exaggerating the forensics, that sent a terrible message that ‘anything goes.’”

In a letter this month to the Senate Judiciary Committee, the Justice Department acknowledged that FBI examiners “may

have exceeded the limits of the science by overstating ... conclusions” in some cases.

“The Department and the Bureau believe it is necessary and appropriate that defense counsel and defendants are informed of any inappropriate testimony by FBI Laboratory examiners,” Acting Assistant Attorney General Judith C. Appelbaum wrote.

Powerful testimony

Before DNA profiling, testimony of a hair match was a powerful way for prosecutors to boil down an ambiguous case to a single, incriminating piece of physical evidence left at the scene of a crime.

No other agency in the United States performed as many hair examinations or believed as much in the technique as the FBI lab's 10-member unit of hair examiners.

But The Post's investigation earlier this year showed how agents, prosecutors or both sometimes exaggerated the significance of the evidence they had.

For example, in a 1980 Indiana robbery case, one agent told jurors that he was unable to distinguish between the hair of different people just once in 1,500 cases he had analyzed.

In one of the District cases, federal prosecutors claimed that the agent had been unable to tell hair samples apart only “eight or 10 times in the past 10 years, while performing thousands of analyses.”

In another, the prosecutor said in clos-

ing arguments, “There is one chance, perhaps for all we know, in 10 million that it could [be] someone else’s hair.” That defendant was declared innocent this year.

The problem is, as an expert peer review panel wrote in Melnikoff’s case, “There is not — and never was — a well established probability theory for hair comparison.”

As noted in 2009 by the chief of the FBI hair team, the proper answer to the question of how often hairs from different people might match is, “We do not know.”

Peter Neufeld, co-founder of the Innocence Project, said hair analysis practitioners should end their resistance to putting scientists in charge of setting clear, consistent standards for lab reports and testimony.

“When the stakes involve life and liberty, those scientific parameters and standards should be set by scientists, not by law enforcement,” said Neufeld, whose organization advocates for people trying to prove their innocence through DNA testing. The Innocence Project and the National Association of Criminal Defense Lawyers were consulted by the FBI on the national review.

Vague standards

The FBI has known for decades that hair found at a crime scene is a valuable piece of evidence. Before DNA testing, agents would use a microscope to compare the evidence with a sample of hair from a suspect.

A visual analysis can tell animal hairs from human hairs; human hairs by race and body part; whether hairs were dyed or otherwise treated; and how hairs were removed from the body. Visual comparison, at its best, also can accurately narrow the pool of criminal suspects to a class or group or definitively rule out a person as a possible source.

But it was not possible to declare an absolute match. So the FBI had a problem. Hair comparisons could yield good evidence. But agents struggled to explain to a jury how good.

Morris Samuel “Sam” Clark was the head of the FBI’s hair unit when it began training state and local analysts in 1973. He

said he long believed that examiners could trace hairs from a crime scene to a particular person with a high degree of probability — even though there is no scientific proof that is possible.

But Clark, who did graduate work in biology at Harvard and retired in 1979, said laboratory experience should not be discounted. He did “hundreds and hundreds of comparisons” over nearly 20 years, and he believes that he was a qualified court expert, he said in an interview from his home in Spotsylvania County.

The FBI’s training regimen, which required agents to compare hairs side-by-side under high-powered microscopes for a year before working on live cases, gave lab veterans confidence that they could tell the difference between individuals’ hairs just as an ordinary person could distinguish between their faces.

They embraced a set of vague standards. In written lab reports, FBI agents would include the caveat that hair examination was not a basis for positive identification.

In court, however, they could suggest that it would be highly unlikely for an examiner’s match to be wrong. The bureau left it up to individual labs and examiners to explain matters to jurors. Agents were trained to say that in their “personal experience” they had rarely seen hairs from different people that looked alike.

That evolved into jurors’ hearing numbers that had a huge impact even if they lacked scientific grounding. After a slaying in Tennessee in 1980, an FBI agent testified in a capital case that there was one chance in 4,500 or 5,000 that a hair came from someone other than the suspect.

But as experts from around the world would later note, the FBI-taught answer was misleading. In reality, FBI examiners did not compare every hair to every other hair they had ever examined. They simply compared crime-scene hairs and hair samples from individuals relevant in each case.

Examiners kept no “database” of samples, which went back to police evidence

files. And differences between hairs are so fine that a person can generally keep only a handful of hairs in mind at any time.

“The claim you could keep all those hairs in your head and sort them in your mind, that would be hard to do,” said Mark R. Wilson, a 23-year FBI veteran who helped develop DNA testing for hair in 1996. “After about three or four [hairs], it gets confusing.”

The claim was called into question at an international conference hosted by the FBI in 1985, but the training was not overhauled for at least a dozen more years.

“It was not promoted, put it that way,” to give juries a more accurate picture of the limits of the technique, said John W. Hicks, who spent about five years in the hair unit in the 1970s and who directed the FBI lab from 1989 to 1994.

Robillard, the former hair unit chief, said that he always waited for a defense attorney to challenge his claims about the accuracy of hair analysis but that neither they nor judges usually caught the logical sleight of hand.

“You would expect a defense attorney to say, ‘Wait — are you, Robillard, saying you compared every person’s hair to every other one?’ That’s the screaming question for cross-examination,” Robillard said. “I can’t off the top of my head remember ever having a defense attorney say that.”

Like the other agents interviewed, Robillard, now a private expert who lives on Martha’s Vineyard, in Massachusetts, said FBI experts were not trying to mislead but to convey in layman’s terms why they were confident in their hair associations.

Not all former chiefs agreed that examiners should have testified differently. Edward L. “Ed” Burwitz, who led the unit from 1985 to 1988, called that “a legal question that I don’t feel confident to answer.”

Like others, Burwitz said he never got complaints about examiners’ testimony. He called the recent criticism a matter of “Monday morning quarterbacking.”

Clark also defended his work, including

the FBI training.

“This was not fly-by-night stuff, not idle conclusions on our part. I think we made a very significant contribution to the criminal justice system,” Clark said. “If [examiners] made a mistake, it’s a personal mistake, and it’s not a matter of [our] training them... nor the whole science of microscopic hair exams, because we did our best.”

Crash courses

The FBI lab began training state and local hair examiners in 1973, as the bureau worked with the nation’s crime lab directors to expand forensic methods.

Deadman said he trained about 600 examiners from outside the FBI between 1973 and 1987, and others estimated that an additional 450 examiners were trained over the next dozen years.

No one knows how many cases local and state hair examiners handled. Estimates of their collective caseload vary from 20 percent to more than half of all hair exams during the period under review. Most of the rest were federal cases.

Yet, FBI agents and others say they doubt the quality of the training, even as they acknowledge that it was a valued credential for state and local labs.

Instead of working with hairs for an entire year before starting trial work, some local trainees spent a week at the FBI Academy at Quantico and then went back to labs where they were one of one or two designated “criminalists,” analyzing everything from hair to paint chips to glass, Robillard said. They might handle a handful of hair cases a year, using substandard equipment while under constant pressure from investigators.

With 200-plus crime labs serving 18,000 police agencies in the early 1990s, DeForest said, “There was no monitoring of people. ... That whole thing for something this complex was ill-conceived, and maybe [the FBI] should have recognized that.”

In 2004, Melnikoff lost his crime lab job in Washington because of errors whose discovery led to three overturned convictions in

Montana. One of those cases was the child rape conviction of Jimmy Ray Bromgard, who served more than 15 years in prison before DNA tests showed he didn't commit the crime.

At Bromgard's 1987 trial, Melnikoff said he found head and pubic hairs "microscopically indistinguishable" from Bromgard's, and he told the jury that there was less than one chance in 10,000 of a coincidence. He based this assertion on his case experience, multiplying by 100 the 1 in 100 frequency with which he claimed to have seen head and pubic hairs he could not tell apart.

After Bromgard was exonerated in 2002, a five-member panel that included Deadman said Melnikoff made "egregious misstatements not only of the science of forensic hair examinations but also of genetics and statistics."

Melnikoff's defense in a civil suit brought by Bromgard was that he simply acted as he was trained.

Michael A. Howard, a 24-year Oregon State Police veteran who also took the Quantico course, noted that Melnikoff's examination and lab report followed FBI practices.

"I took the [FBI] class in 1982 and was not advised to avoid the use of probabilities. ... We were taught that our own experience was most important, and that is what Mr. Melnikoff was doing," Howard told a federal court in Montana in 2007.

In an interview, Howard elaborated. "They didn't say, 'Use it,' and they didn't say, 'Don't use it,'" he said. Instead, he said, the FBI's position was, "You're going to have to decide for yourself, based on your experience, how strong you can state it."

Gilchrist also was accused of misidentifications, misleading testimony and withholding or destroying evidence. In 2001, she was fired from the Oklahoma City Police Department, and authorities set out to reexamine more than 1,400 assigned cases, including a dozen death row cases.

In one case, David Johns Bryson spent 16 years in prison for a 1982 rape, but he was freed in 1999 after DNA results showed another man committed the crime.

Gilchrist testified that she found four hairs that were like Bryson's and that she never saw hairs from different people with the same characteristics. She said, "I would think it would be impossible not to be able to distinguish hairs from two different individuals."

But in April 2001, Douglas W. Deedrick, then head of the FBI unit, found that Gilchrist's matches were wrong and that by implying hairs were "unique," Gilchrist "misrepresent[ed] the science."

Again, Gilchrist took her cue from bureau training. In her files, she kept a certificate of completion from her January 1981 class, including a session on "Discussion of the significance of hair comparisons, testimony matters and pertinent literature."

In her notes, she copied the FBI caveat that one cannot conclusively determine the source or origin of a hair. But, the notes also showed that instructors were teaching their students how to sidestep the limits of the science — by pointing out their experience.

"Can conclude source — point out however in my experience, have rarely seen hairs from diff. people exhibiting the same microscopic characteristics," the notes say.

Eventual change

FBI veterans pointed out that the hair unit gave up members who helped the agency pioneer forensic nuclear and mitochondrial DNA testing. As DNA testing became more common, the limits of microscopic hair comparison became clearer.

Max Houck, who was the unit's first civilian, non-agent examiner, said he changed the way he testified in the late 1990s after consulting an old statistics textbook. Training of examiners also shifted away from citing numbers, probabilities or statistics by 2001, Houck said, as the lab gained outside accreditation and replaced agents with civilian scientists.

Asked why it took until now to correct errors, Houck, the head of the new D.C. Department of Forensic Science, cited a variety of reasons: The "conservatism" of forensic science, the legal system's dependence on

precedence and, finally, government bureaucracy and the FBI's proud culture.

"Could it have happened sooner? Yes," he said. "Would it have cost more money? Yes. Would it have been more disruptive?

Probably. Would we have gotten a better answer? I don't know."

To his list, Houck added one more question. "Does that mean justice was served? Not necessarily."

hsus@washpost.com