Georgia State Law Review Symposium Keynote Address: Uncovering Forensic Flaws - An Outside Perspective

Spencer S. Hsu
The Washington Post, spencer.hsu@washpost.com

Follow this and additional works at: https://readingroom.law.gsu.edu/gsulr

Part of the Civil Rights and Discrimination Commons, Constitutional Law Commons, Courts Commons, Criminal Law Commons, Criminal Procedure Commons, Evidence Commons, Law and Society Commons, Law Enforcement and Corrections Commons, Legal Ethics and Professional Responsibility Commons, Legal Profession Commons, Public Law and Legal Theory Commons, and the Science and Technology Law Commons

Recommended Citation
Available at: https://readingroom.law.gsu.edu/gsulr/vol34/iss4/9

This Article is brought to you for free and open access by the Publications at Reading Room. It has been accepted for inclusion in Georgia State University Law Review by an authorized editor of Reading Room. For more information, please contact mbutler@gsu.edu.
SYMPOSIUM KEYNOTE ADDRESS:
UNCOVERING FORENSIC FLAWS: AN OUTSIDE PERSPECTIVE

Spencer Hsu*

SH: Thank you very much to Dean Cino and Georgia State University Law School and the Law Review for organizing this remarkable symposium. As others have noted, it’s humbling to be in the same room with so many experts and practitioners in forensic science, not to mention the presenters here, Justice Nahmias, and Barry Scheck. I notice we have a Twitter stream here, so unlike most of you, I have not been trained to speak for more than 140 characters at a time. So, I hope you’ll bear with me, especially after that last panel, which I thought did so much to give you a glimpse of how complicated, and if not structural . . . these problems are.¹ [These problems] have deep institutional roots that are hard to get at.

So, I’ve been invited to give an outsider’s perspective of uncovering forensic flaws, so I thought to start I would turn to the philosopher king of outsiders, Gary Larson, whom some of you may remember as the creator of The Far Side cartoon. You might remember this one: What We Say to Dogs and What They Hear. And I thought of this when I remembered first reading the 2009 NAS report on forensic science, which said that, other than DNA, no other pattern-based technique had been scientifically validated to match evidence to individuals. It recommended removing crime labs from government control, and its powerful opening words were that “[f]or decades, the forensic science disciplines have produced valuable

¹ This transcript is a reproduction of the Keynote Address by Spencer Hsu at the 2017–2018 Georgia State University Law Review Symposium—From the Crime Scene to the Courtroom: The Future of Forensic Science Reform—on April 6, 2018. Spencer Hsu is an investigative reporter at the Washington Post, two-time Pulitzer Prize finalist, and national Emmy Award nominee.

¹ Mr. Hsu’s keynote presentation was preceded by a panel entitled “To Err Is Human,” during which panelists discussed the role of human error in forensic evidence collection, evaluation, and presentation, with particular attention to causes of error and effective methods for minimizing error.
evidence that has contributed to the successful prosecution and conviction of criminals as well as to the exoneration of innocent people.” And it went on to say, though, that “the forensic science system, encompassing both research and practice, has serious problems that can only be addressed by a national commitment,” and it can only be done with a “significant infusion of federal funds.” And I think it was a NACDL official who observed that what prosecutors heard was the first line: forensic science produced valuable evidence; law professors heard that the system had deep systemic problems; and practitioners heard a call for a need for an infusion of resources. Now, Gary Larson had another cartoon, *What We Say to Cats*, and this might apply to what the judiciary took away from the report in terms of questioning the admissibility of forensic science. Now, thankfully, as you all know this is not a problem for journalists since what people say has no bearing on what we actually write.

So, in that spirit let me talk about, if I can, what Gene asked: how I came to write about the FBI hair unit, what I learned about how stakeholders addressed errors found in the system, and the role of the media. As you just heard, after our stories in 2012, they started with 20,000 cases that this hair unit had handled. And, what was remarkable was that this was the gold standard of the gold standard: it’s the FBI lab; it’s the U.S. Department of Justice. This was a unit that actually blossomed after Hoover; it was viewed as the next best thing after fingerprinting, especially in violent crimes when you had biological material that you could test. It’s more specific than blood or semen. And, the most serious crimes: rape and murder. And they got from there to the 2,500 to 3,000 cases where there was a positive match. Three years after our stories appeared, the DOJ and FBI formally acknowledged that nearly every one of twenty-eight examiners in this elite unit gave flawed testimony in almost all trials where they offered evidence and that they overstated matches in ways that favored prosecutors 95% of the time and that included thirty-two defendants sentenced to death. Now, we need to caveat that [for] many of these people, there is abundant evidence of their
crimes, many were convicted of other crimes, and also many pleaded guilty. But, in the [District of Columbia] alone, there have been six DNA exonerations for people who’ve served decades in prison. The district has about 1/500 of the population of the U.S.

So let me start in the middle of the story, which is really the story of Donald Gates and his remarkable lawyer, Sandra Levick of the D.C. Public Defender’s Service. Donald Gates was an Ohio man sentenced to life in prison after being convicted for raping and murdering a Georgetown University student in Rock Creek Park in 1981 in D.C. FBI agent Michael Malone found two hairs from an African American male on the victim’s body and matched them to Gates, who was also implicated by a purported confession to a paid jailhouse informant. Gates always maintained his innocence, and in about 1999, he read or heard of DNA fingerprinting on a radio show, and he wrote the trial judge. So, if you’re counting references to a media role, here’s one. He said to the judge, “You don’t want to convict an innocent man. There is this thing called DNA fingerprinting, but if you give me a lawyer then I will show you I didn’t commit the crime.” So, the trial judge, now Chief D.C. Superior Court Judge Fred Ugast, appointed an attorney, Roger Durban. They got the physical evidence, tested by cell mark, but the testing technology at that point was unable to get results with the DNA available. And Gates’ petition was dismissed. By this time he had exhausted his appeals. This was sort of his last gasp before he would disappear without anyone hearing from him again.

But then in 2007, inspired by another newspaper article—Sandy Levick thinks it was maybe an editorial in the New York Times noting the two-hundredth DNA exoneration—Roger Durban read that, remembered Mr. Gates’ case, and wrote another letter to the judge, copied to the D.C. [Public] Defender Service. He couldn’t remember Gates’ name. He had details of the case though, and he said something to the effect that there is not a time that goes by when he drives by those woods at the base of the Rock Creek Park near the Kennedy Center, where the crime occurred, that he didn’t wonder if Gates might still be on this earth and hoped that with the advances in
DNA technology that someone would look into his case. He would do it himself except he had Parkinson’s disease, he said, and he couldn’t practice law anymore. So that is how Sandy Levick got the case, and Judge Ugast, after he retired, kept it. There were ups and downs. There was an order in the early round of testing to preserve the evidence, but it had been lost or destroyed. No one could find a trial transcript, but they did recover appellate court decisions, pleadings and motions for re-trial.

But the key break came from Jim Traynor, who was a D.C. homicide detective, and he ran a cold case unit with the help of criminal justice graduate students in Maryland. He found the physical evidence and a duplicate set of slides kept at the D.C. Medical Examiner’s Office. They took duplicates of autopsy slides after sending them off to the FBI; they kept them in refrigerated storage for twenty-seven years. And that is how it was found that Gates could not have committed the crime. He was exonerated just before Christmas 2009, set free after twenty-seven years. Prosecutors in 2012 traced the crime scene evidence to another man who died the year before, an office temp in the building where the victim worked. A federal jury later determined that police had framed Gates; detectives fed incriminating details to the informant. And if the story had ended there, it would have been extraordinary but not unheard of. Even so, look at the many steps it took: the homicide detective, the police, the lawyers, the medical examiner’s office.

While waiting for the DNA results to come back, Sandy Levick’s husband was driving her to a conference at NYU on Brady v. Maryland, and she started reading about hair evidence. So there’s another reference. She had a brainstorm: if we know that hair analysis is kind of sketchy, what do we know about the hair examiner? Just a few days before those lab results came back, she got a package back from the FBI, and it was Gates’ FBI file, which she had FOIA’d months before. On the top sheet it said, “Do not

2. The Freedom of Information Act, or FOIA, “provides that any person has the right to request access to federal agency records or information,” unless the information requested is protected by an exception to the Act. The Freedom of Information Act, U.S. DEP’T OF STATE, https://foia.state.gov/Learn/FOIA.aspx [https://perma.cc/2V6J-8M27] (last visited July 6, 2018). To have

https://readingroom.law.gsu.edu/gsulr/vol34/iss4/9 4
remove.” It was a cover sheet that referred to FBI hair examiner Michael Malone and a confidential government review of his work. Around that time she also got a thumb drive from another former FBI agent, Fred Whitehurst, who had FOIA’d more than 10,000 pages of DOJ and FBI records from that nine-year review of Malone and twelve other FBI agents. His work had been discredited in a 1997 inspector general (IG) report triggered by Whitehurst’s whistleblowing.

To take people back to the bad old days—and the positive side of this is look at how much progress has occurred—Whitehurst was a chemist and lawyer who worked in the FBI’s crime lab. He testified in New York federal court in 1995 that there was a great deal of pressure from his FBI lab superiors to bias evidence or ignore findings that did not support the prosecution’s theory. In fact, Whitehurst had written or passed along scores of memos over the years warning of a lack of impartiality in scientific standards. The IG went on to find that the chief of the lab explosives unit, for example, repeatedly reached conclusions that incriminated defendants without a scientific basis in the 1995 Oklahoma City federal building bombing. The head of toxicology lacked judgement and credibility and overstated results in the 1994 O.J. Simpson investigation. After the first World Trade Center attack in 1993, the key FBI witness worked backwards tailoring his testimony to reach the result he wanted in the trial of “The Blind Sheikh” suspected of plotting the attack, Omar Abdel-Rahman. Other agents “spruced up” notes for trial, altered reports without the author’s permission, or failed to document or confirm their findings.

So the Bromwich Report in 1997, the inspector general report, stopped short of saying any convictions should be overturned or that any convictions were jeopardized, and also of concluding that anyone committed perjury. But it did lead the FBI to overhaul its practices, its procedures, and its accreditation going forward. That was really a paradigm shift that I think led to the others.

“FOIA’d” information is to have submitted a request for the information under the Act.
In that case though, the DOJ also committed to review, by coincidence, another 3,000 defendants’ cases over what eventually took nine years. These were cases handled by the thirteen agents whose work was questioned. And they pledged to turn over any exculpatory information to defendants. They ultimately found 402 cases of defendants whose handling was so problematic that this Brady task force commissioned a new review of findings. But, what we learned in our reporting was that the Department reported its findings only to prosecutors, not to defendants, and never issued a public report. And, the one news story that I could find by AP by John Solomon, who later was an editor at the Post and the Washington Times, said, “DOJ officials, speaking on condition of anonymity, left the impression that anyone whose case had been affected had been notified, and in any case, no convictions have been overturned.”

In reality, as we found, while many prosecutors made swift and full disclosures, in most cases they did incompletely, years later, or not at all. And that was how Sandy Levick found out just before Christmas in 2009, through her FBI FOIA, that Donald Gates should’ve been notified twelve years earlier, in 1997, that Michael Malone’s hair match case on which his conviction hinged, was questioned by the DOJ because Michael Malone had been deemed [not] credible.

Fred Whitehurst’s information was potentially even more troubling. In the thousands of pages of these task force documents, almost all of the hundreds of re-reviewed cases were by that agent, Michael Malone. As is better understood now, but what was reported over the years in the Wall Street Journal and Florida newspapers, is that he routinely gave identical assurances to juries that he had examined the hairs of 10,000 people in his career and that there were only two or three occasions in which the hair from two different people was so similar that it couldn’t be distinguished.

You’ve probably read and heard this now, but those assertions of the likelihood of a match—implications of statistical certainty and reliability—were a series of statements that were misleading if not
outright false or lies. As we know now, justice officials had blamed errors on isolated failures by rogue examiners, careless or inept prosecutors, or defense lawyers, but there was a suggestion here that practices might have been more systemic. And that, as an outsider, is one of the striking things to me: how long it took. And just to unpack this, as Mark Stolorow was just saying to me, progress has been sort of glacial in some areas. In 1974, researchers knew that visual hair comparisons were so subjective that different analysts can reach different conclusions about the same hair. In fact, even the same examiner, examining the same hair, can reach different conclusions at a different time. In 1984, the FBI acknowledged that hair comparisons can’t be used as a conclusive positive match for positive identification to say that “this hair came from this person.” In 1996, the DOJ studied the first twenty-eight DNA exonerations and found that twenty involved hair comparisons. In that same year, the FBI lab stopped declaring matches based on visual comparisons alone and began using them as a screening measure for DNA testing.

Then comes this Brady task force review, where non-FBI scientists were finding that Malone testified beyond the limits of science, drew conclusions without a scientific basis, and exaggerated and misstated research. You’ve heard this, but this claim of personal experience was kind of a lie within a lie or a misleading statement within a misleading statement. In reality, examiners didn’t compare every hair they ever tested to every other one, only a pair at a time. Most times you wouldn’t expect the hairs to match because you’re dealing with maybe a male perpetrator and a female victim. And when comparing a crime scene hair to a hair from a known person, you wouldn’t know if you were mistaken. In other words, you’re rarely comparing hairs between two people such that you would be in a position to know that they didn’t match.

And also agents told me, “You can’t remember more than a couple of hair comparisons once you’re done.” And they didn’t keep full notes or statistics or pictures. And, as we all know now, the key question here is how often might the hairs of two different people appear to match, and the truth is there’s no scientific way to know.
As the chief of the FBI hair team acknowledged in 2009, “The proper answer is we do not know.”

So, my part of the story just was to see if there were other Donald Gateses out there—people who were sentenced to decades or life in prison; who were convicted through flawed evidence; should’ve been told about problems in their case; or might’ve been convicted through testimony by other agents using similar techniques or practices as Michael Malone. And that’s where I started, and what we did compared to everything else done before was fairly straightforward. I met Levick; well, first of all, my experience with forensic science was quite limited. I did hear about it, though, because when I first came to the D.C. federal court, I was meeting with judges. One of them said, “You should take a look at this,” pushed something across his desk, and it was a speech by Judge Harry Edwards, the former Chief Judge of the D.C. Circuit Court who was the author of the NAS report. He had given a speech to the D.C. Superior Court, the criminal trial court in the district where these cases were happening, and in it he summarized the NAS report. My experience before was limited to when I was covering DHS; it was about 2006, five years after the 9/11 attacks, and they had put out a request to industry, to vendors, to develop a portable fingerprint machine. What they came back with was like a Kleenex-size popup box. It had red plate glass on all sides. You could have a suspect put their fingers on it; you could take this to a cave in Afghanistan or a boat, on a Coast Guard cutter or in the Mariel Straits off of Haiti. You know, you’ve got salt, spray, bad weather, the boat is going like this [tilting side to side], but you could beam it, get an instant fingerprint reading on databases.

Meanwhile, the government said, well we like your first draft, industry, but you know all these things you have about incomplete, inconclusive, partial hits—can you delete those? We don’t need those. And you think, well that makes sense because this is for investigative purposes. It’s not to be admitted in a court. You’re looking for national security threats. But it never occurred to me that there was something other than a positive hit. Then the judge showed us this.
I met Levick who had been working on two more DNA exoneration cases involving defendants named Santae Tribble and Kirk Odom who had spent decades in prisons based on the testimony of two different FBI agents. DNA wound up showing that Tribble couldn’t have contributed the hair found at his murder crime scene, and only a different man could have committed the rape for which Odom was convicted. In Tribble’s case, the two different FBI-trained analysts who were asked to re-examine the hairs afterward failed to distinguish human hairs from a dog hair in the sample or agree in a sample whether a hair showed Caucasian or African-American characteristics.

Fred Whitehurst and his lawyers at the National Whistleblower Center in D.C. gave us their files piecing together hundreds of reviewed cases without names attached. With these 10,000 pages of documents, the Post found more than 250 cases in which a scientific review was completed. We built a database after kind of using a Rosetta Stone approach, comparing case number to FBI case number to Justice Department case number to criminal court number and other biographical details of lawyers and dates. We found 137 identified by name. One of them, Benjamin Herbert Boyle, was executed in Texas in ‘97—more than a year after the DOJ began its review but without being notified. He would not have been eligible for the death penalty without the FBI’s flawed work according to the prosecutor’s memo. A Maryland man serving a life sentence for a 1981 double killing was another case where the prosecutor apparently never told the defendant of potentially exculpatory findings. Attorneys for the man, John Norman Huffington, say they learned about it from the Post. And these were just Malone cases. His case was vacated afterward.

And then we found, in appellate opinions, examples of other agents who testified like Malone did, unable to distinguish among different people’s hair only once or twice in 2,000 cases or once in 1,500 cases; others said 4,500 or 5,000 cases; one put the chances of findings two [indistinguishable] hairs at 5,000 times 5,000, so that’s 2.5 million. And, of course, we found cases where prosecutors were
saying, “You know, we don’t know. Maybe it could be one in one million; it could be one in ten million.” So we reported this in 2012, and the Justice Department and FBI, to their credit, took great steps this time: re-reviewed the cases of all investigators, researchers, and forensic analysts in that hair unit; they included the defense, and then they concluded in 2015 that these problems were systematic and didn’t count hundreds of state hair examiners who were also trained by the FBI.

So what are the lessons? After our story, the Justice Department Inspector General would say that as of 1999 the DOJ had enough information to review all hair unit cases, the action that they eventually took. It’s hard to say. We’ve heard the analogy of how often the system works. We’re reminded of the old adage—I can’t remember from which of the Founding Fathers it came, it may have been Thomas Jefferson—better ten guilty men go free than one innocent person be convicted. The assumption had been that there aren’t problems. The working assumption had been fingerprint examination is 100% accurate, 0% error. What we’ve since seen when the FBI re-tested hair matches using DNA, it found that in closed cases it was wrong 11% of the time. There’s been some testing when Virginia discovered a set of biological evidence going back decades. Similarly it was like a second set that had been kept by, I think, an examiner. They went back and looked and found that between 3% and 16%, depending on how you measured, of convictions contained errors.

We’ve started to see that errors are a little bit more than the system may have understood, and that’s because DNA has been able to prove it. Part of the problem here has been—I think this is overstating it, but—sort of a “prove me wrong” attitude. Or certainly an attitude that “we know the evidence is good, but we don’t know how to say how good.” And in the absence of pushback or in the absence of science or research, there is this challenge of not going too far. So we found many FBI and DOJ individuals, not to pick on them, again this is the gold standard. These are the nation’s leading institutions that undertook the responsibility, and the burden of, this
duty to correct and to go backwards, not just to improve the system going forward—which they’ve done—but to go back and try to look at righting past wrongs and admitting error.

But they were working in a system that was bound to defend convictions, uphold finality, and honor precedence. In this case, what the FBI and DOJ did was to lift procedural bars and statutory time bars to let people come back in. They agreed to offer mitochondrial DNA testing in cases where physical evidence remained. But at the same time, this is a story about how agencies limited disclosure of past problems because of the adversarial system. There was a comparison to airplane crashes, and analogies are all over the place, but we’ve also heard the defense community say, “well, when an airplane crashes, we have an NTSB; there’s zero tolerance for error, but there’s an effort to find the root cause of every error.” In medicine, people talk about grand rounds, that error is kept within the community—it’s private. But, within that community, there’s a ruthless pursuit of finding out what went wrong. Obviously, these are systems where there’s not somebody on the other side. The goal is safety; the goal is to save lives. There’s not someone trying to convict somebody and somebody else trying to defend them.

And how do you square that circle? Which is, I think, the problem that you heard the last panel go deep into. But what’s striking to me is how many people have worked to strengthen the scientific integrity of the forensic science system in the United States, especially in law enforcement, despite any professional toll for themselves or how long it might take. Dwight Adams, the former FBI Lab Director, and Howd Headman, the former Hair Unit Chief, suggested to me that it was a good idea to retest using DNA, if possible, in all cases where [hair] evidence was some of the only evidence leading to conviction. It was FBI agents and scientists who pushed through efforts to test a sample of clothes to FBI hair cases using DNA. They are the ones who developed the law enforcement DNA testing protocol for hair. It was examiners who wrote frankly and honestly about what they believed had gone wrong or been done wrong, about both what Malone did and about the lack of openness. One of them said maybe
a cover up (that was my word), but this was her answer, “Maybe it wasn’t the intent, but it did seem to look that way. It was too controlled by the FBI.” It was Assistant U.S. Attorney Michael Ambicino in D.C. who pushed for the hair case review in the district that the Public Defender Service insisted and demanded and held their feet to the fire.

And again it was the FBI and DOJ who committed to a review of bullet lead analysis with the Innocence Project and NACDL the last time that they had a technique that was based on pattern evidence where they thought that there were industry standards and manufacturing practices that they could show this batch of bullets—this bullet at a crime scene—we know could come from this batch of bullets alone that we found in the defendant’s closet, in the suspect’s closet. It was only after the National Academy of Sciences showed that actually there was too much signal to noise, that the chemical trace signatures that the FBI built—this unique among the world—had pioneered the development of this science after the JFK assassination, turned out twenty-five years later that you couldn’t conclude anything more than it was signal to noise.

So, I’ll close with some thoughts from a reader. After our stories ran, he wrote in a letter to the editor: “I was reminded of the importance of investigative journalism, inspectors general, public defenders, public interest groups, and the Freedom of Information Act. Without the supporting roles of each, we might not be educated and enlightened to the degrees we are now about law enforcement’s curious treatment of forensic evidence when lives can hang in the balance.” And the writer went on to call for a continued attention, and really a discussion like we’re having now, of problems. “Anything less would be unforgivable in a democratic nation that prides itself on the enduring, nonnegotiable principles of liberty, equality, and justice for all.”

And thank you to our hosts and everyone here for upholding those values and continuing that conversation.