

Expert Testimony in Criminal Proceedings: Questions *Daubert* Does Not Answer

Margaret A. Berger*

INTRODUCTION

I strongly believe that we need a very stringent standard of proof in criminal cases. I do not think, however, that *Daubert v. Merrell Dow Pharmaceuticals, Inc.*¹ has been productive in effectuating this goal. In civil cases, courts engage in rigorous gatekeeping and often exclude plaintiffs' experts because the theory underlying their testimony has not been adequately validated. But I see no sign of a parallel approach in criminal cases even when there are problems with the assumptions on which the prosecution's expert testimony rests.² I do not think, however, that more stringent attention to the reasoning of *Daubert* would greatly improve matters. *Daubert* overemphasizes how the data underlying the expert's opinion was produced and distracts courts and counsel from carefully analyzing what the evidence proves, and how it is being used. *Daubert* stresses the medium over the message. In criminal cases, *Daubert* may therefore be counterproductive by diverting attention from other evidentiary principles that should be considered in handling expert proof.

I reached these conclusions after looking at expert testimony dealing with forensic identification, the field that encompasses a group of markedly different techniques that have in common the objective of matching a sample associated with the defendant (or victim) to a sample found at the crime scene.³ I limited myself to

* Suzanne J. and Norman Miles Professor of Law, Brooklyn Law School. I wish to thank David H. Kaye for his helpful comments.

¹ 509 U.S. 579 (1993).

² For example, see the following cases which refused to exclude the prosecution's experts despite extensive defense challenges: *United States v. Llera Plaza*, 188 F. Supp. 2d 549 (E.D. Pa. 2002) (fingerprints); *United States v. Prime*, 220 F. Supp. 2d 1203 (W.D. Wash. 2001) (handwriting).

³ To make this easier to read, I will assume throughout this comment that the samples being matched relate to the defendant though, of course they might relate to the victim, as, for instance, bloodstains on the defendant's clothing that match the victim's blood. Sometimes, but rarely, a defendant might seek to introduce this kind

considering some of these techniques for two reasons. First, forensic identification seemed the category of proof with regard to which it would be easiest to discern and to evaluate the impact and usefulness of the Supreme Court's approach to expert testimony. Whether a particular forensic technique can in fact match two tangible samples would seem a question far easier to answer than determining the reliability of other expert proof that cannot be empirically tested because of ethical considerations⁴ or lack of knowledge.⁵ Second, the forensic identification category includes DNA profiling which is undoubtedly our "gold standard" of expertise.⁶ Understanding why DNA evidence has this distinction would, I thought, be helpful in providing a model against which other identification techniques could be judged.

I. DNA EVIDENCE

A look at how DNA typing works and how it became recognized as the preeminent forensic tool in less than twenty years⁷ explains why the legal and scientific communities have so much confidence in DNA as an identification technique. DNA evidence's road to admissibility differed from that of other forensic identification techniques. DNA typing is not like the other forensic specialties which originated within the law enforcement community and whose sole purpose is to facilitate investigations and prosecutions. Forensic DNA typing is the by-product of cutting-edge science. The underlying theory on which it rests—that no individuals other than identical twins have identical DNA profiles—had already been universally ratified by the scientific community before DNA's forensic

of evidence which may raise constitutional issues that I will not address.

⁴ Obviously, one cannot set up controlled experiments to determine whether abused subjects develop various syndromes or repressed memories.

⁵ Perhaps some day we will have physiological markers for liars and those who do not know right from wrong that will lead to a very different type of expert testimony than that provided by psychiatrists battling over an insanity defense, or seeking to testify to predictions of dangerousness.

⁶ DNA has been more thoroughly scrutinized than any other field of criminal investigation. See NATIONAL COMMISSION ON THE FUTURE OF DNA EVIDENCE, THE FUTURE OF FORENSIC DNA TESTING: PREDICTIONS OF THE RESEARCH AND DEVELOPMENT WORKING GROUP 7 (Nat'l Inst. of Justice 2000) [hereinafter FUTURE OF FORENSIC DNA TESTING].

⁷ The first forensic use of DNA testing occurred in England in 1986. See NATIONAL COMMISSION ON THE FUTURE OF DNA EVIDENCE, POSTCONVICTION DNA TESTING: RECOMMENDATIONS FOR HANDLING REQUESTS 1 (Nat'l Inst. of Justice 1999) [hereinafter POSTCONVICTION DNA TESTING]. The first United States appellate court opinion on the admissibility of DNA evidence was issued in 1988 in *Andrews v. State*, 533 So. 2d 841 (Fla. Dist. Ct. App. 1988).

potential was recognized.⁸ And every step of the way, scientists monitored the use of DNA typing as a means of forensic identification.

From the beginning, eminent scientists agreed to testify as expert witnesses⁹ and to serve on committees looking at forensic issues, perhaps for fear of what lawyers would do without proper guidance. The National Academy of Sciences issued a report on DNA in 1992, only six years after the first known use of DNA testing in a criminal investigation anywhere in the world.¹⁰ A second report followed in 1996 because the field was changing so rapidly.¹¹ The scientists' critiques and criticisms led to numerous improvements in DNA technology and caused some courts to exclude DNA evidence until the flaws were corrected.¹²

In order to comprehend the evidentiary issues posed by other forensic identification techniques, it is important to understand what DNA typing proves and how input from the scientific community amplified the probative value of a match between the defendant's DNA and a crime scene sample. Forensic DNA testing does not require matching the entire genetic code found in the two samples from the crime scene and a suspect. Rather, the examiner determines whether a tiny percentage of the available DNA sequence is identical at a number of specific loci or markers shown to be independent and variable.¹³ A match therefore means only that the defendant is a possible source of the crime scene sample. Over time, two developments in particular enhanced the probative value of a

⁸ COMMITTEE ON DNA FORENSIC SCIENCE: AN UPDATE, NATIONAL RESEARCH COUNCIL, THE EVALUATION OF FORENSIC DNA EVIDENCE 9 (Nat'l Academy Press 1996) ("DNA typing . . . is based on a large body of scientific principles and techniques that are universally accepted.") [hereinafter NRC II].

⁹ In *People v. Castro*, 144 Misc. 2d 956, 957 (N.Y. Sup. Ct. 1989), the court conducted a *Frye* hearing over a 12-week period that resulted in a transcript of over 5000 pages; it heard from numerous scientists, including Dr. Eric Lander of the Whitehead Institute for Biomedical Research at MIT, before concluding that evidence could not be admitted because generally accepted scientific techniques for obtaining reliable results had not been employed.

¹⁰ COMMITTEE ON DNA TECHNOLOGY IN FORENSIC SCIENCE, NATIONAL RESEARCH COUNCIL, DNA TECHNOLOGY IN FORENSIC SCIENCE (Nat'l Academy Press 1992).

¹¹ See *supra* note 8.

¹² DAVID L. FAIGMAN ET AL., SCIENCE IN THE LAW, FORENSIC SCIENCE ISSUES § 11-1.2.1 (2002).

¹³ Three results are possible: (1) a match; (2) a nonmatch that excludes the defendant as the source of the crime scene sample, but does not necessarily prove defendant's innocence; depending on the circumstances he might have committed the crime without leaving a crime scene sample; and (3) inconclusive results caused by not enough DNA being available for testing or because the crime scene sample contained a non-analyzable mixture of DNA from two or more individuals.

match. Matching became less subjective when laboratories shifted, as most have done, from the RFLP technology that was originally used to newer methods.¹⁴ Second, far more loci are now used in testing than originally, thereby making it far less probable that two persons would match at all. In 1994, testing usually involved looking at three to five DNA sequences.¹⁵ The FBI now uses 13 core STR sequences in its Combined DNA Index System (Codis) database of convicted offenders.¹⁶

The strength of the inference that two identical DNA profiles come from the same individual cannot be determined without knowing the frequency of defendant's genetic profile in the general population. Some courts were reluctant to admit the results of DNA profiling until the National Academy of Sciences' 1996 report concluded that the statistics being used were scientifically sound.¹⁷ Since then scientists have obtained additional frequency data for numerous population groups and subgroups so that it has now been calculated that the most common profile has an estimated frequency of less than one in ten billion when the 13 STR loci are used.¹⁸ Although testifying experts still may dispute the appropriate probabilities and argue about how they should be expressed, the number of genetic markers now used in forensic typing means that the likelihood that two matching profiles came from the same person or an identical twin approaches certainty, provided that the samples were properly collected and analyzed.

As soon as DNA was put to forensic use, issues arose about laboratory performance. Critics had long complained about the sorry state of American crime labs,¹⁹ but this time some changes were instituted. Quality assurance protocols were designed to maximize the prospect of accurate and reliable testing. In addition, laboratories adopted stringent quality control protocols that require them to monitor and document their performance so that they can

¹⁴ For a discussion of problems that arose in interpreting results when using RFLP technology, see David H. Kaye & George F. Sensabaugh, REFERENCE MANUAL ON DNA EVIDENCE 516-20 (2000).

¹⁵ JUDITH A. MCKENNA ET AL., REFERENCE GUIDE ON FORENSIC DNA EVIDENCE, REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 283 (Fed. Judiciary Ctr. 1994).

¹⁶ POSTCONVICTION DNA TESTING, *supra* note 7, at 28; *see also* FUTURE OF FORENSIC DNA TESTING, *supra* note 6, at 19-20.

¹⁷ *See* NRC II, *supra* note 8, at 52.

¹⁸ FUTURE OF FORENSIC DNA TESTING, *supra* note 6, at 19.

¹⁹ Paul C. Giannelli, *The Abuse of Scientific Evidence in Criminal Cases: The Need for Independent Crime Laboratories*, 4 VA. J. SOC. POLICY & L. 439 (1997); *see also* Paul C. Giannelli, *The Admissibility of Laboratory Reports in Criminal Trials: The Reliability of Scientific Proof*, 49 OHIO ST. L.J. 671 (1988).

verify compliance with their quality assurance objectives.²⁰ Periodic proficiency testing of laboratory personnel is part of this process. Of course, no human endeavor can be error-free, but by requiring a laboratory to document extensively everything it does, the protocols prevent many mistakes and make it easier to detect mishaps that do occur.

Daubert and DNA match very well. Indeed, it has been suggested that the advent of DNA profiling may have paved the way for the Supreme Court's opinion in *Daubert*.²¹ Criteria which Justice Blackmun saw as integral to the scientific method—testing, peer review, publication and standards—all played a role in facilitating the technical advances that placed DNA profiling on such firm ground.²² What needs to be remembered, however, is that this evidence is so powerful because all persons other than identical twins have a unique genetic pattern and the enormous variability of DNA makes it possible to distinguish among individuals. It is the nature of nuclear DNA that gives DNA evidence its enormous probative value. We should not confuse the process of validation with the inferences that the evidence permits. Although courts are now so convinced that DNA evidence is admissible that they generally treat all challenges as going to the weight of the evidence,²³ there may, nevertheless, be instances when DNA evidence should be found inadmissible or insufficient to convict, as when fraud has infected a particular laboratory.²⁴ Expert proof that satisfies *Daubert* should not be immune to other evidentiary challenges.

²⁰ Quality assurance refers to a program conducted by a laboratory to ensure the accuracy and reliability of the tests it performs. Quality control refers to activities the laboratory undertakes to monitor, document, and verify laboratory performance. In other words, a quality control program seeks to demonstrate that the laboratory is meeting its quality assurance objectives. See Charlotte J. Word, *The Future of DNA Testing and Law Enforcement*, 67 BROOK. L. REV. 249 (2001).

²¹ See David L. Faigman, *The Tipping Point in the Law's Use of Science: The Epidemic of Scientific Sophistication That Began with DNA Profiling and Toxic Torts*, 67 BROOK. L. REV. 111 (2001).

²² For a discussion of why no definite error rate has been established or can be established for DNA testing, see Margaret A. Berger, *Laboratory Error Seen Through the Lens of Science and Policy*, 30 U.C. DAVIS L. REV. (1997).

²³ FAIGMAN ET AL., *supra* note 12, at 688.

²⁴ *Houston's Troubled DNA Crime Lab Faces Growing Scrutiny*, N.Y. TIMES, Feb. 9, 2003, at 20 (DNA testing suspended at Houston Police Department's crime laboratory and review ordered of ninety convictions based on DNA testing; allegations that unit is corrupt or incompetent or both).

II. MICROSCOPIC HAIR ANALYSIS

Microscopic hair analysis has existed since the turn of the last century. Until recently forensic examiners regularly relied upon this technique to identify the defendant through hairs found at the crime scene. This mode of examination, however, has come under a good deal of attack since the advent of DNA testing and the decision in *Daubert*. It is now known that evidence obtained through microscopic hair analysis was admitted in a considerable percentage of the cases in which courts vacated convictions on the basis of postconviction DNA testing.²⁵ Courts, commentators, and defense counsel have questioned whether enough studies were conducted to comply with the dictates of *Daubert*.²⁶

Unlike DNA, which is the same throughout the body's cells, hairs taken from the same individual differ from each other. In DNA analysis, a match means that the examiner has determined that the DNA being sampled is identical in the two samples being compared. By contrast, in microscopic hair analysis, there are no true matches. The technician looks at a collection of hairs taken from the defendant, which differ from each other, and the hair from the crime scene. Then, on the basis of a number of microscopic features, the technician determines whether the questioned hair is or is not consistent with that of the defendant.

Since *Daubert*, some research has been conducted with regard to this identification technique.²⁷ We now have the FBI study,²⁸ which

²⁵ BARRY SCHECK ET AL., ACTUAL INNOCENCE: FIVE DAYS TO EXECUTION AND OTHER DISPATCHES FROM THE WRONGLY CONVICTED 166 (2000) (expert testimony about hair analysis in cases leading to DNA exonerations). In December 2002, the FBI issued a report on the hair analysis that had been introduced in the case of a Montana defendant who was cleared by DNA testing after spending fifteen years in prison for raping an 8-year-old girl. 2 *States to Review Lab Work of Expert Who Erred on ID*, N.Y. TIMES, Dec. 19, 2002, at A24. The FBI's trace evidence unit concluded that the crime scene samples did not match samples provided by the defendant; although at the time of the trial, the prosecution's expert, the director of the Montana Crime Laboratory, had testified that the chances that either set of hairs found at the scene were not those of defendant was 1 in 100 and that since head and pubic hairs look different, "it's a multiplying effect, it would be 1 chance in 10,000." *Id.* In an interview after the defendant's exoneration, the expert who has for the past thirteen years worked as a forensic examiner for the Washington State Police, admitted that there were no studies that would authorize quantifications of the kind he used, but that his testimony was based on his experience with hundreds of hair samples. *Id.*

²⁶ See Paul C. Giannelli, *The Supreme Court's "Criminal" Daubert Cases*, 33 SETON HALL L. REV. 1071 (2003).

²⁷ See Clive A. Stafford Smith & Patrick D. Goodman, *Forensic Hair Comparison Analysis: Nineteenth Century Science or Twentieth Century Snake Oil?*, 27 COLUM. HUM. RTS. L. REV. 227 (1996) (extensively analyzing and critiquing studies conducted by Barry Gaudette).

Richard Friedman discusses at some length in this symposium,²⁹ in which FBI examiners compared for accuracy the results of microscopic DNA analysis with those obtained through mitochondrial DNA analysis. Friedman states that the likelihood ratio for this evidence is higher than that for blood type A evidence. By this he means, if I understand him correctly, that an A type blood match increases by 2.5 the likelihood that defendant is the perpetrator while a finding that two hairs come from a common source increases the odds of defendant being the culprit by 2.89. Consequently he concludes that microscopic hair analysis evidence should be admitted unless the witness makes a misleading statement about what the evidence shows.

Though I too agree that this testimony, flawed as it is, should at times be admitted, I would approach this problem somewhat differently. I do not think that likening microscopic hair results to type A blood tests is an apt comparison. The latter result is not a statement about accuracy—we assume that lab technicians can correctly determine that someone has type A blood—but the likelihood of that characteristic being found in the relevant population. We have no such scientifically accepted statistics with regard to microscopic hair examinations; the few studies that purport to establish the frequency with which hair patterns are distributed are generally viewed as flawed. That to me is the heart of the problem with this identification technique. We have some proficiency results but no way of assessing the probative value of testimony that defendant's hair is consistent with hair found at the crime scene.

Now, of course, there is no rule of evidence that states that evidence of unknown probative value is inadmissible. But we do have a rule, Rule 403, that states that probative value must be balanced against other specified factors to determine admissibility.³⁰ It should not be forgotten that *Daubert* itself acknowledges the applicability of Rule 403, and adds a comment of Judge Weinstein's:

Expert evidence can be both powerful and quite misleading

²⁸ Max M. Houck & Bruce Budowle, *Correlation of Microscopic and Mitochondrial DNA Hair Comparisons*, 47 J. FORENSIC SCI. 964 (2002).

²⁹ See Richard D. Friedman, *Squeezing Daubert Out of the Picture*, 33 SETON HALL L. REV. 1047 (2003).

³⁰ FED. RULE EVID. 403 states:

Although relevant, evidence may be excluded if its probative value is substantially outweighed by the danger of unfair prejudice, confusion of the issues, or misleading the jury, or by considerations of undue delay, waste of time, or needless presentation of cumulative evidence.

Id.

because of the difficulty in evaluating it. Because of this risk, the judge in weighing possible prejudice against probative force under Rule 403 of the present rules exercises more control over experts than over lay witnesses.³¹

How would a Rule 403 approach to microscopic hair evidence work? Whatever probative value the judge assigns to a microscopic hair identification should be discounted by the possibility of false positives, which is probably even higher when specially trained FBI examiners do not make the comparisons.³² On the other side of the scale, the judge must assess the extent to which the jury is misled if the expert expresses an opinion about the probative value of the microscopic hair analysis. We know from cases and news accounts that hair examiners often embellish their testimony with probability estimates based on their personal experience for which there is no empirical basis.

Then there is unfair prejudice. The original Advisory Committee Note to Rule 403 speaks of the risk of “inducing decision on a purely emotional basis.” I think this condition is fulfilled in a rape case in which there is no DNA evidence when the prosecution introduces a hair found on the victim and calls an expert who testifies that this is a pubic hair which is consistent with the defendant’s. Real evidence can have a powerful effect. The jury will overestimate the evidence’s shaky probative value, especially if it is the only evidence that ties the defendant to the scene.³³ We will never know whether it was microscopic hair analysis testimony—now conceded to have been inaccurate—that distracted the jury in the Central Park Jogger case from concerning itself with the vast inconsistencies in the defendants’ confessions.³⁴ Certainly pubic hair testimony is graphic enough to make a jury sit up.³⁵

Take for instance, a case like *State v. Butler*,³⁶ in which defendant was charged with sodomizing a 16-year-old boy. The perpetrator of

³¹ *Daubert*, 509 U.S. at 595.

³² See *supra* note 25.

³³ See 2 JOHN HENRY WIGMORE, EVIDENCE § 1157 (2d ed. 1923) (“[T]here is a natural tendency to infer from the mere production of any material object, and without further evidence, the truth of all that is predicated of it.”).

³⁴ See Jim Dwyer & Kevin Flynn, *New Light on Jogger’s Rape Calls Evidence into Question*, N.Y. TIMES, Dec. 1, 2002, at 1.

³⁵ See JOHN TUCKER, MAY GOD HAVE MERCY: A TRUE STORY OF CRIME AND PUNISHMENT 75 (1998), for an account of the Coleman trial, in which the author reports that the trial judge thought the pubic hair testimony had the greatest affect on the jury.

³⁶ 24 S.W.3d 21 (W.D. Mo. 2000) (*en banc*), *vacated*, 2003 WL 41708 (W.D. Mo. 2003) (conviction vacated for ineffective assistance of counsel).

this crime had approached the victim and a 14-year-old at a mobile home court and asked their help in getting a boat out of storage. He then took the boys to a secluded area of the mobile park where the assault took place. Neither boy managed to get a good look at the assailant. The defendant was arrested 20 months after the attack. His first trial ended in a mistrial.

At the second trial in 1996, the boys could not identify the defendant by sight or voice. They gave a general description of a white assailant of average height and weight that more or less tallied with defendant's appearance. Their estimate of age also corresponded with that of the defendant. The defendant, who had no alibi, was a resident of the mobile trailer park and therefore could have known of the isolated area in which the boys were attacked.

A forensic chemist employed by the state regional crime lab testified that she was unable to find "any significant differences" between one head hair found on the 16-year old's T-shirt and one pubic hair taken from his underwear and samples collected from the defendant "that would cause her to exclude defendant as the source of the unidentified hair."³⁷ She acknowledged that neither she nor the forensic community were able to positively identify individuals based on hair comparison.³⁸ She also testified, however, that both the samples from the defendant and the crime scene samples "had spots on the medulla" and that she had done over 1200 examinations without ever previously finding such spots.³⁹ She also stated that both the head and pubic hair matched, and believing that two matches were twice as meaningful, she therefore felt there was a "very strong probability" that the crime scene hairs came from defendant, and that that the hairs came from defendant "within a reasonable degree of certainty."⁴⁰ Defendant testified, denied committing the trial, and offered a rather unlikely story about how his hairs could have been transferred to the victim in the mobile home park's swimming pool. Defendant was convicted and sentenced to 107 years.⁴¹

³⁷ *Id.* at 47.

³⁸ *Id.* at 46.

³⁹ *Id.* at 32.

⁴⁰ *Id.* at 37.

⁴¹ On appeal, the defense argued that the evidence was insufficient to convict but a majority of the court held that since counsel had failed to object to the hair evidence it had to be considered in ruling on sufficiency, and, therefore, there was sufficient evidence to convict. *Butler*, 24 S.W.3d 21. In 2003, a panel of the intermediate appellate court found that defense counsel's failure to object to the hair evidence was not "reasonable trial strategy" and reversed and remanded for ineffective assistance of counsel. *Butler*, 2003 WL 41708.

Why I do not simply say, as I think Paul Giannelli might,⁴² that microscopic hair analysis should never be admissible; it is unreliable because it has not been tested adequately to satisfy *Daubert*? In the first place, as the FBI study shows, trained examiners do have some proficiency in comparing samples accurately, especially when the evidence is used for purposes of exclusion, a use I discuss below. For purposes of inclusion, however, the probative value of the testimony may be substantially outweighed by undue prejudice and/or the possibility of misleading the jury. The prejudice has to be evaluated in the “evidentiary context”⁴³—it may be considerably less than in *Butler* if the microscopic hair analysis is offered in a dissimilar kind of case with a different type of hair and additional evidence to evaluate.⁴⁴ But the main reason why I would not adopt a *per se* rule excluding microscopic hair analysis testimony is because I cannot find a principled reason to distinguish this class of evidence from other forensic identification testimony.

III. MITOCHONDRIAL HAIR ANALYSIS

Why even discuss microscopic hair analysis when mitochondrial or mtDNA is now available to analyze hair that contains insufficient nuclear DNA to be subjected to nuclear DNA testing?⁴⁵ Since 1996, it has been possible to examine DNA obtained from the mitochondria of a cell rather than from the nucleus. Unlike nuclear DNA typing which samples the DNA found at selected loci, mtDNA typing compares the entire sequence of mitochondrial DNA found in the samples being compared.⁴⁶ Like nuclear DNA typing, MtDNA was initially developed for non-forensic uses and has been tested and peer reviewed. Laboratories have adopted protocols and standards to ensure that contamination does not occur, and examiners are required to undergo proficiency tests.⁴⁷ The FBI maintains a database against which it checks the MtDNA sequence in the case under

⁴² See Giannelli, *supra* note 26, at 1075-76 (though he does not quite say so).

⁴³ *Old Chief v. United States*, 519 U.S. 172, 183 (1977) (suggesting that rather than viewing an item of evidence “as an island,” when undertaking Rule 403 balancing a judge should “take account of the full evidentiary context of the case as the court understands it when the ruling must be made”).

⁴⁴ For instance, hairs on a mask used in a bank robbery or hairs found on the driver’s seat in a vehicular homicide prosecution in which defendant claims he was the passenger, not the driver.

⁴⁵ Nuclear DNA analysis usually can not be done unless the hair has roots.

⁴⁶ FAIGMAN ET AL., *supra* note 12, § 11-2.1.2.

⁴⁷ See *State v. Pappas*, 256 Conn. 854 (2001) (containing an elaborate discussion of mtDNA technology).

analysis. In *State v. Pappas*,⁴⁸ a Connecticut case, the FBI agent testifying for the prosecution stated that “although the most common MtDNA type probably has a population frequency of 4 percent, the database is not yet large enough to know the population frequency of rare types, that is types that have not been seen in the data.”⁴⁹ In the case at issue he concluded that approximately 99.75% of the Caucasian population could be excluded as the source of the crime scene sample.⁵⁰

Despite the defense expert’s objection to the FBI validation studies and statistics, the Connecticut court found that Daubert’s criteria were satisfied.⁵¹ From all this it appears that *Daubert*’s criteria are satisfied, as indeed the Connecticut court found. As compared to microscopic hair analysis, mtDNA typing is a far more objective technique for matching that has been vetted by the scientific community, that receives the same attention as nuclear DNA with regard to appropriate laboratory techniques (the same labs do both types of DNA tests), and that is supported by a database that allows for estimates of the frequency of the hair pattern in the relevant population. There is only one problem. Unlike nuclear DNA testing which identifies a profile that is highly likely to be unique when matches occur at each of the 13 STR loci, mtDNA testing cannot unequivocally achieve such singular identifications. All descendants in the maternal line share a common mtDNA profile. In other words, barring mutations which are not very common,⁵² all siblings, cousins and distant cousins will have an identical mtDNA profile if they are descended from the same maternal ancestor.

Suppose that in a case like *State v. Butler*, mtDNA testing established that the crime scene hairs match the defendant’s hairs. The only evidence in the case other than defendant’s familiarity with the scene of the attack was the fact that his physical characteristics were not inconsistent with the very general description provided by the victims. Of course, siblings and maternal cousins may share a family resemblance as well as a mtDNA profile, and they may reside

⁴⁸ 256 Conn. 854 (2001).

⁴⁹ *Id.* at 872.

⁵⁰ *Id.* at 873.

⁵¹ *Id.* at 874; *see also* *State v. Council*, 515 S.E.2d 508 (S.C. 1999) (death penalty case).

⁵² FUTURE OF FORENSIC DNA TESTING, *supra* note 6, at 48 (“[E]volutionary studies have estimated that the average fixed mutation rate for the mtDNA control region is one nucleotide difference per 300 generations, or one difference every 6,000 years. Consequently, one would not expect to observe many examples of nucleotide differences between maternal relatives.”).

or visit in the same community. Therefore, in a given case, the four per cent frequency rate that relates to the entire population pool may be very misleading.

In the early days of nuclear DNA testing, commentators were concerned about the ramifications of kinship because close relatives, especially full siblings, have more genes in common than non-relatives and may match at the loci being sampled.⁵³ With the larger number of loci now being sampled, the chance of a match occurring at each and every locus is reduced substantially. However, nothing can alter the reality that maternal relatives share the same mtDNA profile. We may therefore need to reconsider some of the suggestions that were made in the context of nuclear DNA testing for dealing with the possibility that a sibling of the defendant may have committed the charged crime.⁵⁴ For instance, should the prosecution have the burden to account for defendant's relatives?⁵⁵ This could be a much more difficult task with mtDNA testing than with nuclear testing because of the far larger numbers that might have to be accounted for. Can the prosecution compel maternal relatives to provide hair to be tested? May that hair also be tested by microscopic hair analysis, which the FBI claims can often distinguish between samples from different individuals that are maternally related.⁵⁶ Can one put the burden on the defendant to bring possible maternal relatives to the court's attention, or is this an impermissible shifting of a burden to the accused? Should mtDNA evidence be considered insufficient in a case in which the prosecution introduces no other, or virtually no other evidence, merely because of the possibility that a maternal relative exists? It is, of course, entirely possible that a

⁵³ See FAIGMAN ET AL., *supra* note 12, § 11-2.6.2[2] (“[T]he probability that an untested brother (or sister) would match at four loci—with alleles that each occur in 5% of the population—is about 0.006.”).

⁵⁴ Recommendation 4.4 of the NRC II REPORT, *supra* note 8, at 6, provides: “If possible contributors of the evidence sample include relatives of the suspect, DNA profiles of those relative should be obtained.” The Report does not explain how to determine or who determines the identity of “possible contributors” or whose obligation it is to obtain DNA profiles.

⁵⁵ See Richard Lempert, *Some Caveats Concerning DNA As Criminal Identification Evidence: With Thanks to the Reverend Bayes*, 13 CARDOZO L. REV. 303, 312-13 (1991) (discussing a number of solutions: 1) putting burden on defendant “to make plausible case that a relative committed the crime” but this runs counter to general principle that burden is on state and defendant would be incapable of doing this unless provided with funds for an investigation; 2) requiring state with cooperation of defendant to identify relatives who belong in the suspect population and examine their DNA, but expensive and relatives may be unwilling to cooperate; and 3) ignore the issue if there is strong evidence against the defendant).

⁵⁶ Houck & Budowle, *supra* note 27, at 964, 966 (stating that microscopic hair analysis was far more accurate in making exclusions rather than inclusions).

defendant might not know all his maternal relatives even if they live in the same community.

Clearly this new type of forensic identification technique satisfies *Daubert* and courts are holding accordingly.⁵⁷ Although admissibility is not a problem this does not dispose of the question of what the evidence proves. MtDNA evidence has the potential to raise all the issues of the blue bus hypothetical⁵⁸ so beloved of Evidence professors—but in a criminal context. *Daubert* does not supply the answer.

IV. HANDWRITING

I certainly do not want to discuss handwriting identification at length at a symposium run by Professors Risinger and Denbeaux at which Professor Saks is a speaker. I do know true expertise when I see it. But I do want to make a few points. First, their seminal paper that exposed the flaws of handwriting expertise predates *Daubert*.⁵⁹ Although *Daubert* undoubtedly heightened judges' sensitivity to the need to scrutinize expert testimony, and led to the current climate in which handwriting testimony is no longer universally admitted, in theory this result could have been reached under Rule 702's original helpfulness test. Second, even if we assume that examiners cannot make accurate matches and that there is insufficient data on the significance of a match, the real problem is what should a court do?

The issue is complicated because the questioned handwriting about which the expert wishes to testify is an item of real proof. Rule 901 of the Federal Rules of Evidence contains a number of provisions that arguably may apply. Rule 901(b)(9) gives as an example of identification that conforms with the rule's requirements, "evidence describing a process or system used to produce a result and showing that the process or system produces an accurate result." This is precisely what the handwriting expert claims he is doing—that he is instructing the jurors on a process that produces an accurate result. The problem is that the standard of proof that Rule 901 instructs the

⁵⁷ See, e.g., *United States v. Coleman*, 202 F. Supp. 2d 962 (E.D. Mo. 2002).

⁵⁸ *Smith v. Rapid Transit, Inc.*, 317 Mass. 469 (1945) (whether purely statistical proof—in this case that defendant operated the only bus line in town—would be sufficient to prove that defendant's bus caused the plaintiff's accident. See Daniel Shaviro, *Statistical-Probability Evidence and the Appearance of Justice*, 103 HARV. L. REV. 530, 538 (1989) (stating that the possibility of unwarranted convictions "suggests raising the burden of proof for all cases. It does not support a special rule for statistical probability cases.").

⁵⁹ D. Michael Risinger et al., *Exorcism of Ignorance as a Proxy for Rational Knowledge: The Lessons of Handwriting Identification "Expertise"*, 137 U. PA. L. REV. 731 (1989).

court to apply is evidence “sufficient to support a finding,” the conditional relevancy standard that also appears in Rule 104(b). This is a standard that is considerably lower than the Rule 104(a) preponderance of the evidence standard assigned the gatekeeping judge in *Daubert*. In addition, Rule 901(b)(3) explicitly provides that an object may be identified through the testimony of an expert witness who has compared the item in question with “specimens that have been authenticated.” The Advisory Committee’s Note to Rule 901(b)(3) explains that it meant to reject the common law rule which required a judge to use a higher standard than that specified in Rule 104(b) when deciding whether to admit expert testimony with respect to matching handwriting exemplars.⁶⁰

In *United States v. Saelee*,⁶¹ the court, after hearing from Professor Saks, excluded the entire testimony of the prosecution’s forensic document analyst who had compared hand printing on package labels with exemplars provided by the defendant. The court mentioned the Advisory Committee note but concluded, “Rule 702 and Rule 901 must be read together. Rule 901(b)(3) contemplates testimony by an expert—but before an expert’s testimony can be admitted, it must pass through the gates of Rule 702.”⁶²

Even if we agree that this decision is correct—that Rule 702 trumps Rule 901, although perhaps Rule 901 should say so—that still does not solve the other Rule 901(b)(3) problem. The provision contemplates that a writing may be authenticated through “comparison by the trier of fact . . . with specimens which have been authenticated.” Is the defendant better protected when jurors during deliberation compare the uncrossed t’s or undotted i’s in the crime scene samples and the provided specimens, as they have probably seen done on some TV show, than if the court allows a prosecution expert to testify who is then demolished by Professor Saks? If trained examiners lack proficiency, it would seem that the laypersons who serve on juries must also (and indeed some studies, but not all the studies cited in the *Saelee* case, have so found). What can a court do if the jury convicts in a case in which handwriting is central after the court has excluded the prosecution’s expert? Suppose there is little evidence apart from the handwriting exemplars and the questioned document which the jurors compared? Should the court grant a motion of acquittal or set aside a verdict on insufficiency grounds?

⁶⁰ “Example (3) sets no higher standard for handwriting specimens and treats all comparison situations alike, to be governed by Rule 104(b).”

⁶¹ 162 F. Supp. 2d 1097 (D. Alaska 2001).

⁶² *Id.* at 1107.

Again *Daubert* furnishes no answers.

V. FINGERPRINTING

Looked at through the lens of *Daubert*, fingerprints clearly should not be admissible and yet fingerprint matches obviously are often accurate and corroborated by other evidence. I think it highly unlikely that any court will issue a *per se* ruling that all testimony about fingerprints must be excluded despite the undisputable fact that remarkably little research about fingerprints has ever been done.⁶³

Matching fingerprints is more complicated than sampling DNA because we know that a limited number of patterns made up of four nucleotide bases can be sequenced at each of the loci used in DNA testing. With fingerprints, there seem to be an endless number of possible permutations consisting of loops, whorls, arches and deltas. It may well be possible to identify an individual by comparing complete sets of prints because the abundance of detail probably makes each individual's fingerprint pattern unique, though this has never been scientifically verified.⁶⁴ The forensic issue, however, is whether partial, latent, blurred, and perhaps contaminated fingerprints found at a crime scene can be matched with fingerprints obtained from the accused. And if there were a match, what would such a match prove? Obviously we need research to determine how much of a print fragment, and of what quality, is needed in order for an expert to reach a reliable conclusion about the degree of probability that this is the defendant's print. The fact that fingerprint examiners do well on easy and nonchallenging proficiency tests does not answer this question.

Hopefully, research into the validity of fingerprint matches will be adequately funded and will commence in the near future. But what should be done now? Generic attacks seeking to exclude all fingerprint evidence are bound to fail. Defense counsel who are challenging fingerprint evidence might do better relying less on arguments about how fingerprint evidence is produced, and more on what the evidence proves in their specific case. Instead of trying to show that the lack of research violates *Daubert*, it might be more fruitful to litigate the admissibility of fingerprint evidence with regard to a particular blurred, small print fragment found at a crime scene. A court that will not reject all fingerprint evidence might still find

⁶³ See FAIGMAN ET AL., *supra* note 12, ch. 2.

⁶⁴ In *United States v. Llera Plaza*, 179 F. Supp. 2d 492 (E.D. Pa. 2002), the court took judicial notice of the uniqueness of fingerprints.

that a particular match is unwarranted, as has happened in some handwriting cases. Should the court then permit the jurors to compare photographs of the defendant's fingerprint with photographs of the fingerprint lifted from the crime scene?

CONCLUSION

In criminal cases, unlike civil cases, courts have generally been unwilling to exclude expert proof on the ground that it has not been sufficiently validated. When the evidence is admitted it is then labeled "reliable" because, according to *Daubert*, expert testimony must be reliable in order to be deemed admissible. That should not, however, mean that no further analysis of the evidence is required. Admissibility and sufficiency determinations rest on more than satisfaction of a reliability component; they require careful attention to what the evidence proves and how the trier of fact will use it.

Each of the identification techniques discussed above raises different issues in this respect even though the underlying inference in each instance is the same—that a match makes it more probable that the accused is the perpetrator. Even nuclear DNA testing—the gold standard for expert proof—may, under some circumstances, produce results that are completely wrong. MtDNA testing easily meets *Daubert* standards but may be extremely misleading. Handwriting, microscopic hair analysis, and fingerprinting all suffer from serious flaws, but are not amenable to the simple solution of excluding all such expertise. *Daubert* is not sensitive to these nuances. Furthermore, because an assumption of reliability now accompanies expert proof that is admitted, judges may overestimate the evidence's probative value, thereby making it more difficult for a defendant to succeed on Rule 403 grounds or on an insufficiency motion.

What criminal defendants need in order to deal more effectively with the forensic identification expertise proffered against them is not more *Daubert*, but tools that would enable them to make more cogent evidentiary arguments—better counsel, access to expert assistance and more discovery.