

A Sociological Perspective on the Science of Forensic DNA Testing

*William C. Thompson**

According to Professor Redmayne, disputes over scientific evidence are not just about science.¹ These disputes also involve broader questions, such as who should bear the risk of scientific uncertainty and how certain one must be to offer scientific conclusions to a jury. To understand these disputes, one must look beyond science per se and consider the institutional and social context in which they arise. One must consider sociology as well as science.

I agree fully with Professor Redmayne's analysis and commend him on his insights. My only complaint about his Article is that it fails to carry the sociological analysis far enough. In particular, it fails to distinguish forensic science from what might be called academic science and thereby fails to recognize some important sociological factors that help account for many persistent problems in forensic science. In this Comment, I endeavor to expand upon Professor Redmayne's analysis by offering a sociological perspective on forensic science, demonstrating how this perspective can account for problems with forensic DNA evidence.

I. A SOCIOLOGICAL PERSPECTIVE ON FORENSIC SCIENCE

Forensic scientists play a fundamentally different role in society than do academic scientists. The major imperative of the academic scientist is to advance scientific knowledge — to find truth through the use of the scientific method. Forensic scien-

* Professor of Criminology, Law & Society, University of California, Irvine; J.D., University of California, 1982; Ph.D., Stanford, 1984. The author has published articles about DNA evidence in a number of scientific and legal publications. He has served as co-counsel for several criminal defendants who were incriminated by DNA evidence, including O.J. Simpson.

¹ See Mike Redmayne, *Expert Evidence and Scientific Disagreement*, 30 U.C. DAVIS L. REV. 1027, 1027-28 (1997).

tists also use scientific techniques and seek to find truth in specific contexts. However, forensic scientists' major purpose is to provide a service to a client by answering specific questions about evidence.

The primary clients of the vast majority of forensic scientists are law enforcement agencies. Most forensic scientists are employed directly by law enforcement agencies. Their role in litigation is typically, and often exclusively, to provide evidence in support of criminal prosecutions. Forensic scientists who work in private laboratories may occasionally be employed by criminal defense lawyers. However, the bulk of their work is for law enforcement as well. The major market for commercial laboratories that develop new technology for forensic testing also consists of law enforcement personnel.

The role of forensic scientists as service providers has important implications. Like most service providers, forensic scientists must convince their clients that their services are valuable in order to succeed professionally. This imperative is most obvious for private laboratories, which often advertise and promote their services with commercial zeal. It is also important for employees of government laboratories who gain prestige and justify larger budgets by convincing public officials of the value of their services.

Because they need to sell their services, forensic scientists have incentives to put the best possible face on their work, to promote the impression that their techniques are accurate and reliable and that their conclusions are trustworthy. I believe that these incentives cause forensic scientists to act more like members of a trade guild than participants in a scientific discipline. For example, they may avoid openly raising questions about the reliability of forensic tests, avoid public discussion of technical problems or concerns, and refrain from publicly criticizing the work of other forensic scientists. They may also avoid publishing anything that might reflect negatively on their field, thereby making forensic science journals forums for self-promotion rather than self-criticism.

The problem with these credibility-enhancing efforts is that they are inconsistent with traditional forms of scientific self-scrutiny. This self-scrutiny is needed to maintain the quality of forensic science as a science. Forensic scientists may be slow to identi-

fy and solve problems with their techniques because problems are rarely discussed openly. The insularity of forensic science makes matters worse. The work of individual forensic scientists often receives little or no external scrutiny, therefore errors are often difficult to detect.

The role of forensic scientists as service providers for law enforcement creates another problem as well. There is a tendency for service providers to be coopted, that is, to adopt the goals of their clients as their own. I believe that forensic scientists tend to identify closely with the goals of police and prosecutors; seeing themselves as part of the law enforcement team.

This identification is problematic because the goals of law enforcement sometimes conflict with the goals of scientific objectivity and neutrality. The conflict is usually not severe during the early investigative stage of cases, when police and prosecutors want to find out what happened and need forensic scientists who can provide neutral, objective analysis of physical evidence. There comes a point in most criminal investigations, however, when police and prosecutors think they know what happened. At that point, their primary goal is to achieve justice by convicting the person they believe is guilty. They need forensic scientists who can be effective advocates of the prosecution's scientific theories. Thus, for forensic scientists, the desire to achieve justice by helping their team win in court may conflict with the goal of maintaining scientific detachment and neutrality.

Evidence that forensic scientists sometimes make science subservient to the goals of law enforcement is not hard to find. The most extreme examples are forensic scientists who lie and fabricate data in order to help obtain convictions. The notorious case of Fred Zain² illustrates both the potential for cooptation that exists in forensic science and the evils that follow from it. Zain worked as a forensic serologist for a number of years in the West Virginia State Police Crime Laboratory and thereafter worked for the Bexar County (Texas) Medical Examiner's Office. He had a formidable reputation as the serologist who could find (incriminating) results when no one else could. As it turned out, he achieved his success dishonestly. An investigation

² See *In re West Virginia State Police Crime Lab.*, 438 S.E.2d 501 (W. Va. 1993).

of his work was prompted when Glen Woodall, a man convicted based on serological evidence produced by Zain, was later proven innocent by DNA tests.³ In a special report on Zain's misconduct in over 130 criminal cases, the West Virginia Supreme Court of Appeals stated that his behavior included "overstating the strength of results; . . . reporting inconclusive results as conclusive; [and] repeatedly altering laboratory records."⁴ His misbehavior was "the result of systematic practice rather than occasional inadvertent error."⁵ Zain was also charged with perjury, tampering with government records, and fabricating evidence in Texas.⁶

Zain may have been motivated in part by a desire for professional recognition, but his primary goal was to help prosecutors convict people they thought were guilty. His misconduct thus reflects a context in which making the case against criminal defendants is seen as more important than objective assessment of the evidence. In this context, forensic scientists are valued more for effective advocacy than objectivity or even honesty. Indeed, there was evidence that "supervisors may have ignored or concealed complaints of his misconduct."⁷

Zain's efforts to grease the wheels of justice ended up producing injustice because he sometimes framed the wrong people. Besides Glen Woodall, two other West Virginia men who had been convicted based, in part, on Zain's serological testimony were later exonerated by DNA tests.⁸ An important lesson of this case is that egregious misconduct by a forensic scientist can go undetected (or worse, be tolerated) in the criminal justice

³ See EDWARD CONNORS ET AL., CONVICTED BY JURIES, EXONERATED BY SCIENCE 74-76 (1996) (discussing cases in which persons convicted of crimes on basis of serological evidence are later exonerated by DNA evidence).

⁴ *State Police Crime Lab.*, 438 S.E.2d at 503; see *Court Invalidates a Decade of Blood Test Results in Criminal Cases*, N.Y. TIMES, Nov. 12, 1993, at A20 (stating ruling of West Virginia Supreme Court of Appeals that state police may have fabricated blood tests).

⁵ *State Police Crime Lab.*, 438 S.E.2d at 503 (quoting Judge Halliday, who was quoting report by American Society of Crime Laboratory Directors).

⁶ See CONNORS ET AL., *supra* note 3, at 18 (stating that forensic scientist was charged with perjury, tampering with government records, and fabricating evidence).

⁷ *State Police Crime Lab.*, 438 S.E.2d at 504.

⁸ See CONNORS ET AL., *supra* note 3, at 18, 48-49, 55-57 (discussing cases of William O'Dell Harris and Gerald Wayne Davis, whose convictions were overturned by DNA evidence).

system. "It is sobering to reflect that but for the adventitious appearance of DNA typing, Glen Woodall would still be languishing in prison and Fred Zain might still be sending innocent persons to prison."⁹

Misconduct as serious as Zain's is probably rare, but other forensic scientists are subject to the same institutional pressures and incentives that gave rise to it. While few are likely to acquiesce to those pressures as completely as Zain, many may be influenced in more subtle ways. Those who are not seduced into an advocacy role by their desire to achieve justice may be pushed into that role by supervisors. The recent allegations of misconduct at the FBI Crime Laboratory indicate the nature of the institutional pressures that may come to bear on forensic scientists. According to news reports, laboratory analysts told Justice Department investigators that they were pressured by their supervisors to distort and misrepresent their findings. Additionally, analysts' conclusions were sometimes changed by supervisors to support criminal prosecutions.¹⁰ Similar problems have been reported at other forensic laboratories.¹¹

To summarize, it is my thesis that many problems in forensic science arise from the unique role that forensic scientists play in society and from the institutional context in which forensic scientists work. I believe that these factors foster in forensic scientists a guild mentality and an identification with the goals of law enforcement, and that these attitudes are inconsistent with doing high quality scientific work. I propose that these sociological factors exert a pervasive influence on the way forensic science is done. These factors influence the development and validation of new testing procedures, the interpretation of results, and their presentation in court. The next sections of the Article show how this perspective helps account for problems with forensic DNA evidence.

⁹ *Id.* at xvii.

¹⁰ See Richard A. Serrano, *FBI Lab Probe Yields Reports of Misconduct*, L.A. TIMES, Jan. 30, 1997, at A1 (stating that FBI encouraged subordinates to misrepresent findings).

¹¹ See Michael J. Saks, *Prevalence and Impact of Ethical Problems in Forensic Science*, 34 J. FORENSIC SCI. 772, 775-83 (1989) (discussing ethical state of forensic sciences and providing results of proficiency tests).

II. INADEQUATE VALIDATION

One reason the dispute over the proper statistical characterization of DNA evidence has been difficult to resolve, Professor Redmayne tells us, is that there was true scientific uncertainty on the fundamental issue of the statistical independence of the genetic characteristics examined by DNA tests. His analysis is correct as far as it goes, but it begs the question why there was so much uncertainty on such a fundamental point. According to Professor Redmayne, forensic laboratories used procedures for estimating the frequency of DNA profiles that *assumed* statistical independence for at least five years before sufficient research had been done to “tip the balance of scientific opinion” toward acceptance of these assumptions.¹² Why was this important validation research done only after statistical estimates, premised on the assumption of independence, had been presented in thousands of cases?

Validation of forensic DNA tests has been deficient in other areas as well. The 1992 report of the National Research Council noted that “some testing laboratories initially used methods (for such fundamental steps as identifying patterns, declaring matches, making comparison with a data bank, and correcting for band shifting) that they later agreed were not experimentally supported.”¹³ Important validation studies were done only after

¹² See Redmayne, *supra* note 1, at 1031-38 (stating that forensic laboratories used DNA procedures that assumed statistical independence long before that independence was scientifically acceptable).

¹³ COMMITTEE ON DNA TECHNOLOGY IN FORENSIC SCIENCE, NATIONAL RESEARCH COUNCIL, DNA TECHNOLOGY IN FORENSIC SCIENCE 56 (1992) [hereinafter NRC1]. The National Research Council (NRC) is the principal operating agency of the National Academy of Sciences, a non-profit society of distinguished scholars established by Congressional Charter in 1863 with a mandate to advise the federal government on scientific and technical matters. See *id.* The NRC has published a number of authoritative reports on scientific issues. See *id.* The 1992 report was the product of a two year study by a distinguished committee of scientists, judges, and lawyers. See *id.* at 173-76. The Committee was established in 1990 in order to respond to “a crescendo of questions concerning DNA typing.” See *id.* at vii. A second committee appointed by the NRC issued a report on forensic DNA evidence in 1996. See generally COMMITTEE ON DNA FORENSIC SCIENCE: AN UPDATE, NATIONAL RESEARCH COUNCIL, THE EVALUATION OF FORENSIC DNA EVIDENCE (1996) [hereinafter NRC2] (updating reports on forensic evidence). This report was prepared by a second panel of experts, appointed in 1994 to update the analysis and recommendations of NRC1 with respect to “statistical and population genetics issues in the use of DNA evidence.” See *id.* at 49.

laboratories had established protocols for DNA testing and begun presenting results in court. For example, Cellmark Diagnostics performed research needed to validate its criteria for declaring a "match" between DNA profiles after its match criteria were established.¹⁴

I believe that forensic scientists were slow to validate DNA tests because they saw no need to do so until courts began demanding it. They were satisfied to rely on the assumption of statistical independence, for example, because this assumption made it possible to generate the extreme frequency estimates that make DNA test results so impressive. If one assumes statistical independence, then frequency estimates of one in millions or billions are easy to derive from relatively small data bases. Forensic scientists saw no need to examine such a convenient assumption so long as no one effectively challenged it in court. To those interested in marketing forensic DNA testing, the possibility of population structure was a can of worms better left unopened. It was only when criminal defense lawyers began raising the issue, and a few courts excluded DNA evidence as a result, that an earnest effort was made to study the issue.¹⁵

There is a similar explanation for the slow and inadequate validation of "match" criteria. Forensic scientists were satisfied with vague, ad hoc standards for matching because they allowed maximum flexibility and discretion in interpreting test results. So long as they could get away with declaring matches on an entirely subjective basis, they had little incentive to establish more rigorous criteria. Consequently, it was only after some embarrassing courtroom debacles called attention to the problem¹⁶ that forensic scientists made serious efforts to establish and validate formal match criteria.¹⁷

¹⁴ See William Thompson, *Evaluating the Admissibility of New Genetic Identification Tests: Lessons from the "DNA War"*, 84 J. CRIM. L. & CRIMINOLOGY 22, 46 (1993) (showing criticism of Cellmark for faulty scientific procedures).

¹⁵ Even then, law enforcement officials sought to control the research effort by channeling research funds to scientists known to be sympathetic, denying scientific critics access to data. See *id.*; see also Peter J. Neufeld, *Have You No Sense of Decency?*, 84 J. CRIM. L. & CRIMINOLOGY 189, 189-90 (1993) (discussing problems of unchallenged and unregulated DNA results).

¹⁶ See, e.g., *People v. Castro*, 545 N.Y.S.2d 985, 999 (Sup. Ct. 1989) (holding scientific evidence inadmissible for not complying with generally accepted scientific techniques); see also Thompson, *supra* note 14, at 42-44 (discussing *Castro*).

¹⁷ Most laboratories now have formal criteria specifying how closely the bands in two

The research to validate formal match criteria was often of poor quality. Apparently, the research was designed to provide scientific window-dressing for court hearings rather than useful answers to important questions. For example, in *United States v. Yee*,¹⁸ the first case in which defense counsel gained access to the FBI's validation research, defense experts demonstrated serious scientific deficiencies in the research.¹⁹ The deficiencies led the federal magistrate to comment on the "remarkably poor quality of the Bureau's [validation] work and infidelity to important scientific principles."²⁰ In an article published in 1991, Dr. Simon Ford and I noted the danger that "retrospective validation by a laboratory already heavily invested in a particular technology or procedure will be half-hearted, goal-directed, self-congratulatory, or even deceitful."²¹ The evidence that later emerged in *Yee* provided an apt illustration of this danger.

Inadequate, goal-directed validation of DNA tests appears to be a continuing problem. For example, when concerns were raised that PCR-based DNA tests might be unduly susceptible to contamination, which could produce false results, the FBI research laboratory performed a series of studies to demonstrate that contamination was not a serious problem.²² The FBI scientists reported that it is difficult, if not impossible, to inadvertently transfer DNA from one sample to another, and any cross-contamination that does occur will be detected by controls. These sunny conclusions have been cited in court by forensic scientists from around the country as proof that contamination is not a problem with PCR-based DNA tests.²³ However, these

DNA patterns must align in order for them to be declared a match. As noted below, however, the standards for determining matches still allow a great deal of room for subjective judgment.

¹⁸ 134 F.R.D. 161 (N.D. Ohio 1990), *aff d sub nom.* *United States v. Bonds*, 12 F.3d 540 (6th Cir. 1993).

¹⁹ For a detailed review of the problems, see Thompson, *supra* note 14, at 46-50.

²⁰ *Yee*, 134 F.R.D. at 210. The court also commented on the "troublesome questions about the quality of the Bureau's work." *See id.* at 207.

²¹ William C. Thompson & Simon Ford, *The Meaning of a Match: Sources of Ambiguity in the Interpretation of DNA Prints*, in *FORENSIC DNA TECHNOLOGY* 150-51 (Mark A. Farley & James J. Harrington eds., 1991).

²² Catherine Theisen Comey et al., *PCR Amplification and Typing of the HLA Dqalpha Gene in Forensic Samples*, 38 J. FORENSIC SCI. 239, 239-40 (1993).

²³ *See, e.g.*, *State v. Russell*, 882 P.2d 747, 768 (Wash. 1994) (holding that PCR procedures meet *Frye* standard), *cert. denied*, 115 S. Ct. 2004 (1995).

conclusions were seriously challenged during the O.J. Simpson trial by defense expert Dr. John Gerdes, a molecular geneticist who heads a DNA typing laboratory in Denver.

Gerdes was hired by the defense to review the results of DNA testing in the Los Angeles Police Department (LAPD) Crime Laboratory during the fourteen months prior to the testing in the Simpson case. Regardless of one's views on Simpson's guilt or innocence, Gerdes's conclusions about problems in the LAPD laboratory must be taken seriously. He found that DNA was frequently transferred from one sample to another.²⁴ For example, control samples that were supposed to contain no DNA often produced typeable results on DNA tests. Samples from known individuals, used as exemplars in the laboratory, were sometimes found to have the wrong genotype due to the appearance of "extra alleles" from extraneous DNA. Furthermore, the occurrence of contamination often was not detected by experimental controls designed for that purpose. Gerdes's detailed review of work in the LAPD laboratory thus produced compelling results that show that contamination can be a serious problem in PCR testing. Gerdes's conclusions are strikingly inconsistent with the FBI validation studies' conclusions, raising more questions about the adequacy and quality of the Bureau's work.

A revealing sidelight on Gerdes's conclusions is that no one has yet challenged them. Forensic scientists surely are aware of Gerdes's work because Gerdes was a prominent witness in the highest profile trial ever. Thereafter, Gerdes presented his scientific findings at a major conference attended by over four hundred experts in forensic DNA testing.²⁵ In addition, Gerdes published his findings in the conference proceedings. Gerdes's conclusions present a fundamental challenge to accepted wisdom in the field of forensic DNA testing. Nearly two years have passed since the Simpson trial, and nearly eighteen months have

²⁴ Gerdes's conclusions are reported in J.C. GERDES, PROCEEDINGS OF THE SIXTH INTERNATIONAL SYMPOSIUM ON HUMAN IDENTIFICATION, CRITICISM/CONCERNS REGARDING DNA FORENSIC TESTING SUBSTANTIATED BY EVIDENCE PRESENTED IN THE O.J. SIMPSON CASE (1995).

²⁵ His findings were presented at the Sixth International Symposium on Human Identification, Oct. 14, 1995.

passed since Gerdes published his findings. Yet no one has commented publicly on Gerdes's work. There have been no articles about Gerdes's work in forensic science journals and no papers about it at professional meetings. The community of forensic DNA experts is acting as if Dr. John Gerdes does not exist.

I believe that if forensic science were really a science, the response to Gerdes would be quite different. Experts in forensic DNA testing would be competing with each other to publish the earliest, strongest refutation of Gerdes's conclusions (if they think he is wrong) or to propose the most effective way to solve the problem (if they think he is right). At the very least, forensic scientists would work to better identify circumstances under which contamination is a problem. Instead, the forensic DNA testing community has been silent. Perhaps they are waiting to see whether Gerdes's criticisms will be taken seriously by the courts, fearing that a response might give Gerdes more credibility. If so, they are acting more like a trade guild than a scientific discipline. In my view, Gerdes's criticisms have provided an interesting test of the scientific character of the field of forensic DNA testing and the test has been failed. I predict that no forensic scientist will respond to Gerdes's criticisms unless, or until, an appellate court rules a PCR test inadmissible based on concerns about contamination.

A final example of inadequate validation of a forensic DNA test came to light just recently. It concerns a relatively new procedure known as the D1S80 test.²⁶ Experts at the FBI laboratory, which has been using the test for case work, published several studies purporting to validate the procedure.²⁷ A key issue for validation was whether the analyst can assume that bands detected by the test are from human DNA rather than some

²⁶ See Bruce Budowle et al., *Analysis of the VNTR Locus D1S80 by the PCR Followed by High-Resolution PAGE*, 48 AM. J. HUM. GENETICS 137, 139 (1991) [hereinafter *Analysis of Locus D1S80*] (discussing using allelic data for D1S80 locus); B. Budowle et al., *D1S80 Population Data in African Americans, Caucasians, Southeastern Hispanics, Southwestern Hispanics, and Orientals*, 40 J. FORENSIC SCI. 38, (1995) (discussing differences in allelic data among ethnic groups); A. Kloosterman et al., *PCR-Amplification and Detection of the Human D1S80 VNTR Locus: Amplification Conditions, Population Genetics, and Application in Forensic Analysis*, 105 INT'L J. LEG. MED. 229 (1993).

²⁷ See Budowle et al., *Analysis of Locus D1S80*, *supra* note 26, at 137-44 (discussing several validation tests).

other source, such as bacterial DNA. One of the FBI studies concludes that the test does not pick up bacterial DNA.²⁸ However, a later study by forensic scientists in Spain found that the test picked up extra bands from six of twenty-six species of bacteria commonly found in forensic samples.²⁹ This finding raises concerns about the reliability of the D1S80 test because the appearance of false bands (due to bacterial contamination) could potentially cause a laboratory to declare a false match between samples from different people. Concerns about this problem recently led a trial court in California to exclude the results of a D1S80 test under the *Frye* standard.³⁰ Why did the FBI scientists fail to find this problem? The answer is not entirely clear, but it appears that they concluded bacterial contamination was not a problem after testing only three bacterial species.

III. THE PERSISTENCE OF POOR SCIENTIFIC PRACTICES

I believe the most serious problem with forensic DNA testing, at present, is the use of poor scientific practices for interpreting tests. I am particularly concerned about the heavy reliance of forensic laboratories on the subjective judgment of analysts to resolve ambiguities in test results.³¹ Whether a test is interpreted as a damning incrimination or a complete exculpation may depend entirely on a subjective determination. The analysts making these determinations are not blind to the expected results of the test. They often are in direct contact with detectives and know all about the case (at least from the police perspective). Consequently, there is a danger that the analysts may intentionally or unintentionally be biased toward the police theory of the case when making subjective determinations.

²⁸ See F. Samuel Baechtel et al., *D1S80 Typing of DNA from Simulated Forensic Specimens*, 40 J. FORENSIC SCI. 536, 542-44 (1995) (discussing results of reliability testing of D1S80 procedure).

²⁹ A. Fernández-Rodríguez et al., *Microbial DNA Challenge Studies of PCR-Based Systems Used in Forensic Genetics*, 1996 ADVANCES IN FORENSIC HAEMOGENETICS 177, 177.

³⁰ Order re: Findings on Admissibility of DNA Evidence, *People v. Scott*, No. CR-48638 (Riverside County, Cal. Sup. Ct., Jan. 23, 1997)

³¹ See W. Thompson, *Subjective Interpretation, Laboratory Error, and the Value of Forensic DNA Evidence: Three Case Studies*, 96 GENETICA 153, 165-67 (1995) (discussing subjective interpretation of ambiguous forensic DNA tests).

Of course this problem does not arise in every case. In many cases the results are sufficiently clear that interpretation is straightforward and uncontroversial. The results of DNA tests can properly be divided into three categories: clear inclusions (i.e., comparisons plainly showing a "match" between the genetic characteristics of two samples); clear exclusions (i.e., comparisons plainly showing a difference between the genetic characteristics of two samples); and ambiguous or uncertain comparisons. It is the third category that concerns me.

As an illustration, I refer to the DNA test results in a case that I litigated in 1996. A woman was raped by two men. Police obtained blood samples from two suspects and sent these samples, along with vaginal aspirate from a rape kit, to a forensic laboratory for DNA testing. The laboratory used RFLP analysis, currently the most common DNA testing procedure, to examine the genetic characteristics of the sample at five genetic loci. According to the laboratory report, "DNA banding patterns obtained from the male fraction of the vaginal aspirate demonstrate DNA from two individuals consistent with the patterns obtained from [the two suspects]." The DNA patterns of suspect 1 "occur with a frequency of one in 641,100,000 in the North American Black population" and the patterns for suspect 2 "occur with a frequency of one in 636,500,000 in the North American Black population." The laboratory report gives no indication of any uncertainty about the "match" between the suspects and the vaginal sample, so it would appear that the DNA test provides damning evidence against the two suspects. But let's look at the underlying results.

Figure 1³² shows one of five autoradiograms (autorads) produced by the laboratory to show the DNA banding patterns of the samples tested in this case. This autorad shows DNA profiles for a particular genetic locus known as D4S139. Three lanes of the autorad (on the far left, far right, and middle) display multiple bands called size markers. These bands are produced by fragments of bacterial DNA of known size. They are compared to the bands in the other lanes (which are produced by fragments of human DNA) in order to allow the size of these fragments to be estimated. The band patterns of the victim and the

³² See *infra* p. 1132.

two suspects appear in vertical lanes on the left side of the autorad. Each individual has two bands. The position of these bands within their lanes indicates the genotype³³ (for this locus) of the individual who provided the sample. Because this locus is an area where DNA is polymorphic (variable among individuals), the banding patterns tend to vary from person to person, as can be seen for the victim and two suspects. Banding patterns of the female and male portions of the vaginal aspirate appear on the right side of the autorad, along with a sample from a known individual run as a control. The banding pattern of the female vaginal extract is difficult to see due to a dark smear in its lane which was probably caused by degradation of the DNA.

The key comparison is between the suspects' patterns and the pattern in the male vaginal extract. Two bands corresponding to those of suspect 2 are clearly visible, indicating that he is a possible source of this DNA. Whether bands corresponding to those of suspect 1 are also present is less clear. The two dots on the left side of this lane are felt-tip pen marks placed by the forensic analyst to indicate where he thought he saw bands. Other experts, however, were skeptical about whether the presence of bands could be reliably determined. One expert thought the upper-most "band" in the male vaginal lane, if present, did not align closely enough with the upper band of suspect 1 to be called a match. So the results shown on this first autorad, although incriminating for suspect 2, are equivocal for suspect 1.

Figure 2³⁴ shows a second autorad produced in this case. The layout of the lanes is the same as the first autorad, although it shows the genotypes of the samples at a different locus, called D10S28. Again the victim and two suspects have distinct banding patterns. Notice, however, that suspect 1 has only one band and that this band is in the same position as the lower band of suspect 2. The male vaginal extract lane again contains two bands corresponding to those of suspect 2, which provides additional evidence against him. However, it is impossible to tell from this autorad whether a pattern corresponding to

³³ Genotype: the assortment of genes that compose the genetic make-up of an individual. DORLAND'S ILLUSTRATED MEDICAL DICTIONARY 607 (24th ed. 1965).

³⁴ See *infra* p. 1133.

that of suspect 1 appears in the vaginal extract because the only band matching his could be accounted for by the DNA of suspect 2. Additionally, the upper portion of the male vaginal extract lane contains dark blotches, caused by technical problems in the assay, that may obscure bands of the second rapist. So the evidence remains equivocal as to suspect 1.

Figures 3, 4, and 5³⁵ show the remaining autorads produced in this case. Each autorad shows the DNA banding pattern of the samples for a specific genetic locus. For each locus, the victim and two suspects have distinct banding patterns. For each locus the male vaginal extract lane contains a clear pattern corresponding to that of suspect 2, providing very strong evidence against him. Whether the male vaginal extract lane also contains a banding pattern matching suspect 1, however, is ambiguous at each locus. In the autorad shown in Figure 3,³⁶ for example, several experts saw no band corresponding to the lower band of suspect 1. Even the forensic analyst, the only person who claimed to see a band there, admitted uncertainty about it. The forensic analyst had more confidence in the presence of a band corresponding to the upper band of suspect 1. However, other experts dismissed this putative band as meaningless "schmutz" (a smudge) on the autorad, saying it lacked the morphology of a true band. The situation was the same for the autorad shown in Figure 4. For the autorad shown in Figure 5,³⁷ no one, not even the hyper-vigilant analyst, detected a band corresponding to the lower band of suspect 1 in the male vaginal extract lane. The forensic analyst thought he saw a band corresponding to the upper band of suspect 1, but other experts were adamant that no band was there.

To summarize, the laboratory report indicated that the DNA test had produced powerful evidence against *both* suspects — a five locus match between each suspect and the DNA found in semen extracted from the victim. The report gave no indication that the evidence against suspect 1 was weaker than that against suspect 2. Indeed, because the DNA profile of suspect 1 was slightly rarer than that of suspect 2, one might infer that the

³⁵ See *infra* pp. 1134-36.

³⁶ See *infra* p. 1134.

³⁷ See *infra* p. 1136.

DNA evidence against him is slightly stronger. Examination of the underlying autorads confirmed a clear, unambiguous match with suspect 2, but indicated the evidence against suspect 1 was ambiguous and equivocal. Suspect 1 happened to be my client.

My initial suspicion was that the forensic analyst had fallen victim to examiner bias (i.e., a tendency to see what one expects). The analyst knew my client was a suspect and could see my client's DNA pattern while making the judgment of whether bands corresponding to his were present in the vaginal extract. I feared the analyst may intentionally or unintentionally have conformed his judgment to the police theory of the case, which held that my client was one of the rapists.

When I raised concerns about examiner bias during the pre-trial phase of the case, however, the prosecution took the position that the autorads had been scored objectively by a computer-assisted imaging device. The prosecution claimed a scanner was used to create a digital image of each autorad. These images, according to the prosecutor, were scored by a computer program that detects the presence of bands in each lane according to their optical density, making the process entirely objective. Because I was skeptical of this claim, I obtained a court order that required the forensic laboratory to re-score the autorads with the computer imaging device while an independent expert and I watched.

During this re-scoring, the claim that the process was objective evaporated. In order to detect bands in the male vaginal extract lane that corresponded to those of suspect 1, the analyst had to increase the sensitivity of the computer to the point that it detected many additional "bands" that matched neither suspect. The analyst then performed a "manual override" of the computer's scoring, instructing the computer to "delete" (i.e., ignore) all of the bands that matched neither suspect.³⁸ When asked to state the basis for "deleting" some bands while leaving others, the analyst responded that he could tell by looking that

³⁸ An image of the autorad appeared on a computer screen, with green lines indicating places where the computer had detected a "band." The analyst was able to delete any bands he did not judge to be "true" bands through a simple point-and-click operation with the computer mouse. The software program also allows an analyst to re-position the "bands" using the mouse.

the undeleted bands, which happened to match my client, were "true" bands while the others were not. A number of the bands he deleted had higher optical densities than the bands scored as matching my client. So much for objectivity.

The re-scoring also resolved another issue. As the computer scored the bands, it compared their position to that of the size markers in order to estimate the size of the underlying DNA fragments designated by each band. Forensic laboratories use these sizings to determine whether bands of different samples align closely enough to be called a match. The laboratory had previously reported that the upper band of suspect 1 for locus D4S139 (see Figure 1)³⁹ was a perfect match with the highest band in the male vaginal extract lane. However, the re-scoring showed that the two bands differ in size by over nine percent. Because the policy of the laboratory called for a match to be declared only if the sizes of bands differ by less than four percent, this new scoring arguably excluded suspect 1 as a potential contributor of the DNA that gave rise to the upper band in the male vaginal extract lane. The re-scoring thereby confirmed the suspicion of the independent expert who, based on visual examination, doubted that there was a match.⁴⁰

This case shows that DNA test results are not always clear cut. More importantly, it illustrates how an analyst may draw damningly incriminating conclusions from data that are ambiguous or even exculpatory. In my view, innocent people are far more likely to be falsely incriminated through biased interpretation of ambiguous DNA test results than through coincidental matches with persons having the same profile. Consequently, I

³⁹ *Infra* p. 1132.

⁴⁰ Additionally, independent scorings of both the digital image of the autorad and a photographic copy of the autorad confirmed that the initial scoring (showing a perfect match) was wrong and the re-scoring (showing an exclusion) was correct. How was it, then, that the laboratory had initially scored these non-matching bands as a perfect match? My theory is that, during the initial scoring, the analyst performed a manual override of the computer to re-position the bands and make them match. Perhaps the analyst saw enough similarity between the DNA pattern of suspect 1 and faint bands in the male vaginal lane to confirm his suspicion that suspect 1 was guilty, and then took steps to improve the quality of the match to help police make the case against him. In other words, the analyst took a step down the pathway trodden by Fred Zain.

believe that the issue of how frequently ambiguities arise in DNA tests, and how laboratories deal with them, warrants far more attention than it has received.⁴¹

Ambiguous DNA test results can arise in a number of ways. Faint results, such as those just discussed, are quite common, particularly in cases involving mixed DNA samples. Minor inconsistencies between DNA profiles are also common. Bands may be somewhat misaligned (as in the autorad shown in Figure 1), or the number of bands observed may differ. The analyst must decide whether the discrepancies reflect true genetic differences or are simply the result of variability in the assays. In PCR-based tests, where the results are sometimes shown in a pattern of dots on test strips, ambiguities are even more common. The analyst must decide whether to "score" faint dots. When there are discrepancies between the patterns of two samples, the analyst must decide whether to attribute them to true genetic differences between samples or to technical problems in the assays. Technical problems which may give rise to discrepancies include the failure to detect certain alleles due to degradation of the DNA and the appearance of spurious extra dots due to cross-hybridization or contamination. Whether or not experimental controls have failed (and thereby invalidated the test) is sometimes also an issue that turns on subjective judgment.

Because ambiguities in the test results are resolved based on subjective judgment, the analyst has license to invoke all manner of ad hoc, unverified scientific reasoning to reach whatever interpretation is preferred. Analysts sometimes dismiss inconsistencies between profiles or problems with the test results (such as failed controls) by invoking ad hoc explanations that they fail to test empirically. Moreover, their ad hoc explanations sometimes shift with changing circumstances, making them inconsistent from case to case. To make matters worse, analysts sometimes rely on other evidence in a case to resolve ambiguities in DNA test results. I heard one forensic analyst defend the scoring of an ambiguous band (a judgment that incriminated a defen-

⁴¹ That this issue has received almost no attention, while the controversy over population structure has received a great deal of attention, is yet more evidence that the research agenda of the field is designed to meet the requirements of appellate courts rather than to seriously address potential problems.

dant in a rape case) by saying "I must be right, they found the victim's purse in [the defendant's] apartment."

Inferential bootstrapping of this sort is inevitable when analysts fail to use blind or objective scoring procedures. It can be terribly prejudicial to the defendant because it allows the analyst, by relying on other evidence in the case, to convert otherwise equivocal DNA results into seemingly damning evidence. To the trier of fact, it appears that the DNA test results is an independent piece of evidence against the defendant. In fact, the power of the DNA evidence may be derived in part from other evidence that the jury may already have considered. Consequently, the jury may double-count evidence against the defendant.

The danger of examiner bias has led scientists in many fields to insist that procedures for interpretation of potentially ambiguous data be performed either blindly or objectively.⁴² Both reports of the National Research Council call for the use of blind or objective "scoring" procedures by forensic DNA laboratories.⁴³ However, forensic laboratories have not followed the NRC's recommendations in this area. Their failure to do so cannot be explained on scientific grounds. If there is a scientific justification for the continued use of subjective interpretive procedures in forensic DNA testing in the face of contrary recommendations from the broader scientific community, it has yet to be articulated in the forensic science literature.

To understand the persistence of poor interpretive practices in forensic science, we must look beyond science to the sociology of the field. In my view, forensic scientists persist in relying

⁴² See, e.g., T.G. Tape & R.J. Panzer, *Echocardiography, Endocarditis, and Clinical Information Bias*, 1 J. GEN. INT'L MED. 300, 300 (1986) (finding that, in medical diagnostic setting, clinical information provided to interpreter of imaging tests can "bias the interpreter towards certain diagnoses, increasing the chances of false positives"). One of my faculty colleagues at UCI who uses RFLP analysis to study the mating habits of finches (by comparing the profiles of parents and offspring) is adamant that her assistants be "blind" when scoring autorads. She expressed concern that her staff might unintentionally be biased in their interpretations by knowing the hypotheses of the study, saying "there is too much at stake here not to use rigorous procedures — this research will affect our understanding of the entire evolutionary history of the finch." That blind procedures are considered essential for studying the mating habits of birds but not for determining the guilt and innocence of criminal defendants is another measure of the difference between forensic and academic science.

⁴³ See NRC1, *supra* note 12, at 72; NRC2, *supra* note 12, at 143.

on subjective judgment because this is the way they want to work. Their highest imperative is not to maintain scientific rigor, but to help their primary clients do justice. Faced with a choice between interpretive procedures that are scientifically rigorous and procedures that maximize the analyst's discretion to control the outcome of ambiguous cases, forensic scientists will opt for discretion over rigor whenever they can get away with it.

CONCLUSIONS

The sociological perspective I have offered here has implications for legal policy. If I am right, then it is a mistake for courts to view forensic science as a science. Lawyers and judges should not expect forensic scientists, left to their own devices, to follow the most rigorous scientific practices for validating test procedures and interpreting test results. General acceptance of a technique or procedure within the field of forensic science may not guarantee its reliability. Courts, therefore, should look beyond forensic science to the broader scientific community when seeking to determine whether a new method is sufficiently well accepted to be considered reliable. Furthermore, lawyers and judges should recognize that courts play the primary role in setting the research agenda for forensic science as a field and in raising the level of scientific practices. The reason we have seen a great deal of good research on population structure over the last five years, for example, is that some courts insisted that concerns about population structure be settled before DNA evidence was admitted. The reason we have seen little good research on PCR contamination or interpretation of ambiguous tests is that courts have not insisted that these issues be resolved as a condition for the admissibility of the tests. Lawyers and judges should recognize their profound influence over the character of forensic science, and exercise this power wisely. By setting high standards for the admissibility of forensic science evidence, courts provide a necessary incentive for better scientific work.

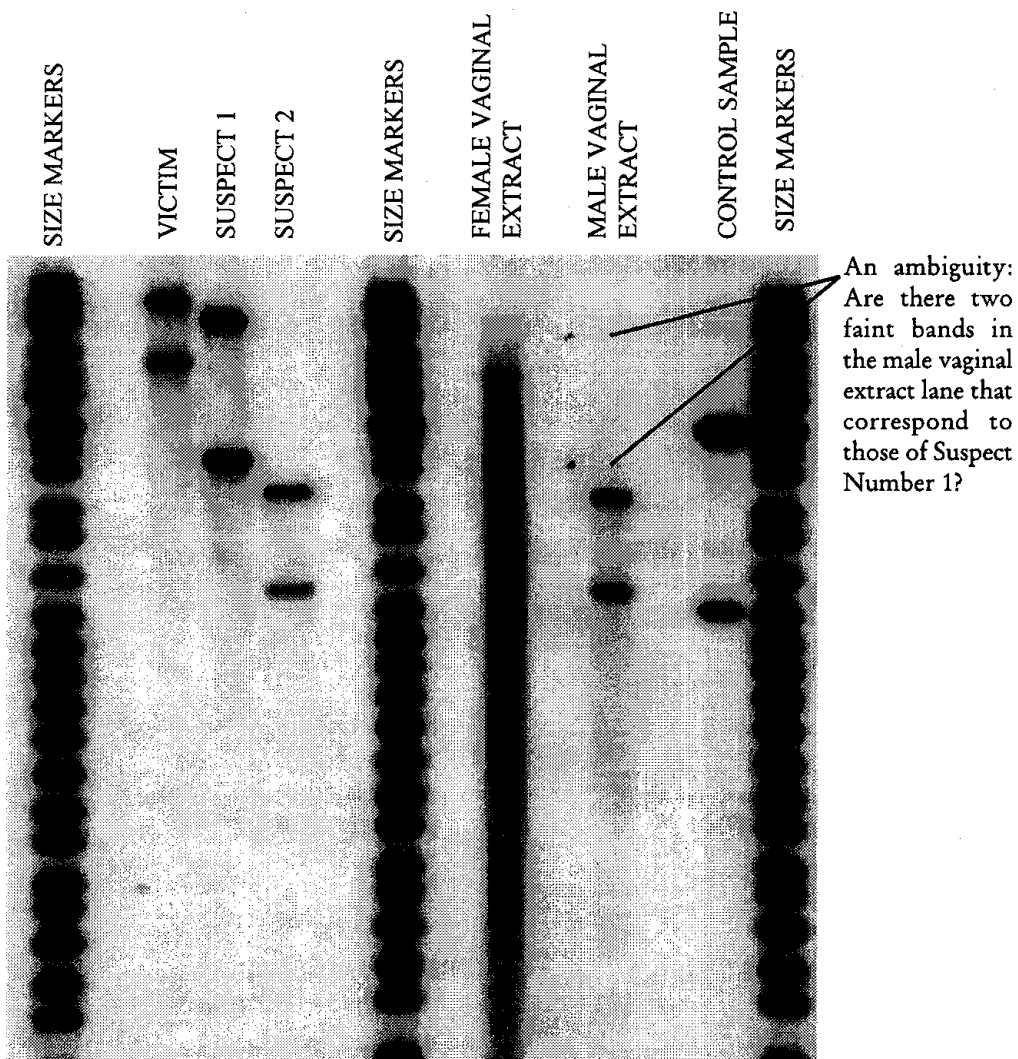


FIGURE 1

RFLP Autorad, Locus D4S139. In this case the victim was raped by two men. Banding patterns of the victim and two suspects appear on left side of autorad. Banding patterns of female and male extract from a vaginal sample appear on right side, along with a sample from a known individual run as a control. The key comparison is between the suspects' patterns and the pattern in the male vaginal extract. Two bands corresponding to those of suspect 2 are clearly visible. Whether bands corresponding to suspect 1 are also visible is a judgment call on which experts may differ. Dots to the left of the lane are felt-tip pen marks placed by a forensic analyst to indicate where he thought he saw bands.

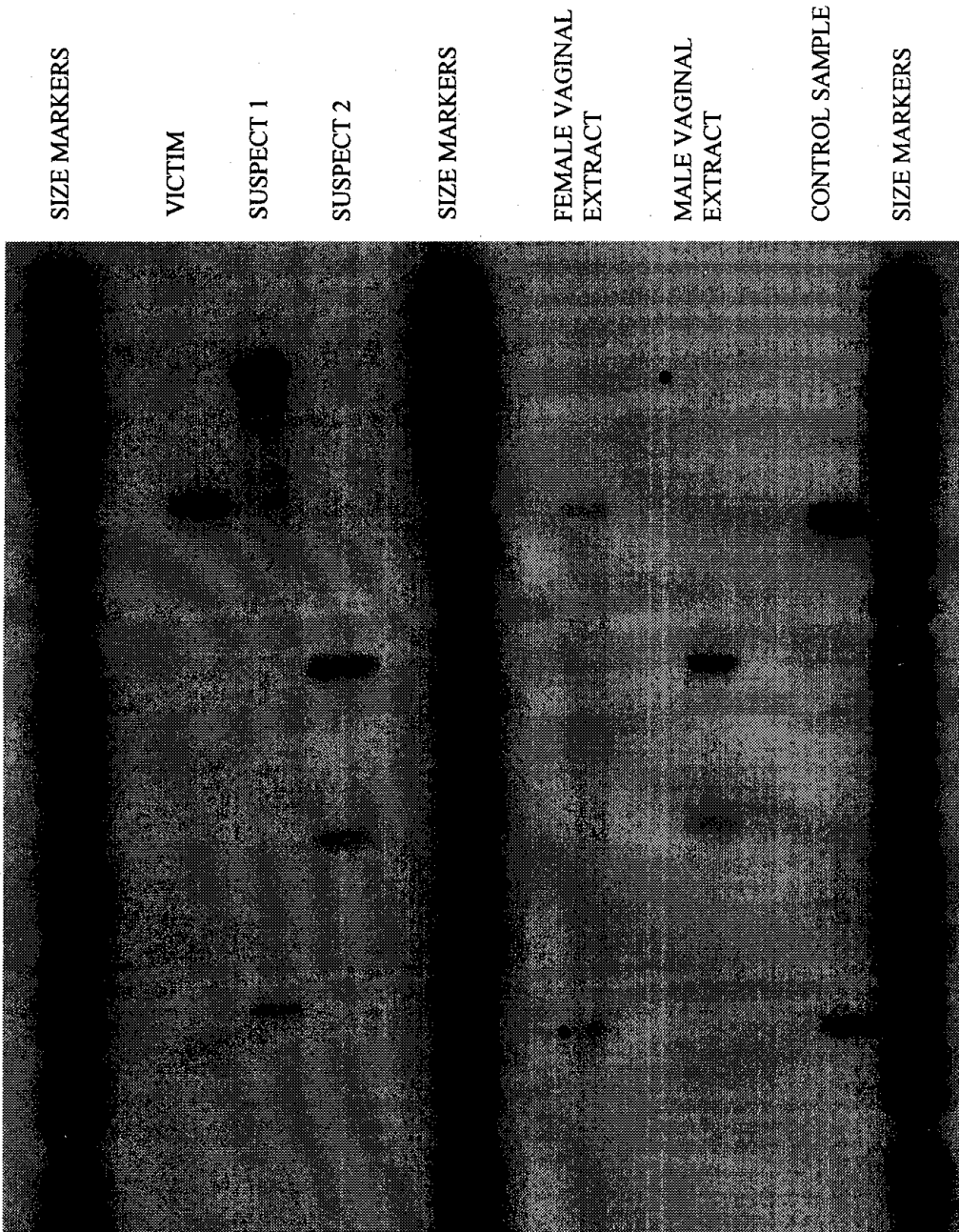


FIGURE 5

RFLP Autorad, Locus D17S26. A banding pattern corresponding to that of suspect 2 again appears in the male vaginal extract. The forensic analyst thought he also saw a band corresponding to the upper band of suspect 1 (and indicated the position of this putative band with a felt-tip pen mark). No band corresponding to the lower band of suspect 1 was detected in the male vaginal extract lane.