Organization economics explains many forensic science errors

Roger Koppl

Fairleigh Dickinson University, Madison, NJ, United States

koppl@fdu.edu

I thank Maria Minniti and David Croson for comments and William C. Thompson for providing the images used in Figures 2 & 3.

Abstract:

Judge Posner has identified an important lacuna in law and economics, namely, a tendency to ignore organization theory. I will apply the tools of organization theory to an area almost completely neglected in law and economics, forensic science. Posner points us to tools we should make use of; I am pointing to an application we have neglected. Forensic science today is characterized by a twofold monopoly. First, evidence is typically examined by one crime lab only. Second, that same lab will normally be the only one to offer an interpretation of the results of the examination it performs. Crime labs today are typically organized under law enforcement agencies, which may create conscious and unconscious biases in favor of police and prosecution. These organizational features of forensic science today encourage errors and wrongful convictions.

1. Introduction

Judge Posner has identified an important lacuna in law and economics, namely, a tendency to ignore organization theory. He provides a masterful summary of the literature and gives us several examples of how concepts such as agency costs, and information costs illuminate important issues in law and economics. "Many organizational features of the legal system," Posner concludes, "remain to be studied with the tools of organization economics" he surveys. I will apply those tools to an area almost completely neglected in law and economics, forensic science. Posner points us to tools we should make use of; I am pointing to an application we have neglected.

The application of organization economics to forensic science is timely in light of the recent release of an important study of forensic science by the National Academy of Sciences (NAS 2009). The study was commissioned and funded by Congress. The NAS report (as I will continue to call it) contains a sharp criticism of current practice in forensic science. "With the exception of nuclear DNA analysis," it says, "no forensic method has been rigorously shown to have the capacity to consistently, and with a high degree of certainty, demonstrate a connection between evidence and a specific individual or source" (p. S-5). Thus, the nature of forensic science has been widely misunderstood. It is far less "scientific" than commonly thought. In part, perhaps, this misapprehension reflects deficiencies in the popular model of "science." More importantly, however, it reflects the genuinely unscientific practices of many crime labs and the lack of scientific foundations for many forensic science disciplines.

According to the NAS report, the courts have not adequately disciplined bad practice within forensic science. "The bottom line is simple: In a number of forensic science disciplines, forensic science professionals have yet to establish either the validity of their approach or the accuracy of their conclusions, and the courts have been utterly ineffective in addressing this problem" (NAS 2009, p. 1-14). The NAS report contains 13 recommendations to Congress for corrective action (NAS 2009, pp. S-14 to S-24). Unfortunately, the perspective of organization economics did not inform these recommendations, which revolve around the idea of oversight.

My argument builds on three facts not all of which are generally appreciated. First, forensic science today uses more subjective judgment than generally understood. As we shall see, many forensic-science disciplines, including fingerprint identification, rely on the subjective judgments of examiners. Even DNA profiling often requires subjective judgment (Thompson 2009, Thompson & Cole 2007, Risinger et al. 2002). Second, "Most forensic science methods, programs, and evidence are within the regulatory province of state and local law enforcement entities or are covered by statutes and rules governing state judicial proceedings" (NAS 2009, p. S-9). Thus, performance measures for law enforcement, such as "number of arrests, prosecutions, [and] convictions" (Posner MS, p. 32) will influence the decisions of crime lab employees. As the NAS report correctly notes, "Forensic scientists who sit administratively in law enforcement agencies or prosecutors' offices, or who are hired by those units, are subject to a general risk of bias" (NAS 2009, p. 6-2). Third, forensic evidence is generally examined by one crime lab only, creating a kind of monopoly on the examination and interpretation of such evidence.

2. Subjective judgments in forensic science

Forensic science examinations use more subjective judgment than commonly recognized. A forensic scientist's decision to declare a "match" is a subjective judgment in many forensic science disciplines. (The NAS report discusses the variety of terms used to describe what is commonly thought of as "matching." See page S-15.) According to the NAS report, these disciplines include "impression evidence" such as shoeprints and tire tracks, toolmarks and firearms identification (the later commonly called "ballistics"), traditional hair microscopy, the handwriting comparisons of questioned document examiners, bloodstain pattern analysis, and fingerprint examinations (NAS 2009). Clearly, the last item is the most surprising, especially considering the existence of automated fingerprint identification systems such as those now commonly built into laptops. It is probably appropriate, therefore, to spend some time discussing the subjective nature of latent print identification.

Original research by Francis Galton, Henry Faulds, and others supports the hypothesis that friction ridge patterns on human hands and feet are unique and essentially unchanging over time (Beavan 2001). Apparently, the friction ridge patterns of an individual will change over time only to the extent that cuts, diseases, other accidents, or willful manipulations such as surgery or acid burnings induce a permanent change. Thus, both human examiners and automated systems can reliably decide whether two sufficiently clear and distinct images of friction ridge skin have a common source. The subjectivity of fingerprint analysis derives from the ambiguity of many "latent" fingerprints gathered at crime scenes. These latent prints may be smudged, distorted, covered with other latent prints, or otherwise unclear and indistinct.

Figure 1 illustrates. The clear and distinct image on the left is a rolled print taken from Brandon Mayfield by law enforcement officers. The unclear and indistinct image on the right is latent print taken from the scene of the 2004 Madrid train bombing. In 2004, the FBI declared a "100 percent match" of Mayfield to the latent lifted from the Madrid crime scene. The Spanish authorities objected this identification. They seem to have been correct as the FBI later withdrew its identification and released Mayfield (Office of the Inspector General 2006). The Mayfield error is one of a growing list of known false positive fingerprint errors (Cole 2005, NAS 2009, Williams 2007, Rubin & Winton 2008). On October 19th, 2007, Maryland judge Susan Souder cited the Mayfield case and other evidence in her decision to exclude incriminating fingerprint evidence from a capital case, *State of Maryland v. Bryan Rose* (In the Circuit Court for Baltimore County, Case No. K06-545). She described the standard technique as "a subjective, untested, unverifiable identification procedure that purports to be infallible" (p. 31).

Place Figure 1 about here.

The NAS report contains little to suggest that "nuclear DNA analysis" may often involve substantial use of subjective judgment. But, "DNA tests sometimes produce ambiguous results that are subject to multiple interpretations" and "[w]hen interpreting ambiguous results . . . human analysts rely heavily on subjective judgments to distinguish signal from noise, explain anomalies, and account for discrepancies" (Thompson and Cole 2007, 34). Subject judgment is likely to enter when more than one person has contributed to the DNA sample, or the sample is contaminated, degraded, or very small.

Figure 2 illustrates unambiguous DNA evidence. The biological sample is prepared and run through a genetic analyzer, which produces data that is represented as an electropherogram, which is then interpreted by a forensic scientist. As Figure 2 reveals, an electropherogram is a squiggly line. The figure shows only three loci, whereas most standard tests in the US examine 13 loci. Thompson and Cole (2007), from whom Figures 2 & 3 are lifted, explain. "As can be seen, the profile of Suspect 3 corresponds completely to that of the crime scene sample, hence it is a match that indicates Suspect 3 is a possible source of the blood at the crime scene. Suspects 1, 2, and 4 are eliminated as possible sources because one or more of their alleles differs from the crime sample" (p. 34).

Figure 2 about here

Figure 3 illustrates some of the ambiguities that can enter DNA profiling. When DNA is degraded or the original sample contains a small quantity of DNA, it can be difficult to distinguish between signal and noise. The figure, Thompson and Cole explain, "shows a comparison between the DNA profile of a saliva sample from the skin of a sexual assault victim and the profile of a suspect. Experts differed over whether these two profiles match. For example, some experts thought the peak labeled '12' at locus 'D3S1358' was a true allele, others thought it was merely noise in the system. The experts," they continue, "also differed over whether the peak labeled 'OL allele' at locus 'FGA' was a spurious anomaly that could be safely ignored, or whether it might be hiding another allele. When interpreting ambiguous results like those shown in [Figure 3] human analysts rely heavily on subjective judgments to distinguish signal from noise, explain anomalies, and account for discrepancies" (p. 34).

Figure 2 about here

Figures 2 & 3 illustrate a common practice in American crime labs. The forensic scientist typically examines the genetic profile of the police suspect(s) before characterizing the unknown sample. They "call the alleles" only after they have seen the answer key. As Figure 2 illustrates, the DNA examination is sometimes a multiple-choice test. Krane et al. (2008) object to observing the genetic information of suspects and victims before interpreting the genetic information of the unknown sample; they say it may induce "observer effects" that bias the outcome of the examination.

3. Dependence on law enforcement

The NAS report says, "The majority of forensic science laboratories are administered by law enforcement agencies, such as police departments, where the laboratory administrator reports to the head of the agency" (p. 6-1). We do not have good quantitative evidence on how much work is done in police labs rather than independent labs. (In some counties in California, the crime lab is organized directly under the prosecutor's office.) It is clear, nevertheless, that the statement in the NAS report is right.

Crime lab budgets and performance reviews often come from the police, the FBI, or the sheriff's office. Recent literature (Baicker and Jacobson 2007; Benson & Rasmussen 1995; Benson, Rasmussen, & Sollars 1995; Makowsky and Stratmann 2007) recognizes that law enforcement agencies have an incentive to expand their budgets in more or less the way described the public choice analysis of Niskanen (1971), notwithstanding the presence of competing motives. As I have noted elsewhere "If a crime lab's budget and performance reviews come from a budget maximizing law enforcement agency, personnel in the lab have an incentive to serve the interests of that agency, rather than abstract and pristine truth" (Koppl 2010).

4. Monopoly

Forensic science today is characterized by a twofold monopoly. First, evidence typically examined by one crime lab only (Koppl 2005a, 2005b). In this sense the crime lab receiving a bit of evidence has a monopoly on examination of that evidence. Second, that same lab will normally be the only one to offer an interpretation of the results of the examination it performs. No other experts in forensic science will be asked to judge what the evidence means. When expert witnesses are called in civil cases, it is common that each side will have its own expert or experts. The same is not true of criminal trials. Typically, only the prosecution will have expert witnesses testifying on forensic evidence (NAS 2009 S-8, Giannelli 2004). Monopoly in examinations may allow errors and even fraud to go undetected. Monopoly in interpretation may allow false interpretations to go unchallenged and alternative hypotheses to go unexamined.

The twofold epistemic monopoly in forensic science, the biases sometimes induced by dependence on law enforcement, and the ambiguity of much forensic evidence seem to have produced a needlessly high rate of "intentional and unintentional" (NAS 2009, p. 1-8) errors of test and interpretation in forensic science (Koppl 2005a; Saks and Koehler 2005; Garrett & Neufeld 2009). The organization of forensic science as a part of law

enforcement encourages interpretations favorable to the police theory or the prosecutor's case (NAS 2009, p. 6-2, Giannelli 1997, Risinger et al. 2002). Whitman and Koppl (2010) develop the logic of such institutional bias with a Bayesian model of rational choice by forensic scientists. The biases, or "observer effects," induced in this way need not be conscious (Dror & Charlton 2006, Dror Charlton & Péron 2006, Risinger et al. 2002, Krane et al. 2008).

5. Objective performance criteria for law enforcement.

Judge Posner notes the use of objective performance criteria to evaluate police agencies. He notes an agency problem with policing, saying, "when agents can appeal to local knowledge as the basis of their decisions, the center, which lacks that knowledge, cannot monitor their decisions effectively" (p. 32). He imagines the problem is mostly solved by objective performance measures. "When the task is criminal investigation, this problem is not acute, because the outputs of FBI agents engaged in criminal investigation – number of arrests, prosecutions, convictions, length of sentences, and amount of property recovered – are quantitative, relatively hard to manipulate (at least legally), and therefore feasibly measured." He imagines these objectives measure permit the successful use of high-powered incentives.

The criteria Posner lists, with the partial exception of "amount of property recovered," make suspects fungible. Law enforcement agencies should attempt to discriminate between the guilty and the innocent. Performance criteria such as number of convictions, however, place no weight on correctly discriminating between the innocent and the guilty. The police can produce arrests, trials, and convictions of innocent persons by arresting persons who are poor, uneducated, or mentally weak or by inducing false confessions. As Holmstrom and Milgrom (1991) point out in a classic contribution to organization economics, high-powered incentives may be inappropriate when relevant aspects of the agent's performance are unobservable, hard to measure, or subject to significant measurement error. (Dewatripont, Jewitt, andTirole 2000 provide a helpful survey.) If we may consider false convictions to be low quality output, we would seem to have an example of their "result that it may be optimal to provide no quantity incentives when quality is poorly measured" (Holmstrom & Milgrom 1991, p. 28).

It is difficult to gather data that would reveal how many poor, uneducated, and mentally weak persons are wrongfully convicted. The literature on false confessions, however, supports the notion that some police investigators in at least some cases may unjustly target such persons. Drizin & Leo (2004) assembled the largest set of "proven" false confessions, consisting of 125 cases. They identify four groups who, compared to others, "are more vulnerable to the pressures of interrogation and therefore less likely to possess or be able to muster the psychological resources or perspective necessary to withstand accusatorial police questioning" (p. 919). They are children under the age of 14 (pp. 963-968), Juveniles under the age of 18 (pp. 968-970), the "mentally retarded" (pp. 970-973), and the mentally ill (pp. 973-974). (The issue at hand is not related to the debate, reviewed by Novak and Engel 2005, on "criminalizing mental illness.")

Techniques readily available to the police allow them to do easily what Posner says they can do only with difficulty, namely, manipulate the performance criteria. In some cases,

such manipulation is illegal. Long standing "allegations . . . that Chicago police detectives routinely tortured murder suspects," for example, may be legitimate given the perjury and obstruction of justice charges filed against a retired police commander in October 2008 and the appointment of a special prosecutor in April 2009 for "the cases of five murder suspects seeking new trials" in connection with the commander's alleged abuses (Coen & Mills 2008, Wallberg 2009). The techniques may also be applied legally. Police brutality is not required to extract a false confession from many vulnerable persons. The poor are less able to marshal resources to defend themselves even when innocent. And so on. It seems likely that unconscious biases could cause police to concentrate attention on persons who are poor, uneducated, or mentally weak even when the police have no conscious desire to manipulate the system.

Monopoly forensic science is an important tool for police investigators and prosecutors seeking (consciously or unconsciously) to maximize convictions. As we have seen, the current organization of forensic science encourages crime labs to support the police and prosecution. When forensic evidence is ambiguous this disposition may lead to wrongful arrest and conviction.

Public records reveal at least some use of several inappropriate practices that support the epistemic monopoly of crime labs. These include 1) evidence destruction (Greene & Moffiet 2007); 2) resistance to case review when errors are revealed or suspected (Thompson & Dioso-Villa 2008, pp. 283-289); 3) failure to comply with Coverdell oversight requirements (Office of the Inspector General 2008); 4) over-claiming in forensic science testimony (Cole 2007, especially pp. 821-825; Garrett & Neufeld 2009, especially pp. 55-60); 5) resistance to redundancy and oversight (Gestring 2009); strong demarcations between outsiders and insiders (Gestring 2009); 7) asymmetric models of defense and prosecution experts (CLPEX 2007); 8) resistance to blind testing (Wells 2009), 9) reliance on subjective techniques (NAS 2009), and 10) efforts to exclude defense experts (Kreeger 2002).

6. Conclusion

Objective police performance measures such as number of arrests, prosecutions, and convictions are of limited utility in overcoming agency problems when the agent enjoys an epistemic monopoly. Such performance criteria create an incentive to treat all persons as fungible inputs to a production process aimed at maximizing the number of case resolutions, rather than discriminating between persons who have and have not committed crimes. The epistemic monopoly in forensic science hampers error prevention, detection, and correction, thus allowing errors to go unnoticed. When making a monopolistic subjective judgment about ambiguous evidence, forensic scientists in the employ of law enforcement agencies have an incentive to interpret evidence in ways that support the interest of law enforcement. Law enforcement agencies are judged largely by the number of arrests, prosecutions, and convictions they generate. Thus, crime labs have an incentive to interpret evidence to maximize case closures rather than to discriminate between the guilty and the innocent or report accurately on the evidence.

More work applying the insights of organization economics to forensic science would help to fill an unfortunate gap in the scholarly literature in law and economics. More importantly, such work might help to prevent wrongful convictions based on forensic-science testing errors and false or misleading forensic-science testimony.

References

Baicker, K. and M. Jacobson. 2007. "Finders keepers: Forfeiture laws, policing incentives, and local budgets," 91(11) *Journal of Public Economics*, 2113-2136.

Beavan, Colin. 2001. *Fingerprints: The Origins of Crime Detection and the Murder Case that Launched Forensic Science*, New York: Hyperion.

Benson, B. and D. Rasmussen. 1995. "Predatory Public Finance and the Origins of the War on Drugs," 1 *The Independent Review,* 163–195.

Benson, B., D. Rasmussen, and D. Sollars. 1995. "Police Bureaucracies, their Incentives, and the War on Drugs," 83 *Public Choice*, 21–45.

Coen, Jeff & Steve Mills. 2008. "Feds catch up with Burge Notorious ex-Chicago commander charged with lying about torture, *Chicago Tribune*, 22 October 2008. Downloaded 16 April 2009 from http://archives.chicagotribune.com/2008/oct/22/local/chi-burge-22-oct22.

Cole, Simon. 2005. "More Than Zero: Accounting for Error in Latent Fingerprint Identification," 95(3) *Journal of Criminal Law & Criminology*, 985-1078.

Cole, Simon. 2007. "Where the Rubber Meets the Road: Thinking About Expert Evidence as Expert Testimony," 52 *Villanova Law Review*, 803-842.

CPLEX. 2007. Complete Latent Print Examination Message Board http://clpex.com/phpBB/viewtopic.php?t=616, downloaded 13 November 2007.

Dewatripont, Mathias, Ian Jewitt, Jean Tirole. 2000. "Multitask agency problems: Focus and task clustering," 44 *European Economic Review*, 869–877.

Drizin, Steen A. & Richard A. Leo. 2004. "The Problem of False Confessions in the Post-DNA World," *North Carolina Law Review*, 82: 891-1004.

Dror, Itiel E., and David Charlton. 2006. "Why Experts Make Errors." 56(4) *Journal of Forensic Identification* 600-616.

Dror, Itiel, David Charlton, and Ailso E. Péron. 2006. "Contextual information renders experts vulnerable to making erroneous identifications." 156 *Forensic Science International* 74-78.

Garrett, Brandon & Peter J. Neufeld. 2009. "Invalid forensic science testimony and wrongful convictions," *Virginia Law Review*, 95(1): 1- 97.

Gestring, Brian. 2009. "The Dawn of the 'Forensic Science Provocateur,'" *CAC News*, 1st quarter 2009: 25-28.

Giannelli, P. C. 1997. "The Abuse of Evidence in Criminal Cases: The Need for Independent Crime Laboratories," *Virginia Journal of Social Policy & the Law*, 4: 439-478.

Giannelli, Paul C. 2004. "*Ake v. Oklahoma*: The right to expert assistance in a post-*Daubert*, post-DNA world," *Cornell Law Review*, 89: 1305-1419.

Greene, Susan & Miles Moffiet. 2007. "Bad faith difficult to prove," *The Denver Post*, 22 July 2007. Downloaded 28 January 2009 from http://www.denverpost.com/evidence/ci_6429277.

Holmstrom, Bengt and Paul Milgrom. 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design," 7(Sp) *Journal of Law, Economics, & Organization*, 24-52.

Koppl, Roger. 2005a. "How to Improve Forensic Science." 20(3) *European Journal of Law and Economics*, 255-286.

Koppl, Roger. 2005b. "Epistemic Systems." 2(2) *Episteme: Journal of Social Epistemology*, 91-106.

Koppl, Roger. 2010 (forthcoming). "Romancing Forensics: Legal Failure in Forensic Science Administration," in: Edward Lopez (Ed) *The Pursuit of Justice: Law and Economics of Legal System*, New York: Macmillan.

Krane, Dan and ten others. 2008. "Sequential Unmasking: A Means of Minimizing Observer Effects in Forensic DNA Interpretation," *Journal of Forensic Sciences* 53(4): 1006-1007.

Kreeger, L. 2002. "Preparing for defense experts," *Silent Witness* (newsletter of the National District Attorneys Association), 7(2). Downloaded 16 April from http://www.ndaa.org/publications/newsletters/silent_witness_volume_7_number_2_2002.html.

Niskanen, W. 1971. *Bureaucracy and Representative Government,* Chicago: Aldine-Atherton.

Novak, Kenneth & Robin S. Engel. 2005. "Disentangling the influence of suspects' demeanor and mental disorder on arrest," *Policing: An International Journal of Police Strategy & Management*, 28(5): 493-512.

Office of the Inspector General (OIG), Oversight and Review Division. 2006. *A Review of the FBI's Handling of the Brandon Mayfield Case: Unclassified and*

Redacted, U. S. Department of Justice. Available at http://www.usdoj.gov/oig/special/s0601/final.pdf.

Office of the Inspector General (2008) *Review of the Office of Justice Programs' Paul Coverdell Forensic Science Improvement Grants Program*, U.S. Department of Justice.

Risinger, Michael, Michael J. Saks, William C. Thompson & Robert Rosenthal. 2002. "The Daubert/Kumho Implications of Observer Effects in Forensic Science: Hidden Problems of Expectation and Suggestion." 90(1) *California Law Review*, 1–56.

Rubin, Joel and Richard Winton (2008) LAPD finds faulty fingerprint work, *Los Angeles Times*, 17 October 2008.

Saks, Michael & Jonathon J. Koehler. 2005. "The Coming Paradigm Shift in Forensic Identification Science," *Science*, 309: pp. 892- 895.

Thompson, William C. 2009. "Painting the Target Around the Matching Profile: The Texas Sharpshooter Fallacy in Forensic DNA Interpretation, *Law, Probability and Risk*, forthcoming.

Thompson, William C. and Simon A. Cole. 2007. "Psychological Aspects of Forensic Identification Evidence," in Costanzo, Mark, Daniel Krauss, and Kathy Pezdek (eds.), *Expert Psychological Testimony for the Courts*, Mahwah, NJ: Erlbaum, pp. 31-68.

Thompson, William C. & Rachel Dioso-Villa. 2008. "Turning a blind eye to misleading scientific testimony: Failure of procedural safeguards in a capital case," *Albany Journal of Science and Technology*, 18: 151-304.

Wallberg, Matthew. 2009. "Former Cook County judge named special prosecutor in murder cases," Chicago Tribune, 8 April 2009. Downloaded 16 April 2009 from www.chicagotribune.com/news/local/chi-burge-ruling-08-apr08,0,374954.story.

Wells, Jeffrey D. 2009. "Commentary on: Krane DE, Ford S, Gilder JR, Inman K, Jamieson A, Koppl R, Kornfield IL, Risinger DM, Rudin N, Taylor MS, Thompson WC. Sequential unmasking: a means of minimizing observer effects in forensic DNA interpretation. J Forensic Sci 2008;53(4):1006–7, 54(2) *Journal of Forensic Sciences*, 500.

Whitman, Glen and Roger Koppl. 2010. "Rational Bias in Forensic Science," *Law, Probability, & Risk*, forthcoming.

Williams, Joy. 2007. *Administrative Review* PC-07-0018, Seminole County Sheriff's Office, 4 June 2007.

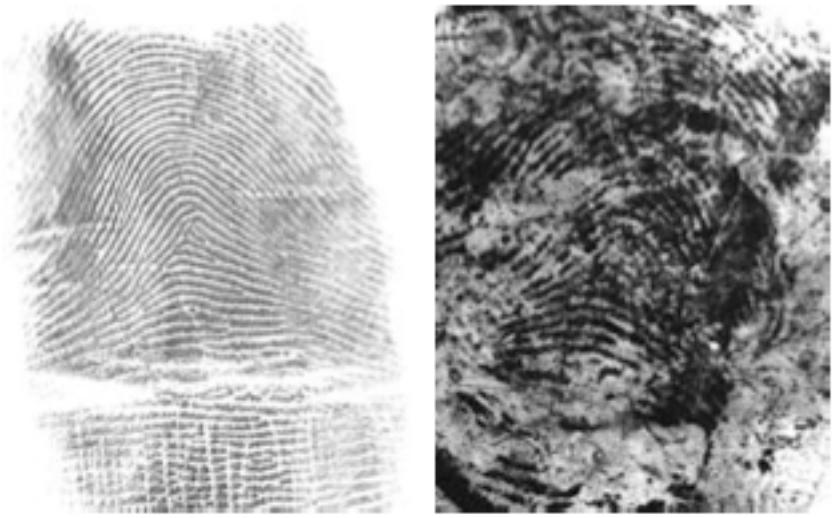

Figure 1: Ambiguous and Unambiguous Fingerprint Images

The clear and distinct image on the left is a rolled print taken from a suspect by law enforcement officers. The unclear and indistinct image on the right is latent print taken from a crime scene by law enforcement officers. FBI examiners declared a "100 percent match" between them.

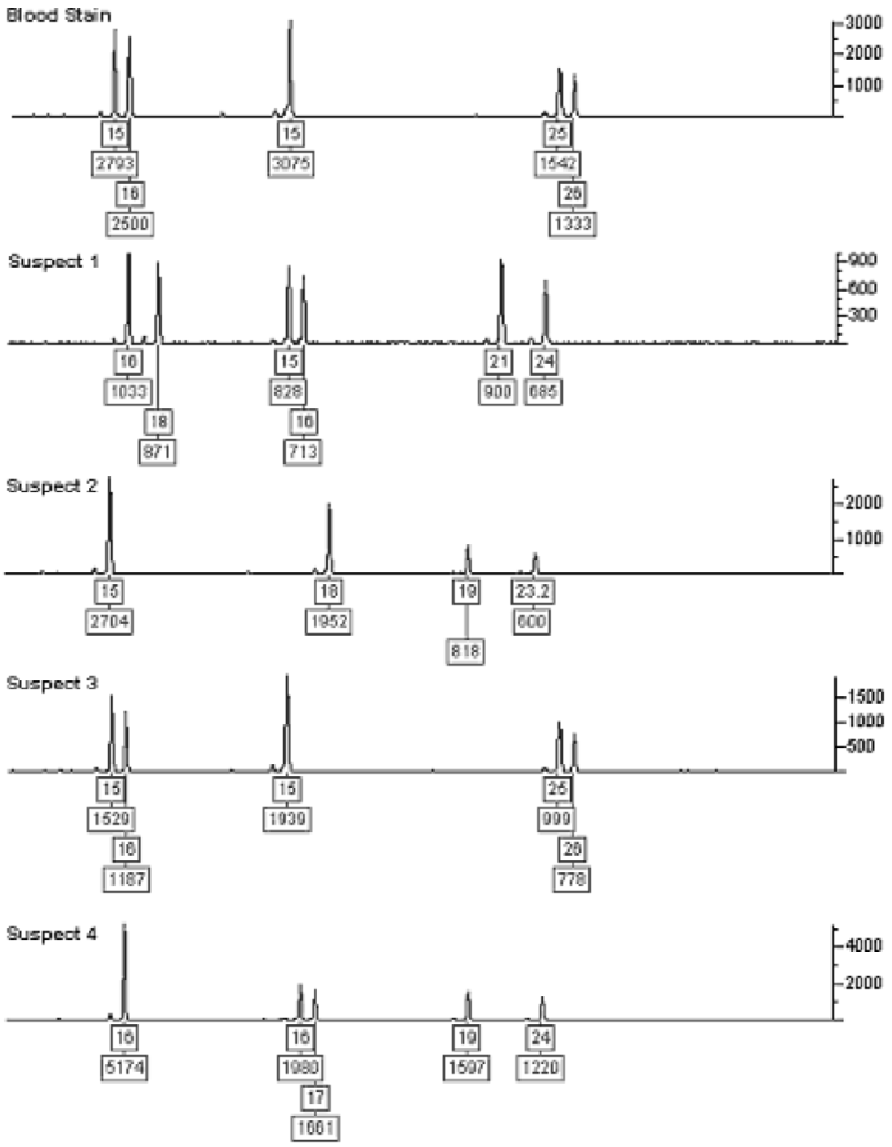

Figure 2: Unambiguous DNA Evidence

The electropherogram of the crime scene evidence is clear and distinct. It matches suspect 3, who is included as a possible source. Suspects 1, 2, & 4 are clearly excluded.

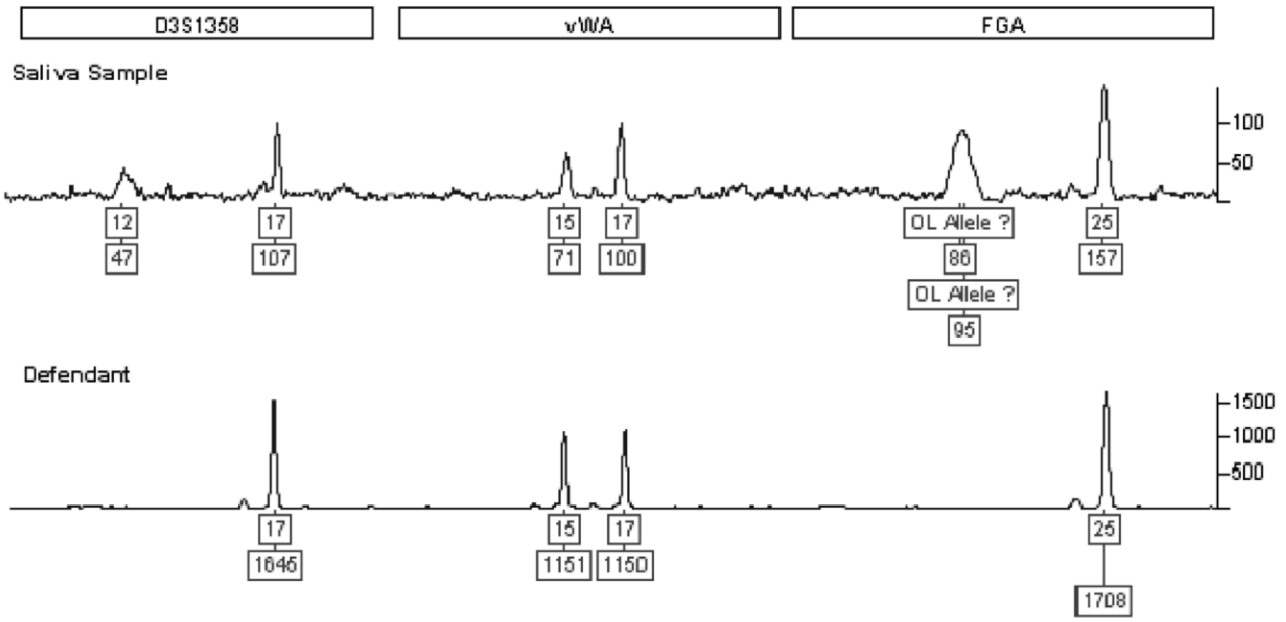

Figure 3: Ambiguous DNA Evidence

The electropherogram of the crime scene evidence is not clear and distinct. It is not clear whether the suspect should be included or excluded as a possible source.