

THREE CARD MONTE, MONTY HALL, *MODUS
OPERANDI* AND “OFFENDER PROFILING”: SOME
LESSONS OF MODERN COGNITIVE SCIENCE FOR
THE LAW OF EVIDENCE

D. Michael Risinger and Jeffrey L. Loop***

“I want the truth!”

“You can’t handle the truth!”¹

Nowhere has the thorny relationship between science and the law been more difficult than in regard to the law of proof. In the last century, the products of science have indeed been allowed to enter the courtroom as evidence in individual cases (along with much masquerading as the product of science). But even in this context, until recently the law has shown little inclination to come to grips with either the nature of the enterprise of modern science, or of its special epistemic claims. Perhaps even more important, however, is the general failure of the law to reflect virtually any of the insights of modern research on the characteristics of human perception, cognition, memory, inference or decision under uncertainty, either in the structure of the rules of evidence themselves, or the ways in which judges are trained or instructed to administer them. Those rules of evidence that functionally depend on such questions were derived by accretion from common sense notions over the course of three or four hundred years. The great syntheses of such notions into an integrated system is a product of the 19th century, which came to virtually complete fruition by 1904 with the publication of the first edition of Wigmore’s great treatise. In the last century, while there have been refinements which made the underlying common-sense notions somewhat more explicit, the general contours of the relevant evidence law have remained virtually unchanged, while the scientific exploration of these topics has exploded, yielding many well documented conclusions concerning a variety of human tendencies to

* Professor, Seton Hall University School of Law.

** Associate, Debevoise & Plimpton, New York. The opinions expressed herein do not necessarily reflect the views of Debevoise & Plimpton or any of its clients.

¹ A FEW GOOD MEN (Columbia Pictures & Castle Rock Entertainment 1992) (courtroom exchange between actors Tom Cruise and Jack Nicholson).

accurately and inaccurately handle information under various conditions.

The collision between the insights of experimental psychology and some of the important commonsense underpinnings of standard evidence doctrine began just as the systematic integration of standard doctrine was being completed. This is most dramatically shown in the rather vicious conflict in the first decade of the 20th century between Wigmore and the pioneer cognitive psychologist Hugo Münsterberg over the law's handling of eyewitness testimony.² Suffice it to say that while the last century has seen the accumulation of literally thousands of studies on the weaknesses of eyewitness testimony, the controversy concerning the proper response of the law of proof remains heated, and in most jurisdictions the law remains functionally unchanged.³ This is not to say that there have not been adjustments in legal doctrine as insights from science have become so well documented as to render them un-ignorable. However, the results of those adjustments have rarely altered outcomes. Take for example, the creation of the doctrine that eyewitness identifications which are the product of suggestion are excludable.⁴ Like an oyster dealing with an irritant by coating it with nacre, the law has recognized the inconvenient phenomenon of suggestion, and has covered it with a doctrine of surface luster which has changed the way the system operates almost not at all, since identifications are rarely found to be the product of suggestion except in the most extreme

² For a description of the collision, see WILLIAM TWINING, *THEORIES OF EVIDENCE: BENTHAM AND WIGMORE* 135-36 (1985). See also generally John H. Wigmore, *Professor Muensterberg and the Psychology of Testimony*, 3 ILL. L. REV. 399 (1909).

³ In 1926, Robert Maynard Hutchins, then a 27-year-old member of the Yale Law faculty, gave a speech at the annual meeting of the Association of American Law Schools calling for the overhaul of the rules of evidence in light of the findings of psychology. See HARRY S. ASHMORE, *UNSEASONABLE TRUTHS: THE LIFE OF ROBERT MAYNARD HUTCHINS* 46-47 (1989) (describing Hutchins' criticism of the prevailing practice). Wigmore responded in a critical letter to the Dean of the Yale Law School. This did not prevent Hutchins from co-authoring a series of articles on the subject with his colleague Donald Slesinger. See Robert M. Hutchins & Donald Slesinger, *Some Observations on the Law of Evidence—Consciousness of Guilt*, 77 U. PA. L. REV. 725 (1929); Robert M. Hutchins & Donald Slesinger, *Some Observations on the Law of Evidence—Family Relations*, 13 MINN. L. REV. 675 (1929); Robert M. Hutchins & Donald Slesinger, *Some Observations on the Law of Evidence—State of Mind in Issue*, 29 COLUM. L. REV. 147 (1929); Robert M. Hutchins & Donald Slesinger, *Some Observations on the Law of Evidence—State of Mind to Prove an Act*, 38 YALE L. J. 283 (1929); Robert M. Hutchins & Donald Slesinger, *Some Observations on the Law of Evidence—Spontaneous Exclamation*, 28 COLUM. L. REV. 432 (1928); Robert M. Hutchins & Donald Slesinger, *Some Observations on the Law of Evidence—Memory*, 41 HARV. L. REV. 860 (1928); Robert M. Hutchins & Donald Slesinger, *Some Observations on the Law of Evidence—The Competency of Witnesses*, 37 YALE L.J. 1017 (1928). These articles repay reading today, though they do at times seem quaintly poised between the excesses of Freud and the early Behaviorists. However, they were not much cited by courts (a Westlaw "allcases" search shows 40 citations total for all seven articles in over a span of 73 years, the plurality of which were string cites). In the end, Hutchins' call for reform accomplished little beyond drawing down the wrath of Wigmore.

⁴ See, e.g., *Manson v. Brathwaite* 432 U.S. 98 (1977); *Neil v. Biggers*, 409 U.S. 188 (1972); *Stovall v. Denno*, 388 U.S. 293 (1967).

cases.⁵

Perhaps we should not be too surprised at this state of affairs. The law is, all other things being equal, a profoundly conservative enterprise.⁶ Socialized in a long tradition of *stare decisis*, judges in general do not depart easily from the way things were done yesterday. In addition, the claimed ideology of the proof system, the standard “search for truth” model, accounts very imperfectly for the realities of the system as it actually operates. The actual operation of the system may be better seen as polyvalent, with rectitude of decision being but one of a number of constituent beasts with claims to be fed by the system, along with crime control, vengeance, protection of established wealth and power, and others.⁷ The compromises which are made to feed all these demanding beasts yield a system which is prone to anomic error which can be seen as at different levels of magnification as either random or biased: Random in that it is often unpredictable in either specific content or direction in the individual case, and biased in that certain trends can be derived from large samples of cases.⁸

The last quarter century has seen mounting evidence that humans manifest specific and predictable weaknesses in dealing with certain

⁵ Professors Saltzburg and Capra conclude that, despite legal doctrines that would suggest a different obligation, “many courts are not very careful in their handling of eyewitness evidence.” STEPHEN A. SALTZBURG & DANIEL J. CAPRA, *AMERICAN CRIMINAL PROCEDURE: CASES AND COMMENTARY* 771, 769 (6th ed. 2000). Additionally, they observe that “[t]here is certainly strong evidence that the *Manson* test, at least as applied by the courts, does little to deter the police from using suggestive identification procedures.” *Id.* One area in which psychology research may have had a larger practical effect is in the law’s dealings with child witnesses. See Gail S. Goodman & Jodi Quas, *Innovations for Child Witnesses, A National Survey*, 5 PSYCHOL. PUB. POL. & L. 255 (1999). Another may be line-up identification, following the recent promulgation of federal guidelines based on current research. See Technical Working Group for Eyewitness Evidence, United States Department of Justice: EYEWITNESS EVIDENCE: A GUIDE FOR LAW ENFORCEMENT (1999). Thus far, however, only a small number of states have adopted them, and only New Jersey has adopted them fully. See *Witnesses, Victims Get New Way to ID Suspects*, SUNDAY RECORD (Bergen County, N.J.), July 22, 2001, at A3.

⁶ “The [evidence] rulemaking process, like the judiciary itself, is inherently conservative We are resistant to change. The judiciary doesn’t jump in and like things changed” The Hon. Fern Smith, U.S.D.J., N.D. Cal, Former Chair of the Federal Advisory Committee on the Rules of Evidence and Current Director of the Federal Judicial Center, Address Before the Section on Evidence of the Association of American Law Schools on “The Politics of (Evidence) Rule-Making,” in New Orleans, La. (Jan. 4, 2002).

⁷ While our stated ideal is truth-finding and proof beyond a reasonable doubt, our operative delivery is often more easily squared with feeding the competing beasts. See D. Michael Risinger, *John Henry Wigmore, Johnny Lynn Old Chief, and “Legitimate Moral Force”: Keeping the Courtroom Safe for Heartstrings and Gore*, 49 HASTINGS L.J. 403, 436-38 (1998) [hereinafter Risinger, *Heartstrings and Gore*]. This insight is hardly new. See RICHARD MAXWELL BROWN, *STRAIN OF VIOLENCE: HISTORICAL STUDIES OF AMERICAN VIOLENCE AND VIGILANTISM* 146-47 (1975), quoting Herbert L. Packer, *Two Models of the Criminal Process*, 113 U. PA. L. REV. 1 (1964). Nevertheless, we must remain convinced that the law can be held to its highest ideals when sufficiently clear exposure of its shortcomings in their light makes the shortcomings impossible to ignore.

⁸ For evidence of one such bias in favor of the prosecution in criminal cases, see D. Michael Risinger, *Navigating Expert Reliability: Are Criminal Standards of Certainty Being Left on the Dock?*, 64 ALB. L. REV. 99 (2000) [hereinafter Risinger, *Navigating Expert Reliability*].

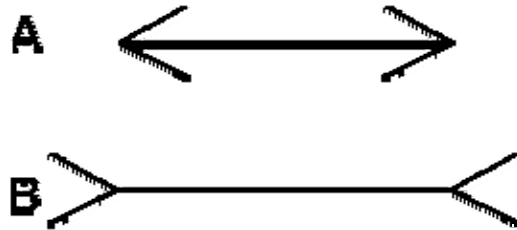
kinds of information under definable conditions, which weaknesses are reflected in traditional proof law imperfectly or not at all. These increasingly well-documented weaknesses have been called cognitive “tunnels.”⁹ They appear to be the result of hard-wired processing heuristics—that is to say, built-in recipes for dealing quickly with a general class of problems from input information.¹⁰ Clearly, such heuristics must give usefully right answers to most problems of the class, or else they would have been so counterproductive as to have been eliminated by natural selection. However, in order to be useful, such heuristics don’t have to be actually right even a high percentage of time—as long as the errors favor survival. Think of a bird which always flies up from the ground when a certain type of movement is perceived. While it wastes some energy, if there is no cognitive circuit which can accurately separate the movement of a real predator from that of a non-predator, survival favors the bird who flies up every time over the bird that does not, even though by not flying up the non-flying bird may be right ten or a hundred times to the flying bird’s once. It is the once that counts, determining survival for one and death for the other.

The bird flight example is easily understood, and generally taken to reflect some neurally mediated heuristic existing below what we would ordinarily call a cognitive level, closer to a “reflex.” Such “startle” reflexes may result in action prior to anything normally called “cognition” in humans. A step up the ladder toward cognition are “optical illu-

⁹ MASSIMO PIATTELLI-PALMARINI, INEVITABLE ILLUSIONS: HOW MISTAKES OF REASON RULE OUR MINDS 45-52 (1994). See Amos Tversky & Daniel Kahneman, *Judgment under uncertainty: Heuristics and biases*, in JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES (Daniel Kahneman et al. eds., 1982); Michael J. Saks & Robert F. Kidd, *Human Information Processing and Adjudication: Trial by Heuristics*, 15 LAW & SOC’Y REV. 123 (1980).

¹⁰ See PIATTELLI-PALMARINI, *supra* note 9, at 19. In a recent review of Paul Slovic’s *THE PERCEPTION OF RISK*, Cass Sunstein gives a useful definition of heuristic: “Heuristics are rules of thumb, substituting a simple question for a more difficult one.” Cass R. Sunstein, *The Laws of Fear*, 115 HARV. L. REV. 1119 n.24 (2002) (book review), citing Daniel Kahneman & Shane Frederick, *Representativeness Revisited: Attribute Substitution in Intuitive Judgment*, in HEURISTICS OF INTUITIVE JUDGMENT: EXTENSIONS AND APPLICATIONS (T. Gilovich et al. eds., forthcoming 2002) (manuscript at 2-3, on file with the Harvard Law School Library). It should be noted that there is an ongoing debate over the extent to which ordinary humans are subject to processing errors from “probability blindness” in circumstances of decision presented by everyday life in the modern world. Some, most notably Gerd Gigerenzer, assert that the poor performance of people in laboratory experiments are more an artifact of the artificiality of the way information is presented in the experiment than a function of inaccurate judgment in normal circumstances. See, e.g., Gerd Gigerenzer & Peter M. Todd, *Fast and Frugal Heuristics: The Adaptive Toolbox*, in SIMPLE HEURISTICS THAT MAKE US SMART (Gerd Gigerenzer et al. eds., 1999); Gerd Gigerenzer, *How to Make Cognitive Illusions Disappear: Beyond “Heuristics and Biases,”* in 2 EUR. REV. SOC. PSYCHOL. (Wolfgang Stroebe & Miles Hewstone eds., 1991). The debate appears to be a debate over whether our cognitive cup is half or more empty, or half or more full, since both sides concede that there are some problems we solve well, and some problems we deal with poorly. See GEORGE LAKOFF & MARK JOHNSON, *PHILOSOPHY IN THE FLESH: THE EMBODIED MIND AND ITS CHALLENGE TO WESTERN THOUGHT* 527-28 (1999). Whether the task of specific inference from propensity in the rather artificial setting of the courtroom is a task that humans do well is open to debate.

sions,” a phenomenon with which all are familiar. It would be easy to generate a list of well-known optical illusions which have been extensively catalogued and studied in the past century: stimuli which will cause the brain to perceive the existence of things which are not present (by reference to more objective standards of evaluation). A simple one will suffice, the two lines illusion:¹¹



The two central lines are the same length, but the bottom one appears much longer. This illusion belongs to a class which exhibits certain important characteristics, chiefly, that it persists even when the observer knows that it is an illusion, and it is virtually impossible to learn to accurately judge when it is present and when not without resort to more objective instruments, even when one knows of the existence of such a phenomenon. Nevertheless, when presented with a ruler, it is not hard for the observer to conclude that the two lines are the same length, and to accept the error of his own perception.

Optical illusions are generally called perceptual illusions, to separate them from heuristic errors that are dependant on reasoning, analysis or reflection. That is not to say that the source of the error is in the eye and not the brain, but merely to say that they manifest themselves at the same time as, or *as*, perception, unlike reflexes that can be pre-perceptual, and cognitive tunnels, which are specifically (and reflectively) post-perceptual.

The well known fool's gamble called "Three Card Monte" is based on an optical illusion. In this "game," the dealer displays three cards, two red and one black, so that the observer knows their starting positions as they lay on a surface. The dealer then rapidly moves them back and forth across each other, and invites the observer to pick the black card. In this case, not only is "the hand quicker than the eye," but the "toss" will predictably lead the observer to be wrong about the position of the black card.¹²

The products of optical illusions can enter the courtroom through

¹¹ The illusion is also called the "Müller-Lyer Illusion," after its creator.

¹² For an explanation of the toss in "Three Card Monte" (also known as "the "Three Card Trick" or "Find the Lady"), see EDWIN A. DAWES & ARTHUR SETTERINGTON, *THE ENCYCLOPEDIA OF MAGIC* 118-20 (1989).

testimony. However, the possibility of their existence is sufficiently well known that cross-examination may be an appropriate tool to deal with the errors in conclusion they may cause, at least in theory. The same cannot be said of “cognitive tunnels.”

Cognitive tunnels are hard-wired heuristics that operate at the level of conscious reflective analysis and reasoning. They are like optical illusions in two important respects—they lead us to wrong conclusions from data, and their apparent rightness persists even when we have been shown the trick. They are not only logical errors, they are logical errors that resist revision by information that (logically) ought to be sufficient to dispel them, and this phenomenon is independent of intelligence.

Many cognitive tunnel phenomena seem to involve probability estimates and risk judgments. Humans appear to be wired to resist treating some classes of phenomena that they encounter probabilistically even when that is clearly the optimum way to deal with those phenomena, at least for the purposes of the modern world.¹³

While cognitive tunnels resist modification through learning, they are not impervious to it. Humans, or some of them, display an ability to learn to deal with a probabilistic world. Those that do have a substantial advantage over those that don’t in many activities. In gambling games, for instance, those that can accurately figure and act upon probabilities can turn those who can’t into what has been referred to as “money pumps,” doomed to repeat judgment errors and lose as long as they play the game. Indeed, with some games, the only rational course of action is not to play.

It is one thing to allow adults to lose their own money through erroneous conclusions arrived at through cognitive tunnels. Personal freedom includes the freedom of adults to be suckers to a very great degree. However, when the legal system allows judges and juries the same freedom at the expense of parties, particularly criminal defendants, that is (or ought to be), another thing entirely. Yet not only is this the case today as it has been time out of mind, recent developments threaten to allow certain claimed experts to become pied pipers, leading factfinders into cognitive tunnels toward unjustified results in a novel and sinister way. This article is written as a warning.

In order to map this particular tunnel, and the escape from it, it is useful to begin with an analysis of one of the most famous puzzles illus-

¹³ See LAKOFF & JOHNSON, *supra* note 10; PIATTELLI-PALMERINI *supra* note 9, at 130-32. Our mental *armamentum* of default heuristics has obviously served us well in the environment in which we evolved and though it continues to serve us well in most situations, societal and technological evolution have placed us in a context in which more of our store of historically helpful heuristics may be counterproductive when applied to modern tasks. As previously noted, globally looking at the downside or the upside of this situation is like deciding whether a glass is half empty or half full. Compare Kahneman & Tversky with Gigerenzer, *supra* note 10. See the discussion in STEVEN PINKER, *HOW THE MIND WORKS* 343-51 (1997).

trating the cognitive tunnel phenomenon, the Monty Hall problem.¹⁴ This problem is named after the Master of Ceremonies of the old “Let’s Make a Deal” television show, who often presented the problem’s choices to contestants on the show. However, a full exposition of the ground conditions of the problem requires description more detailed than was given on the show itself, as follows:

Assume that you are faced with three doors. Behind one of the doors is a desirable prize (say a million dollars), and behind the other two are nothing. The million dollars has been placed behind the winning door by random selection. The Master of Ceremonies knows which door hides the prize and which doors are empty. You are allowed to choose any door you wish. Having chosen a door (say door 1 for this example), the MC must show you what is behind one of the other two doors. However, he may not show you the prize (or else the game would be over), so he always shows you an empty door (say door 2 in this particular example). That leaves two closed doors unopened, door 1 (the door you chose initially), and door 3. He then asks you if you would like to switch to door 3, or stick with door 1. What should you do?

Most people say that they would stick, since door 1 is as likely as door 3 (1 out of 2) and it was their first choice, so they have no reason to change. Of course, this is totally wrong. In this game, you should always switch, because the odds of the prize being behind door 3 are actually two out of three.

The Monty Hall problem is like three card monte in one fundamental way. It is dependant on misdirection and invited misprocessing of information. However, in three card monte, as we have previously noted, the misprocessing takes place on a level close to perception. The card toss in “Three Card Monte” depends on “sleight of hand” creating what can be accounted for as essentially an optical illusion, while something more profoundly puzzling and threatening takes place in the Monty Hall situation. There, our misprocessing is not perceptual, but cognitive. The Monty Hall problem concretely illustrates what we have above described as a hard-wired human tendency to process information according to heuristics which, under certain conditions, predictably lead even extremely intelligent people confidently to the wrong result, and which resist revision even upon the most detailed of explanation.

The misdirection involved in the Monty Hall problem depends upon the display of apparently new information apparently affecting probability of outcomes (which, by the way, makes the display apparently a core example of a relevant proffer under Federal Rule of Evidence 401). However, neither apparent condition is true, because the

¹⁴ Piattelli-Palmerini credits Martin Gardner of Scientific American with the invention of this puzzle in 1959. See PIATTELLI-PALMERINI, *supra* note 9, at 161.

displayed information is fully entailed in the original information, and therefore is neither new nor an appropriately considered Bayesian up-dater.¹⁵ The so-called information is already fully accounted for. When you choose door 1, you conclude properly that your odds of success are 1 in 3. You also know that if Monty offered to let you either keep your initial bet, or switch to “the set of doors 2 & 3” (so that you would win if the prize were behind either 2 or 3), you would jump at the chance to switch because your odds after the switch would be 2 out of 3 (remember, the rules require that Monty cannot “game” you, that is, he cannot offer his switches only when his superior knowledge allows him to know that if he can’t get you to switch you will win because you picked the right door the first time). But what else do you already know. Well, you know that even if there is a prize behind one of the doors in the set “2 & 3,” there is always at least one door with nothing behind it. You also know that Monty knows which door that is. And since Monty is not obliged by the rules to show you any particular door (such as always having to show you the lowest numbered door remaining), and he won’t show you the prize if it is behind either door 2 or door 3, this means he can always select and open the door without a prize behind it, and is indeed obliged by the rules of the game to do exactly that. So when Monty opens door 2 and shows you there is no prize behind it, he has supplied no new information relevant to the odds of winning if you switch.¹⁶ Your original choice remains a 1 in 3 winner, and your elec-

¹⁵ Bayes’ Theorem deals with the sequential revision of probabilities from a starting point (the initial or prior probability) through the integration of new probability-affecting information. See David H. Kaye & David A. Freedman, *Reference Guide on Statistics*, in REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 83, 151-59, 160 (Federal Judicial Center, Washington, D.C., 2d ed. 2000). One need not go into the details of Bayes’ Theorem to see that information fully entailed in previous information is not “new” and cannot properly be taken to change probabilities.

¹⁶ Monty has supplied some relevant new information on another matter, essentially a linguistic one, that is, the proper meaning of the words “do you want to switch to door 3.” If Monty utters the words “do you want to switch to door 3” before he opens door 2, those words represent a different question than they would if they are uttered after door 2 is opened. By opening door 2, the meaning of those words is informationally equated with the words “would you like to switch to the set of doors 2 & 3” if those words were said before door 2 was opened. What should be clear is that the probability of winning with door 1 never changes, and therefore the offer to switch to “everything that is left that might contain a prize” (which is what Monty’s post door opening offer is, even when phrased in terms of door 3, which is all that is left after door 2 is opened) also remains unchanged. The only way that Monty can change your odds on door 1 is either to show you what is behind it, or show you some other door which actually displays the only prize. Only then has he given you new probability-affecting information in regard to door 1. However, if the question *ex ante* is whether or not to exercise an option to shift to door 3 and only door 3 under any circumstances, Monty’s display of door 2 is extremely important and valuable new information, unless you know that by the rules of the game, when Monty inquires the first time about a switch apparently *ex ante*, he must offer the switch to the door that has the prize if either remaining door does, in which case the game is informationally the same as in the first instance. In the real game, he never offered the switch until displaying the door. Note also that if Monty is required to show a door at random, or to show the lowest unchosen door, the effect of an empty door is the same as in the basic game, although 1/3 of the time he will show you that

tion to switch still presents a 2 out of 3 chance of success. But the human instinct (and we use the term advisedly) to believe that the odds of winning are now 50-50 between doors 1 and 3 (and that therefore you might as well stick with door 1) is amazingly powerful.

Keeping in mind the lessons of Monty Hall, we will move a step closer to the central point of this article by examining the following hypothetical, which we call the “Two Room” problem. Assume there are two rooms containing a thousand fair roulette wheels each, with 100 slots on each wheel. The wheels are numbered 1-1000, and the slots are numbered 1-100. Start with wheel one in room 1 and wheel one in room 2. Spin the wheels simultaneously. What are the odds that both will come up with the same number? Well, the odds that they will both come up with a number predicted in advance are 1 in 100 for each wheel, and so only 1 in a ten thousand for both wheels together. But if no number need be predicted, so that we merely look at correspondence between wheels after the spin (whatever the corresponding same numbers may be), the odds are one in a hundred that they will correspond. It is as if wheel 1 were spun first. Whichever number it displays (and it must display *some* number) becomes the target for wheel 2, which wheel 2 will hit one time in 100. But of course, wheel 2 could be spun first with no change in analysis, and so the same holds for simultaneous spins: a 1 in 100 chance of what we will call the state of “paired correspondence.”

Before any spins, what is the chance that any two randomly selected pairs of wheels will manifest paired correspondence (both wheels of the pair showing the same number, though not necessarily the same number for pair 1 as for pair 2)? Again, 100×100 , or 1 in ten thousand. Add a third wheel in advance and the odds of three paired correspondence hits becomes one in a million, and so on. BUT, spin all the wheels together, and out of the two rooms, there will be (on average) ten pairs of wheels that are in the state of correspondence (odds on any one wheel, 1 in a 100 . . . out of a thousand pairs spun, average expected corresponding pairs, 1000 divided by 100, or 10). After the fact, it is easy to go through the two rooms, identify and produce all ten paired wheels, and then claim that the coincidences mean that something wildly significant has occurred or is being shown when in fact nothing significant has occurred or is being shown. Like Monty Hall, you have presented as new information, information that was fully entailed in the original known conditions, and therefore it is not new at all and cannot legitimately have an effect on updating probabilities. It is, in a wholly accurate sense, irrelevant under the very terms of Rule 403.

you have already lost by displaying the prize. Only when he is allowed to game you by electing to offer a switch, whether or not he displays a door as part of the offer should you stick, because the only rational gaming strategy is to offer the switch only when you have already won. But then the door display is also totally irrelevant, merely misdirection.

In any circumstance which is informationally rich, that is, which has many potential information variables, multiple *post hoc* correspondences can be identified and pointed out. On top of that, as the two rooms example shows, the correspondences will not necessarily be the trivial kind easily dismissed because of obviously high base rates of occurrence (the perpetrator had two eyes, the defendant has two eyes, the perpetrator had a nose, the defendant has a nose, big deal). On the contrary, as the two room example shows, it doesn't take an extremely high number of variables to get many fairly low base rate correspondences *post hoc*, as long as you have not tied yourself to any particular variable in advance of examination. This phenomenon can create the illusion of significance in any context. It is a well-known danger in science, mainly resulting from uncareful reexaminations of preexisting data by what is referred to as "data-dredging" or "data-trawling."¹⁷

Trawl search problems are by no means unknown in the law. The most dramatic emerging trawl search problem is presented by DNA databases.¹⁸ If a recoverable and analyzable allele grouping of a semen stain in the immediate area of a rape-murder victim is limited to bands which, in combination, give a random match probability in the 1 in 10,000 range, that would at first blush seem to be excellent evidence of presence at the crime scene when it is shown that the defendant possess the same alleles. Indeed, so it would be if, as has been the usual case in the past, the match resulted when a person against whom there was other evidence was tested. However, once databases are large enough to allow trawl searches, the meaning of such a match (a so-called "cold hit"¹⁹) becomes more troublesome, because no other information is required to generate such a match. All the "one in 10,000" number tells us is that, in a candidate population of 100,000 (say, the adult male population of the city in which the murder occurred), we would expect ten persons with the same genetic markers.²⁰ The fact that a particular

¹⁷ See Charles Q. Socha & H. Andrew Rzeppiennik, *All Journal Articles Are Not Created Equal: Guidelines for Evaluating Medical Literature*, 67 DEF. COUNS. J. 61, 67-68 (2000). See generally Mark Klock, *Finding Random Coincidences While Searching for the Holy Writ of Truth: Specification Searches In Law and Public Policy or Cum Hoc Ergo Propter Hoc?* 2001 WIS. L. REV. 1007 (2001).

¹⁸ See Peter Donnelly & Richard D. Friedman, *DNA Database Searches and the Legal Consumption of Scientific Evidence*, 97 MICH. L. REV. 931 (1999).

¹⁹ Ross E. Eisenberg, *The Lawyer's Role When the Defendant Seeks Death*, 14 CAP. DEF. J. 55, 71 nn.128-29 (2001).

²⁰ This insight is sometimes rather disingenuously called the "defense fallacy," to contrast it with the "prosecutor's fallacy," which involves claiming that a random match probability of 1 in 10,000 means that the odds are 10,000 to 1 that the defendant is guilty. However, the so-called "defense fallacy" is not really a fallacy, though of course it would be fallacious to argue that, because one would expect ten persons in candidate population of 100,000 to manifest the same DNA characteristics as defendant, this affirmatively establishes that the odds of guilt are 1 in 10, no matter what other evidence there is against the defendant (which is a position so facially stupid we are unaware of anyone ever actually putting it forth at trial). See RONALD J. ALLAN ET AL., EVIDENCE: TEXT, CASES AND PROBLEMS 806 (3d ed. 2002).

DNA searchable database may contain fewer than that number (even only one) does not alter this. It would be a gross error to use the search result as the sole basis to convict the first person who turns up based on nothing else, though it would be an appropriate investigative lead. Distinguishing between its treatment as an investigative lead, and the forcing of information into the frame of the DNA match to give the appearance of other evidence, creates non-trivial problems we are only now starting to understand. However, the DNA cold hit problem is itself trivial compared to the two room problem, because you know in advance the variables for which you are trawling. You do not have the added problem of being able to select variables based on matches.

At this point the reader may be starting to get a glimmer of the direction in which we are headed. Many issues concerning the admissibility of proffered evidence turn on claims by the proponent that two events are significantly similar because of shared characteristics or details. This is particularly true in regard to those doctrines embedded in FRE 404 and its outriders and cross referents, rules 405, 607, 608, 609 (and arguably the “habit” parts of rule 406), which collectively define what is usually styled in academic writing the “propensity rule,” but which is more commonly referred to in everyday practical discourse by reference to the context and the specific problem which gave rise to the more general doctrine: “other crimes evidence.” We do not intend to essay a complete exploration of the snarled thicket that constitutes the propensity rule. That has been done, to the extent it can be done, by Professor Imwinkelried in his monumental treatise on the subject.²¹ However, in order to set the specific problem with which we will deal in its proper legal setting, it is necessary to make certain general observations on what one of us has referred to in another place as “this miasmic doctrine.”²²

The modern propensity rule grew out of rules of admission and exclusion developed almost exclusively in criminal trials, as courts tried to come to grips with intuitions concerning the relevance, weight, and potential accuracy-harmful effects of evidence concerning the kind of person a defendant might be, especially when this was to be proved by showing specific prior criminal acts of the defendant not in themselves forming any part of the charged crime.²³ It appears that it was always recognized that such proof was not wholly irrelevant, at least by the

²¹ See EDWARD J. IMWINKELRIED, *UNCHARGED MISCONDUCT EVIDENCE* (2d ed. 1999).

²² Risinger, *Heartstrings and Gore*, *supra* note 7, at 428.

²³ The foundations of the rule, rejecting the admissibility of evidence of bad character against the accused in a criminal prosecution, are traceable back at least to the early eighteenth century, but the elaboration of the modern structure of the doctrine appears largely to be a mid-nineteenth century development. See I JOHN HENRY WIGMORE, *A TREATISE ON THE ANGLO-AMERICAN SYSTEM OF EVIDENCE IN TRIALS AT COMMON LAW* § 57 (3d ed. 1940 & Supp. 2001-02) and authorities collected therein.

more reflective jurists and commentators,²⁴ though it might be so labeled as a shorthand justification for its exclusion. Instead of complete irrelevance, the problem was in the remoteness and indeterminacy of the inference, the small weight to be assigned to it, its potential overvaluation, and in many instances, its potential for inflammation or inducement of practical reduction in the applied standard of proof.²⁵ People often do not act according to the general propensities we might assign to them by observing their past behavior. If such proof were admissible, it might reduce the incentive on the part of the authorities to find and proffer more particularized and more reliable proofs of guilt. Judges suspected that juries would be prone to overvalue such evidence, and when it took the form of showing prior criminal behavior, especially behavior of the same general type as the crime charged, such overvaluation might be coupled with a conclusion that the defendant was a dangerous person who should be locked up unless it were clear that he did not commit the charged crime, practically reducing the functional standard of proof to preponderance or even lower. Consequently, the courts decided that unless there was a particularly compelling reason of policy to allow it, such evidence of "character" could not be shown, and beyond that, when there was a sufficiently persuasive reason to recognize an exception to the prohibition, proof could not (with a couple of exceptions) be made by showing prior actions.²⁶

HOWEVER . . .

Some prior actions of a defendant charged with a crime seemed relevant in more important ways than merely the establishment of the kind of person (or the kind of previous criminal) the defendant was. When prior actions, even actions which were criminal, even actions already the subject of previous conviction,²⁷ were relevant to the crime charged in such a more particularized, inferentially more specific way, it was felt that there was no longer any reason for a special rule to exclude them. So they were admissible.²⁸

Consider what this means in practice, and has meant for at least 150 years. A defendant is on trial, say for burglary. The prosecutor has

²⁴ "[Character evidence] is strictly relevant to the issue, but it is not admissible. . . ." R. v. Rowton [1865] All E.R. Rep. 549, in CROWN CASES 520, 540 (Hon. E. Chandos Leigh & Lewis W. Cave eds., 1866).

²⁵ See IMWINKELRIED, *supra* note 21, § 1.03; Justice Evidence Committee, Criminal Law Revision Commission (England), *Paper on Evidence: Admission of Accused's Record*, (No CLRC/EV/154 Feb., 1969), *quoted in* WIGMORE, *supra* note 23, § 57. See generally Roger C. Park, *Character at the Crossroads*, 49 HASTINGS L.J. 717 (1998).

²⁶ See, e.g., FED. R. EVID. 405(a).

²⁷ See IMWINKELRIED, *supra* note 21, § 1.17.

²⁸ There are various secondary restrictions which may apply, such as requirements of pretrial notice of intent to use specific instances, see IMWINKELRIED, *supra* note 21, § 8.1 *et seq.*, or various so-called "sanitization" requirements. See, e.g., *State v. Cofield*, 605 A.2d 230 (N.J. 1992). Rule 403 notionally applies. See IMWINKELRIED, *supra* note 21, § 9.1 *et seq.* In practical terms, however, the statement in the text is generally true from a functional standpoint.

a fairly weak case against the defendant, independent of any prior actions by the defendant. However, the prosecutor can, if allowed, prove prior criminal acts by the defendant, some of them burglaries. If the jury sees one or two of these, the prosecutor feels confident of a conviction. If they see none, the prosecutor sees a high risk of acquittal. The defense attorney makes a similar evaluation. The prosecutor will proffer the evidence of the prior crimes and the defense will object. They will fight over whether the prior crimes are relevant only by reasoning through “propensity” or “character,” or whether it is fair to say that they are relevant through some inferential connecting path more specific and more particularized than mere propensity. The borderline between propensity and non-propensity uses is ill-defined and indeterminate, and therefore the decision is heavily subject to non-doctrinal influences like the judge’s idiosyncratic personal views and the skills of the lawyers at marshalling facts and engaging in rhetorically persuasive forensic argument. Whoever persuades the judge to come down on their side wins the ruling, and often as a result, wins the case.

In a not-too-successful attempt to give guidance to trial judges making such rulings, appellate courts have discussed various labels for inferential relations they have found to involve non-propensity paths of inference in particular cases. The main products of this process are familiar to every second year law student as the litany of labels contained in FRE 404(b). Unfortunately, the categories themselves are indeterminate, often incoherent, overlapping, elastic, and explicitly non-exclusive. While perhaps the product of good intentions, the effect of such category labels has at least as often been counterproductive to the process of analysis in later cases as it has been helpful, and has neither changed what is at stake, nor changed the generally unpredictable nature of such courtroom disputes even though it has largely determined the rhetoric with which the disputes are fought.

One of the labels on the list is “identity,” which in its main use can be taken to mean “defendant clearly did the uncharged crime, and there is something about both the charged and uncharged crimes which makes it tenably rational to infer that whoever did the uncharged crimes also did the charged crime, not based upon the propensities of the perpetrator, but on the particular similarities in the two crimes and the resulting increase in probability that two crimes with such similarities were both been committed by the same person, as opposed to two different criminal perpetrators.” This form of “identity” argument is sometimes labeled “*modus operandi*,” sometimes “signature crimes.”

At first blush, the reasoning seems straightforward enough, and so it can be in the most extreme examples. Consider a hypothetical variation on the serial killings in the book and movie *Silence of the Lambs*.²⁹

²⁹ THOMAS HARRIS, *THE SILENCE OF THE LAMBS* (1988). The film adaptation of the book

Suppose victim 1 is killed and buried with a Death's Head Moth³⁰ pupa in his mouth. Defendant is arrested, and pleads guilty to murder in return for a life sentence, because he is faced with overwhelming evidence—surveillance camera photos of the abduction by a single person easily identified as the defendant, fingerprints with 20 points of identity found on the murder weapon and on the defendant's skin, a souvenir video of the murder made by a camera built into the defendant's van which shows only the defendant, possessions of the victim found in the defendant's pocket, the defendant's arrest at the burial site as he was finishing the burial, and an uncoerced confession containing many details only later found to be accurate, but omitting any reference to the pupa. However, it is never established where he got the rare pupa, or even specifically that he ever had one. After he is committed to prison on a life sentence, another body is discovered in an adjoining state with a pupa in its mouth, which body, it can be established, was buried months in advance of the defendant's arrest for the first murder. In such a case, it seems reasonable to infer fairly confidently that whoever killed one victim killed both, and since defendant killed one of them, the evidence should be admissible in regard to his guilt upon his trial (perhaps now a capital trial) for the second murder. In such a case, it would be expected that virtually any judge undertaking a conscientious decision pursuant to existing evidence doctrine would admit the evidence.

One must note that the inference of commission by the same perpetrator is very strong here, but not inevitable. Perhaps the defendant was a member of a group of two or more for whom a murder with the pupa detail was a form of initiation. However, the likelihood of such an alternative hypothesis being true in the absence of particular evidence is barely greater than that of the claim that both murders were committed by Martians disguised as the defendant who then brainwashed him to give a false confession. However, one must be extremely careful to identify the variables upon which this strength of inference depends. First, we are dealing with only two linked variables, the pupa, and its placement in the mouth of the victim. As to the presence in the mouth of rape-murder victims of foreign objects, we are unlikely to know very confidently how common or uncommon such a phenomenon is. However, it would be surprising if it were anything near universal, or else we (at least we are likely to believe) would have heard. The pupa, however, has a base rate of availability so low that it jumps out as important when viewing the first murder *even before the second murder is discovered*. In addition, there is a virtually perfect correspondence between the particularities of both variables. By this we mean that the variables

was released in 1991 by Orion Films.

³⁰ *Acherontia atropos*. [NEED MORE EXPLANATION HERE]

have not been generalized and repackaged as generally the same even though they are specifically clearly different in perhaps important detail. The pupa was the same rare species in both cases, and the placement was in the mouth in both cases.

But consider what the enterprising prosecutor might do if the pupas were from different insects and the placement was in the mouth in case 1, and in the rectum in case 2. He could, and predictably would, characterize these things as essentially the same: insect pupae in bodily orifices. You can see where we are going. By expanding the criteria of correspondence, we can claim relevant similarity in the face of specific difference virtually without limit, until the thread of tenability snaps because the family resemblances being asserted are so general that they have ceased to be unusual and become obviously common (the murderer in crime 1 used a weapon, this murderer used a weapon). However, some lawyers display considerable skill in rhetorically concealing how generally they have construed their claimed similarities. And, of course, the more generic the similarities, and the less they would have struck one as diagnostic in advance of comparison even if specific, the more we have to contend with the “two room” problem described above.

None of this has been well dealt with in the exposition of evidence doctrine, either by the courts or by commentators, and such insightful exposition as there has been generally does not significantly influence the course of argument in the average courtroom. However, as long as the game was merely a game of lawyer-asserted speculation about baserates, and lawyer rhetoric in the packaging of similarities and differences, perhaps judges and juries subjected to the arguments had some chance to make rational common-sense assessments (although to the extent the “two room” trawl problem partakes of the Monty Hall cognitive tunnel, it may be overly optimistic to believe this). At any rate, a full-scale exposition of this complicated problem of functional irrationality in standard practice, and its impact on results, is beyond the scope of this article. However, the problems of standard practice are now potentially subject to a powerfully synergistic new factor³¹—a new brand of

³¹ In 1997, FBI offender profiling pioneer John Douglas, the father of such claimed expertise, described his admission in the 1993 California capital prosecution of Cleophus Prince as “a growing trend in courts across the country.” The Prince case (no reported opinion), another unreported California case, *People v. Bogard*, San Diego no cd10027 (1995) (described in STEPHEN G. MICHAUD WITH ROY HAZELWOOD, THE EVIL THAT MEN DO: FBI PROFILER ROY HAZELWOOD’S JOURNEY INTO THE MINDS OF SEXUAL PREDATORS 188-97 (1998) [hereinafter, MICHAUD & HAZELWOOD, EVIL]), and three of the four reported decisions dealing with the claimed expertise, *State v. Russell*, 882 P.2d 747 (Wash. 1994), *State v. Code*, 627 So. 2d 1373 (La. 1993), *State v. Pennell*, 583 A.2d 1348 (Del. Sup. Ct. 1990), have all approved of its admission, and we know of no case categorically rejecting such a witness. This article concentrates on the facts and disposition of the fourth case, *State v. Fortin*, because it is the most recent. The opinion in *Fortin* was the most extensive and reflective, it appeared to reject the claims of expertise, and yet the result was, functionally, virtually the same on remand as in the other cases. See analysis *infra* notes

asserted expertise that claims to be able to identify the right set of variables diagnostic of similarity in such cases, to give at least general testimony on their individual baserates and proper combination, and even to conclude more reliably than the jury that two crimes were committed by the same person. This asserted expertise associated with certain practitioners of what has come to be commonly known, even in the popular culture of television, as “criminal profiling.” The name they give this new asserted expertise is “linkage analysis.”³² If the reader does not immediately understand how powerful a force such a claimed expertise may be, even when there is substantial reason to doubt the claims of those asserting it, consider the murder prosecution of Steven Fortin.

I. *STATE V. FORTIN*: A TRUE-CRIME NARRATIVE³³

In the mid-1990s, the mile-long stretch of old U.S. 1 which runs south from the East Jersey State Prison through the Avenel section of Woodbridge, New Jersey had seen better times, though not recently. It was defined mainly by its multiple motels, which had become the housing of last resort for an ever-changing assortment of welfare mothers, prostitutes, dopers, drug dealers, day laborers, layabouts, and paroled sex offenders of every age, race, sex and description.³⁴ On August 11, 1994, 25-year-old Melissa Padilla was one of the mix, living in a room at the Gem Motel with her four children, ages 2-5, and her boyfriend, Hector Fernandez.³⁵ Unemployed, they supplemented her income by

278-87 and accompanying text.

³² Though virtually all of the theory and methodology of this claimed expertise are attributable to John Douglas, see *supra* note 31, the name “linkage analysis” appears to be a coinage of Robert R. (Roy) Hazelwood, another of the FBI’s profiling pioneers. As Hazelwood has recently said, “[i]n cases where no reliable witnesses, or physical evidence, are available, [linkage analysis] can be a critical factor in establishing guilt or innocence.” STEPHEN G. MICHAUD & ROY HAZELWOOD, *DARK DREAMS: SEXUAL VIOLENCE, HOMICIDE AND THE CRIMINAL MIND* 194 (2001) [hereinafter MICHAUD & HAZELWOOD, *DARK DREAMS*].

³³ All the facts set out below have been carefully assembled from trial transcripts, witness statements, police reports, and other litigation material. An outline of the major facts, or at least what the New Jersey courts found to be major, can be gleaned from the Appellate Division opinion, 724 A.2d 818 (N.J. Super. Ct. App. Div. 1999), and the opinion of the New Jersey Supreme Court, 745 A.2d 509 (N.J. 2000). Conflicts in information have been resolved in the way that appeared most likely to the authors. All cited sources are on file with the authors. Neither of the authors is in any way associated with either the defense or the prosecution in the *Fortin* case. Both the defense and the prosecution were contacted and asked to provide documents for academic use. The defense allowed us to copy the transcripts and the various motion papers and appendices. The prosecution refused to allow access to any documents.

³⁴ See May 4, 1998 Hearing Transcript, *State v. Fortin* [hereinafter *May 4 Hearing*], Vol. I, at 45-48, 56-72 (testimony of Lawrence Nagle) (on file with authors); Trial Transcript, *State v. Fortin*, Nov. 14, 2000, at 71-72 (testimony of Trent Eubanks) (on file with authors). From this point on, all references to the Trial Transcript will be stated as “TT.”

³⁵ See TT, Nov. 3, 2000, at 21 (testimony of Carmen Gonzalez, Melissa Padilla’s mother).

³⁷ *May 4 Hearing*, *supra* note 34, at 56 (Nagle testimony).

dealing minor amounts of dope out of their room.³⁷

Sometime around 11:00 p.m., Melissa had gone down to the office to use the payphone to order pizza³⁸ (there were no phones in the rooms).³⁹ For one reason or another, this had not worked out, and the kids were still hungry.⁴⁰ Leaving Hector (who may or may not have been the father of the last two)⁴¹ with the children, Melissa set out to buy food at the Quik Chek convenience store,⁴² about 350 yards north along route 1, at the corner of Avenel Street.⁴³ She never came back.

Sometime around midnight,⁴⁴ Hector began to worry. The trip to the Quik Chek should have taken Melissa a half an hour at most and she had been gone much longer.⁴⁵ Hector decided to go look for Melissa.⁴⁶ He left the children in the room and went to the front desk to ask Anita McKenzie, the night clerk, if she had seen Melissa.⁴⁷ McKenzie told him that she hadn't seen Melissa since Melissa had unsuccessfully tried to order pizza on the payphone, and left for the Quik Chek.⁴⁸ Hector's friend, Trent Eubanks, who had driven him to New York earlier in the day to buy marijuana,⁴⁹ was also in the office.⁵⁰ He asked Eubanks to help him look for Melissa.⁵¹ They walked all the way to the Quik

³⁸ See TT, Nov. 3, 2000, at 47-48 (testimony of Anita Mackenzie).

³⁹ See *id.* at 45.

⁴⁰ See *id.* at 47-48 (Nagle testimony); *May 4 Hearing, supra* note 34, at 25-26.

⁴¹ The father of the two boys was Francisco Domicilio. A love token (a necklace with the gold letters spelling out the name "Rhonda") given by Domicilio to the victim, Melissa Padilla, was always worn. The token was apparently taken from her body by her killer. See TT, Nov. 3, 2000, at 21-22 (testimony of Carmen Gonzalez). The record is silent concerning the paternity of the two girls. See *id.*

⁴² See TT, Nov. 3, 2000, at 52 (testimony of Anita Mackenzie); *May 4 Hearing, supra* note 34, at 25 (Nagle testimony).

⁴³ See TT, Nov. 3, 2000, at 92-93 (testimony of Investigator James O'Brien).

⁴⁴ The times given are best estimate reconstructions. There are two points in time that can be fixed with some precision. The time on the victim's cash register receipt from the Quik Chek was 11:29 p.m. See *May 4 Hearing, supra* note 34, at 25 (testimony of Nagle). The call went out over the police radio at 12:55. The walk to the Quik Chek should have taken about six minutes, TT, Nov. 3, 2000, at 93 (testimony of James O'Brien), with perhaps a couple of minutes more allowed for good measure, plus five to ten minutes or so to select and have prepared items of food. See *May 4 Hearing, supra* note 34, at 20-21 (testimony of Nagle); TT, Nov. 3 2000, at 116-17 (testimony of O'Brien); TT, Nov. 14, 2000, at 15 (testimony of Sgt. Jos. Joraskie). Thus, it appears likely that she left for the Quik Chek around 11:10-11:15. Allowing a minute to pay and leave after the time on the register receipt, and another three minutes to reach the vacant lot, it appears that she encountered her fate around 11:33. The entire episode is unlikely to have taken more than 5-10 minutes, so the murderer was probably walking away by 11:45-11:50. It does not seem that Hector went looking for Melissa until around midnight, or after. See TT, Nov. 3, 2000, at 53 (Mackenzie testimony).

⁴⁵ See *May 4 Hearing, supra* note 34, at 26 (testimony of Nagle summarizing the Hector Fernandez statement).

⁴⁶ See *id.*

⁴⁷ See TT, Nov. 3, 2000, at 57 (Mackenzie testimony).

⁴⁸ See *id.*; TT, Nov. 14, 2000, at 37 (testimony of Trent Eubanks).

⁴⁹ See TT, Nov. 14, 2000, at 29-30 (testimony of Trent Eubanks).

⁵⁰ See *id.* at 37.

⁵¹ See *id.*

Chek,⁵² passing in front of the Americana Motel and the Premium Diner, across Wiley Street, along the dirt path in front of the vacant lot that was the site of some long-running construction project, past the Avenel Motel and across Avenel Street,⁵³ finally reaching the little store without seeing any sign of Melissa.⁵⁴ They then returned to the Gem.⁵⁵ Eubanks offered to drive Hector around looking for Melissa, if he could get his car started.⁵⁶ The car had conked-out at the end of the New York trip and he had left it sitting on Wiley Street with the hood up.⁵⁷ Eubanks borrowed the battery from Anita's car and went to try and start it.⁵⁸ Hector asked Anita whether her two boys, Antoine, aged five and Christopher, aged eleven, (who were still awake) could help him go look for Melissa.⁵⁹ She agreed.⁶⁰

Hector sent the boys ahead, and after a while he again started out along the path to the Quik Chek.⁶¹ He saw the boys coming back from the Quik Chek, at which point they had seen nothing of Melissa.⁶² But as Hector passed the vacant construction site just north of Wylie Street he noticed a couple of shopping bags near the south side of four eight-foot-long, 30-inch-diameter sewer pipes which were on the ground awaiting installation.⁶³ The pipes were lying next to each other perpendicular to Route 1 about ten feet from the roadway. The shopping bags were on the ground a few feet from the pipes. They looked new, and food from them was strewn on the ground.⁶⁴ Hector went over and bent down to inspect the food.⁶⁵ It was then he saw Melissa's feet sticking out the east end of the northernmost pipe, the end away from the road of the pipe nearest to the Quik Chek.⁶⁶ She was covered in blood, naked from the waist down, motionless and silent.

Hector pulled Melissa out of the pipe and began yelling at her to wake up, and pounding on her chest.⁶⁷ When he got no response he got

⁵² See *id.* at 37.

⁵³ All details of the area layout are taken from the O'Brien testimony. See TT, Nov. 3, 2000, at 92-99.

⁵⁴ See TT, Nov. 14, 2000, at 38 (testimony of Trent Eubanks).

⁵⁵ See *id.*

⁵⁶ See *id.*

⁵⁷ See *id.* at 31-32.

⁵⁸ See *id.* at 38-39.

⁵⁹ See *id.* at 37; TT, Nov. 3, 2000, at 49-50 (Mackenzie testimony).

⁶⁰ See TT, Nov. 3, 2000, at 50 (Mackenzie testimony).

⁶¹ See *id.* at 54 (Antoine Mackenzie testimony). Hector Fernandez was not available to testify. See also TT, Nov. 28, 2000, at 27-28 (statement of Hector Fernandez read into record by Lawrence Nagle at the request of the defense because Fernandez did not show up at trial and could not be found).

⁶² See TT, Nov. 3, 2000, at 60.

⁶³ See TT, Nov. 29, 2000, at 45 (statement of Fernandez, as read by Nagle). For layout details from O'Brien, see *supra* note 53.

⁶⁴ See O'Brien, *supra* note 53, at 114-15.

⁶⁵ See TT, Nov. 29, 2000, at 45 (statement of Fernandez, as read by Nagle).

⁶⁶ See *id.*

⁶⁷ See *id.* at 46-47.

up and began yelling for Eubanks, whom he had just seen on Wylie Street with his head under the hood of his car trying to fix it.⁶⁸ Eubanks heard the yelling and ran over to Melissa's body.⁶⁹ After seeing what had happened, Eubanks ran to the Gem to call the police.⁷⁰ After calling the police, Eubanks returned to the scene. Seeing the children from Motel starting to come over, he took his t-shirt off and gave it to Hector, who used it to cover Melissa's private parts.⁷¹

The call went out over the police radio at 12:55 a.m. The first policeman on the scene was Officer Michael Dalia.⁷² Dalia approached Eubanks and Hector, who was still very excited.⁷³ Hector pointed out Melissa's body. Officer Dalia checked to make sure Melissa was dead, then turned his attention to securing the crime scene as a crowd started to gather. Other officers began to arrive.⁷⁴ All told, there were fourteen officers on scene when the lead homicide investigator Detective Sergeant Lawrence Nagle arrived at 1:22 a.m.⁷⁵ The evidence control officer, Investigator James O'Brien, had already started his crime scene examination, observing evidence in place.⁷⁶

Hector was taken back to the Gem and allowed to make arrangements for care of the children.⁷⁷ Eubanks spoke to the police, but then went to his room and was not transported to the police station to give a formal statement until a few hours later.⁷⁸ Finally, at 2:07 a.m. the Middlesex County Medical Examiner, Dr. Marvin Shuster, arrived and did his on-site examination preparatory to having the body moved to the morgue for autopsy.⁷⁹ He took a swab of the perineum⁸⁰ (presumably so that external fluid evidence wouldn't be rubbed off in transit⁸¹), and covered the hands and feet with paper bags to preserve any trace evidence on the hands, then had the body taken to the morgue. He didn't do the actual autopsy until 9:00 a.m. on August 13, twenty-seven hours later.⁸²

⁶⁸ See TT, Nov. 29, 2000, at 46 (statement of Fernandez, as read by Nagle); TT, Nov. 14, 2000, at 42-43 (testimony of Eubanks).

⁶⁹ See *id.*

⁷⁰ See TT, Nov. 29, 2000, at 46 (statement of Fernandez, as read by Nagle); TT, Nov. 14, 2000, at 44-45 (testimony of Eubanks).

⁷¹ See *id.*; TT, Nov. 29, 2000, at 47 (statement of Fernandez, as read by Nagle).

⁷² See TT, Nov. 3, 2000, at 11 (testimony of Patrolman Michael Dalia).

⁷³ See *id.*

⁷⁴ See *id.* at 11-12.

⁷⁵ See TT, Nov. 8, 2000, at 157-58 (Nagle testimony).

⁷⁶ See TT, Nov. 3, 2000, at 66-67 (O'Brien testimony).

⁷⁷ See TT, Nov. 14, 2000, at 20 (Joraski testimony).

⁷⁸ See *id.* at 22, 89-91 (Eubanks testimony).

⁷⁹ See *May 4 Hearing*, *supra* note 34, at 96-98 (testimony of Marvin Shuster, M.D.).

⁸⁰ See *id.* at 98.

⁸¹ See TT, Nov. 3, 2000, at 73 (O'Brien testimony). All other swabs were taken at autopsy. See *May 4 Hearing*, at 99 (testimony of Marvin Shuster, M.D.).

⁸² See Autopsy Report on Melissa Padilla by Marvin Shuster, M.D., Middlesex County, N.J., Medical Examiners Office, Sept. 20, 1994, at 1 [hereinafter *Autopsy Report*].

The autopsy documented in detail what was generally apparent as Melissa lay on the ground that night. She had been brutally beaten about her upper face. Blood covered her face, was in her hair, and covered the front of her shirt.⁸³ Her eyes and forehead were black-and-blue and swollen, her nose was broken, and there were lacerations and abrasions on her face.⁸⁴ There were bruises on her neck on external examination, and the hyoid bone was fractured and there were numerous areas of internal hemorrhage in the neck tissues.⁸⁵ There were also abrasions on the front and back of her knees and lower legs. Only after the caked blood was cleaned from the body did the medical examiner observe other marks on Melissa's chin and her left breast,⁸⁶ a "few" of which he thought "had a general, vague appearance of possible bite-mark-type".⁸⁷ A dentist, Dr. Jay Kartagener, was called in and concluded that they were probably bite marks.⁸⁸ The M.E. found a few sperm in her vagina, but only a few of those were intact and many were only sperm-heads.⁸⁹ He also found some small superficial lacerations around her anal area with fresh blood present.⁹⁰ The M.E. determined that Melissa's death was the result of assault and strangulation resulting in asphyxiation.⁹¹

In the days and weeks that followed, the detectives assigned to the Padilla murder investigated a variety of suspects,⁹² but the investigation just did not seem to develop any traction. Obviously somebody had encountered Melissa while she was on her way back to the Gem with her Quik Chek purchases. Obviously the interaction between them had been relatively brief and relatively quiet until she was put out of commission, whether she was dead then or not. The initial assault had almost certainly happened where the groceries were scattered around, because her sandals were there also.⁹³ That was a bit odd, because the groceries were about ten feet off the path, just south of the four sewer pipes.⁹⁴ There wasn't much blood outside the pipe, so she had to have been moved pretty far from the point of attack pretty quickly.⁹⁵

It also seemed to be a reasonably safe bet that the attacker or attackers were male. More than one? Maybe, but probably not. The

⁸³ See *id.*

⁸⁴ See *id.* at 2-3.

⁸⁵ See *id.* at 2, 4.

⁸⁶ See *May 4 Hearing*, *supra* note 34, at 99-101 (testimony of Marvin Shuster, M.D.).

⁸⁷ *Autopsy Report*, *supra* note 82, at 5.

⁸⁸ Report of Dr. Jay Kartagener, D.M.D., Nov. 1, 1994.

⁸⁹ See *Autopsy Report*, *supra* note 82, at 7.

⁹⁰ See *id.* at 3.

⁹¹ See Death Certificate of Melissa Padilla, signed by Marvin Shuster, M.D.

⁹² See TT, Nov. 8, 2000, at 128-35 (Nagle testimony).

⁹³ See TT, Nov. 3, 2000, at 116 (O'Brien testimony).

⁹⁴ See *id.* at 68, 115.

⁹⁵ The position of her sandals shows that the victim was literally knocked out of her shoes by her attacker. The groceries were "5-10 feet" south of the pipes, and she ended up in the northern most pipe, bypassing three other 30 inch pipes, a distance of 15-20 feet. See *May 4 Hearing*, *supra* note 34, at 20 (Nagle testimony).

strangulation and sexual assault seemed to have taken place largely in the pipe, where there was barely room for one attacker. However, it was not clear if she met her attacker on the path, if he overtook her, or if he had been hanging around the pipes. It was not clear if she had known him. It was not clear if the reason for the initial contact was an attack looking for money, an attack with a pure sadistic motive, or if it was triggered by some interchange with the victim. There was probably some interchange, because she had left the path with her bags before the blitz. Whatever the original impetus, the attacker had to be pretty strange. He had beaten her face viciously, apparently with his fists, stuffed her in the pipe, beaten her some more, strangled her, ripped off her pants, apparently bit her on the breast and chin, maybe shoved something in her anus, stolen her jewelry, then walked off toward Wiley Street covered in blood, carrying her shorts and panties in one hand and one of her sandwiches from the Quik Chek in the other.⁹⁶ He had tossed the shorts and panties into a tall bush on Wiley Street, then taken a bite out of the sandwich and left it sitting on a slate fence on the corner of Wiley and Jansen Avenue, as he ambled off into the night.⁹⁷

The trouble was, the neighborhood was full of people who, if you found out they did it, you wouldn't be that surprised. And they came and went all the time. The M.E. had taken swabs from her vagina, her anus, and her mouth.⁹⁸ He said there was not much evidence of semen. As previously noted, the vaginal swab had a few sperm, but they were mostly non-motile and broken and looked like the leftovers from a day or two before the murder. The crime scene technicians had also picked up a smoked cigarette from the west end of the pipe she was stuffed into, the end near the road.⁹⁹ There was no telling how long it had been there though. The lot was littered with cans, bottles, cigarettes, and various trash.¹⁰⁰ Still, maybe the hair and fiber guys would find something, or maybe she'd hurt the guy and he'd bled. Maybe the bitemarks on her breast and chin would lead somewhere, assuming they were actually bitemarks. The M.E. hadn't seemed that sure to begin with. If there were ever any forensic results, there were plenty of guys to compare them to, Hector and Eubanks and the seven or eight area perverts they had interviewed,¹⁰¹ but, to be honest, they didn't really feel

⁹⁶ See *id.* at 21.

⁹⁷ See *id.*

⁹⁸ See *Autopsy Report*, *supra* note 82, at 7.

⁹⁹ See TT, Nov. 3, 2000, at 145 (testimony of O'Brien).

¹⁰⁰ See *May 4 Hearing*, *supra* note 34, at 49 (Nagle testimony).

¹⁰¹ Christian Frederick, the clerk at the Quik Chek who waited on Melissa, was a paroled pedophile convicted of targeting boys, but he was in the store at the time of the murder. His roommate, Harry Thomas, was a convicted rapist. Then there was Jeffrey Blain, who had forcibly raped a 13-year-old girl, Robert Crowell, who had sexually molested his 8-year-old daughter, Jeffrey Norich, who was another homosexual pedophile, Mark Anderson, convicted of sexual assault, and that was just the known locals. Even Melissa Padilla's ex-husband, Carlos Qualles, had a sexual assault conviction, but that was a jailhouse homosexual assault. The police also investi-

strongly about any of them. Most did not seem very violent, though there were a couple with violent attacks in their background. They had all co-operated—no one had “lawyer-ed-up.” Most could account for the time of the murder pretty persuasively, all voluntarily gave blood and hair samples,¹⁰² and most took and passed police polygraphs. Not much to go on. So they waited for the lab results.

The results from the state police lab weren’t very helpful either. Whoever smoked the cigarette was a secretor with blood type A,¹⁰³ but it still was not clear that the cigarette was connected to the crime. No other blood type except the victim’s was found. No fiber evidence had been turned up on her body or clothes.¹⁰⁴ All of the pubic hair and all the head hair found had been hers, so far as anybody could prove, with two exceptions.¹⁰⁵ These were a pubic hair found on her stomach, and a short brown head hair. The head hair could have come from almost anywhere.¹⁰⁶ The pubic hair looked potentially helpful. The attacker hadn’t spurted much semen, if any, but that didn’t mean he might not have taken his pants down, or reached in his pants to rub himself and hauled out a pubic hair. Still, the tee shirt of that guy Eubanks had been put on the body, and who knows how many people’s pubic hair might have been on it—he had been living by moving between various friends, including lady friends, almost nightly.¹⁰⁷ Plus, it wasn’t clear that the victim limited herself to her boyfriend.¹⁰⁸ The hairs didn’t match any of the suspects they had investigated. Without DNA, single hairs didn’t yield very strong evidence anyway. Even if it matched someone, it might match half the world. What they had wasn’t going to convict anyone. They could have a forensic dentist look at the bite-marks and the “suspects” teeth, but that was expensive and they really didn’t think any of them were the guy anyhow. Maybe DNA might help.

It wasn’t until January of 1995, after the State Lab had finished whatever it could do,¹⁰⁹ that the various swabs and the cigarette were submitted to Cellmark, Inc. (“Cellmark”), a leading private DNA identi-

gated Theophes Spurlock, who was known to have been abroad in the area that night. All of these details were testified to by Nagle. See TT, Nov. 8, 2000, at 128-37.

¹⁰² See *id.*; TT, Nov. 3, 2000, at 89-90 (O’Brien testimony).

¹⁰³ See TT, Nov. 15, 2000, at 19 (testimony of Patricia R. Prusak, N.J. State Police Laboratory). A secretor is someone whose ABO blood group can be determined from bodily fluids other than blood. See *id.*

¹⁰⁴ See TT, Nov. 14, 2000, at 206 (testimony of Theodor Mozer, N.J. State Police Laboratory).

¹⁰⁵ See TT, Nov. 14, 2000, at 200-07 (Mozer testimony).

¹⁰⁶ See TT, Nov. 14, 2000, at 204-05 (Mozer testimony).

¹⁰⁷ See *id.* at 49-50 (Eubanks testimony).

¹⁰⁸ The DNA from the few sperm on the vaginal swab matched neither Hector nor Fortin. Rule 104 Hearing, Sept. 22, 2000 at 139-42 [hereinafter *Rule 104 Hearing*] (testimony of Paula Yates, Cellmark Diagnostics). The failure of the control dot on this test did not affect the accuracy or significance of the appearance of an unaccounted for allele in this test. See *id.*

¹⁰⁹ See *id.* at 23 (O’Brien testimony). The State Police Laboratory issued two reports, one on August 2 and the other on October 13, 1994. See TT, Nov. 19, 2000, at 11 (testimony of Prusak).

fication laboratory, for DNA testing.¹¹⁰ By that time, the investigation was pretty much on a back-burner. The DNA submissions were mostly just to have the information on hand in case something turned up. They didn't even bother to test samples from most of the "suspects," just Hector the boyfriend,¹¹¹ and they didn't expect anything from that. And they got what they expected. In fact, the report that came back from Cellmark in March looked less potentially helpful than usual. Most of the material seemed to be the victim's DNA. What wasn't hers seemed to have glitches with various control tests, and a lot of the samples might or might not have been mixtures,¹¹² which can create real problems.¹¹³ They were dead ended. No suspects, no leads.¹¹⁴

Then, on April 11, 1995, the phone rang.

Recently-promoted Lt. Lawrence Nagle of the Middlesex County Major Crimes Unit picked up the phone. He had been in charge of the Padilla investigation since the beginning.¹¹⁵ The call was from Detective Theodos of the New Jersey State Police Major Crimes Unit.¹¹⁶ They had received a call earlier from the Maine State Police asking them to do a background investigation on a Steven Fortin, age 30, who had been arrested in Maine for an April 3 attack on a female Maine State Trooper during a routine traffic encounter. Fortin had given Woodbridge, New Jersey as his most recent home address.¹¹⁷ The attack Fortin was charged with in Maine was pretty unusual. Without warning, he had repeatedly punched the Trooper in the face, strangled her with his hands (though she survived), pulled off her sweatpants (she was off-duty and out of uniform, on her way home, though driving a marked patrol car) sexually assaulted her with his hands both vaginally and anally, and bit her on the breast and chin. And when the state police had contacted Dawn Archer, whom Fortin had identified as his girlfriend, she put Fortin, drunk and angry, within two hundred yards of the Padilla murder scene a little over an hour before Padilla was killed.¹¹⁸ Theodos just thought Nagle would like to know.¹¹⁹

Now this was no doubt the hottest suspect they had had, a violent biter who was a local. He was also capable of killing, as he had spent

¹¹⁰ Report of Cellmark Diagnostics, March 24, 1999.

¹¹¹ See *id.* at 3.

¹¹² See *id.* at 2.

¹¹³ See *id.* at 3.

¹¹⁴ See TT, Nov. 8, 2000, at 137 (Nagle testimony).

¹¹⁵ See *id.* at 125.

¹¹⁶ See *May 4 Hearing*, *supra* note 34, at 79 (Nagle testimony).

¹¹⁷ See Fortin Statement to Michael Mitchell, Maine State Police Dept., April 4, 1995, at 4 [hereinafter *Fortin Statement*]. He also talked about his attachment to Dawn Archer at length. See *id.* at 9, 22-25.

¹¹⁸ See *May 4 Hearing*, *supra* note 34, at 35-36 (Nagle testimony). The content is inferred from Archer's story. See *infra* note 121 and accompanying text.

¹¹⁹ See *May 4 Hearing*, *supra* note 34, at 31-32 (Nagle testimony).

seven years in jail for the stabbing death of his own brother.¹²⁰ Nagle and his team turned their full attention to Mr. Fortin.

Dawn Archer's full story was this:¹²¹ She was 28-years-old when she met Steven Fortin through friends in April of 1994 and they had moved in together almost immediately.¹²² In August of 1994, they were living at the Douglas Motel,¹²³ which was about a third of a mile north of the Quik Chek on Route 1 closer to the prison.¹²⁴ Fortin was working for a paving contractor and she was unemployed.¹²⁵

On the night of August 11, 1994, she and Fortin had left their motel and walked south on Route 1 to the Quik Chek, where they bought cigarettes.¹²⁶ They then walked south past the construction site and the other motels (including the Gem), then past the bar and restaurant called Bud's Hut, finally arriving at the Five Oaks apartments, which was across Tappen Street from Bud's Hut on Route 1. They had gone to the Five Oaks to visit Dawn's friend Charlie Bennett.¹²⁷ They arrived around 9:00 p.m and spent the next hour and a half drinking with Charlie at the Five Oaks.¹²⁸ Dawn and Fortin got drunk and began to bicker over their relationship.¹²⁹ They left Charlie's a little before 10:30 p.m and by the time they were walking through the parking lot of Bud's Hut, their argument was getting ugly.¹³⁰ At some point Fortin knocked her down (though she admits she might have done some defensive damage of her own).¹³¹ He jumped on her but she got loose and ran into Bud's Hut and called the police.¹³² The call came at 10:32 p.m.¹³³ When the police arrived a few minutes later, Fortin was gone.¹³⁴ Dawn was clearly drunk and a bit belligerent, but she did have bruises and a bloody nose.¹³⁵ She told the police her story and said that she wanted to press charges against Fortin, but then refused to sign a complaint.¹³⁶

¹²⁰ See *Fortin Statement*, at 29. Fortin pled guilty to involuntary manslaughter. See *id.*

¹²¹ Archer was interviewed formally by Lt.Nagle on April 20, 1995. See *May 4 Hearing, supra* note 34, at 35 (Nagle testimony); Statement of Dawn Archer, April 20, 1995, listed in Report of Robert R. (Roy) Hazelwood, Oct.17, 1997, at 2 [hereinafter *Hazelwood Report*]).

¹²² See TT, Nov. 8, 2000, at 20-21, 22-24 (testimony of Dawn Marie Archer).

¹²³ See *id.* at 25.

¹²⁴ See details of O'Brien testimony, *supra* note 53.

¹²⁵ See TT, Nov. 8, 2000, at 45-46 (O'Brien testimony).

¹²⁶ See *id.* at 26-27.

¹²⁷ See *id.* at 25-26.

¹²⁸ Dawn estimated that they drank at Bennett's for an hour to an hour and a half. See TT, Nov. 8, 2000, at 29 (Archer testimony). Bennett says "an hour, an hour and 45 minutes." *Id.* at 77 (testimony of Charles Bennett).

¹²⁹ See *id.* at 29 (Archer testimony).

¹³⁰ See *id.* at 30-31.

¹³¹ See *id.*

¹³² See *id.*

¹³³ The call about the Fortin-Archer altercation came in at about 10:32 p.m. See TT, Nov. 8, 2000, at 98 (Grimshaw testimony).

¹³⁴ See *id.* at 32 (Archer testimony), 99 (Grimshaw testimony).

¹³⁵ See *id.* at 100-02 (Grimshaw testimony).

¹³⁶ See *id.* at 32 (Archer testimony), 102 (Grimshaw testimony).

Dawn went to the hospital in an ambulance, but refused to go into the hospital.¹³⁷ Instead, she walked to her mother's house in nearby Perth Amboy and did not see Fortin again until Saturday, which was two days later.¹³⁸ He then talked her into moving back in with him.¹³⁹ They stayed together in the area until late December, when they went to Maine to visit his parents.¹⁴⁰ Late in January, they started back to New Jersey, stopping to visit her father in Connecticut.¹⁴¹ There they had another fight, in which she received a black eye.¹⁴² Fortin left and she did not see him again after that.¹⁴³ Dawn added that when she saw Fortin on the Saturday after the fight at Bud's Hut, he had scratches on his face, arms and chest.¹⁴⁴

On April 12, Lt. Nagle called the Maine authorities to get more details on the Maine charges.¹⁴⁵ The full details of Fortin's attack on Trooper Gardner in Maine made the New Jersey case detectives like Fortin for the Padilla murder even more:

At about 8:45 p.m on April 3, 1995, Maine State Trooper Vicki Gardner was traveling south on Interstate 95 near Pittsfield, Maine in a marked Maine State Police cruiser. She was off-duty at the time, but had received permission to use the cruiser to visit her parents near Bangor, some 90 miles to the north, and she was now returning home.¹⁴⁶ She was dressed in black nylon athletic pants, a navy blue turtleneck and a gray sweatshirt.¹⁴⁷ Her service weapon was locked in the trunk of the cruiser.¹⁴⁸

The area she was driving through was essentially wooded countryside.¹⁴⁹ The northbound lanes of Interstate 95 were separated from the southbound lanes by a wide, wooded median strip.¹⁵⁰ There were two traffic lanes north and a broad shoulder or "breakdown lane."¹⁵¹ Around mile-marker 139, Trooper Gardner saw a car stopped on the shoulder.¹⁵² Its lights were on and it was pointing north, in the wrong direction, against traffic.¹⁵³ She decided to stop and investigate.¹⁵⁴

¹³⁷ See *id.* at 32 (Archer testimony).

¹³⁸ See *id.* at 32-33.

¹³⁹ See *id.* at 33-36.

¹⁴⁰ See *id.* at 36.

¹⁴¹ See *id.* at 36-37.

¹⁴² See *id.* at 69-70.

¹⁴³ See *id.* at 68.

¹⁴⁴ See *May 4 Hearing*, *supra* note 34, at 79 (Nagle testimony).

¹⁴⁵ See *id.* at 79.

¹⁴⁶ See TT, Nov. 8, 2000, at 192 (testimony of Vicki Gardner, Maine State Police).

¹⁴⁷ See *id.* at 198.

¹⁴⁸ See *id.*

¹⁴⁹ See *id.* at 216.

¹⁵⁰ See *id.*

¹⁵¹ See *id.* at 193.

¹⁵² See *id.*

¹⁵³ See *id.*

¹⁵⁴ See *id.* at 194.

Trooper Gardner approached the driver's side of the vehicle, where a man was sitting in the driver's seat.¹⁵⁵ He rolled down his window and she identified herself as a State Trooper, which he acknowledged.¹⁵⁶ She asked him for his license, registration, and insurance card.¹⁵⁷ He produced a Maine learner's permit in the name of Steven Fortin, with what later turned out to be his parent's address, but he had no registration or insurance card, and could not legally drive alone on the learner's permit.¹⁵⁸ He explained that he had just moved up from New Jersey, that he had just bought the car, and that he thought the person who had sold it to him would have left registration and insurance documents inside.¹⁵⁹ Trooper Gardner detected the smell of alcohol on his breath.¹⁶⁰ She asked him to exit the car.¹⁶¹ When he did so, she noted that he staggered a bit.¹⁶² She decided that she had to process him on the charge of Operating under the Influence of Alcohol as well as the license, registration, and insurance offenses.¹⁶³ Fortin was being polite and cooperative at that point.¹⁶⁴ It was still quite cold outside, and she placed Fortin in the passenger seat of her cruiser.¹⁶⁵ She got in the driver's seat and finished giving him a number of field sobriety tests, some of which he passed and some of which he didn't.¹⁶⁶ At that point, she radioed in the details of the stop and Fortin's name and address to the dispatcher, and requested that the uniformed on-duty trooper assigned to the area, Trooper Stewart, come and pick Fortin up to transport him to the station for processing.¹⁶⁷ She then gave Fortin the *Miranda* warnings, which she noted at 9:03 p.m.¹⁶⁸ For the next forty-five minutes they sat in the cruiser while Trooper Gardner periodically radioed to inquire about when Trooper Stewart would arrive, and she was continually told that he had other things to do just then, but would be along as soon as he had finished.¹⁶⁹ At around 9:45 p.m., Fortin said he had a proposition for her.¹⁷⁰ She responded that she would listen if he wanted to talk but it wouldn't do any good, he was going to be taken in and would have to make bail on the charges.¹⁷¹ He then said: "My

¹⁵⁵ *See id.*

¹⁵⁶ *See id.* at 197.

¹⁵⁷ *See id.* at 195.

¹⁵⁸ *See id.* at 195-96.

¹⁵⁹ *See id.* at 195.

¹⁶⁰ *See id.* at 195-96.

¹⁶¹ *See id.*

¹⁶² *See id.* at 196.

¹⁶³ *See id.*

¹⁶⁴ *See id.* at 202.

¹⁶⁵ *See id.* at 196.

¹⁶⁶ *See id.* at 197.

¹⁶⁷ *See id.* at 199.

¹⁶⁸ *See id.*

¹⁶⁹ *See id.* at 199-01.

¹⁷⁰ *See id.* at 203.

¹⁷¹ *See id.*

proposition is that you just let me go back to my car and drive away and pretend that nothing ever happened.”¹⁷² She dismissed this as ridiculous and told him that he didn’t understand the seriousness of the offenses he was charged with, and that she felt he needed to be placed under arrest.¹⁷³ She then turned her attention back to her notebook, and the next thing she knew he had jumped across the car and slammed her head into the doorpost.¹⁷⁴ He pounded her face and she felt consciousness slipping away, as she tried to fight back and he had his hands around her throat, and she got a hand up to try and pry them loose and then she went under.¹⁷⁵

When she came to, she was laid out across the passenger side of the seat with her head up against the door.¹⁷⁶ Her pants and underwear had been pulled off.¹⁷⁷ Her sweatshirt and turtleneck had been shoved up above her breasts, and her bra pulled off.¹⁷⁸ She had a vague recollection of Fortin sticking his fingers in her,¹⁷⁹ but otherwise she was just in general pain and fear.¹⁸⁰ The car was in motion and Fortin was fighting to turn the wheel.¹⁸¹ She decided that if she were going to survive she would have to jump.¹⁸² She opened the door and started to roll out. He reached over and pushed her and she rolled free.¹⁸³ She lost a lot of skin on the pavement, but she was alive as the cruiser pulled away.

Fortin had been scared into fleeing by the arrival of Trooper Stewart, who gave chase.¹⁸⁴ Fortin continued down the highway a short distance, lost control of the cruiser and crashed it, fleeing the scene on foot.¹⁸⁵ He was found after a manhunt of some hours, hiding in the restroom of a roadside rest stop.¹⁸⁶

Fortin more or less confessed that afternoon without asking for a lawyer.¹⁸⁷ At least he admitted attacking Trooper Gardner, but claimed

¹⁷² See *id.*

¹⁷³ See *id.*

¹⁷⁴ See *id.* at 204-05.

¹⁷⁵ See *id.* at 205.

¹⁷⁶ See *id.* at 205-06.

¹⁷⁷ See *id.* at 206.

¹⁷⁸ See *id.*

¹⁷⁹ See TT, Nov. 8, 2000, at 224 (“fresh complaint” testimony of Lt. Jackie Theriault, Maine State Police); *May 4 Hearing*, *supra* note 34, Vol. 2, at 115 (testimony of Dr. Lawrence Ricci concerning the April 3, 1995 examination with history of Vicki Gardner, given May 5, 1998). Trooper Gardner did not specifically remember the digital anal penetration by the time of trial, though the objective signs were clear. See *id.*

¹⁸⁰ See TT, Nov. 8, 2000, at 206 (testimony of Vicki Gardner, Maine State Police).

¹⁸¹ See *id.*

¹⁸² See *id.* (testimony of Vicki Gardner, Maine State Police).

¹⁸³ See *Fortin Statement*, *supra* note 117, at 32-33.

¹⁸⁴ See *May 4 Hearing*, *supra* note 34, at 120-21 (testimony of Gerard Madden, Maine State Police Sgt.).

¹⁸⁵ See *id.*

¹⁸⁶ See *id.* at 123.

¹⁸⁷ See *Fortin Statement*, *supra* note 117, at 31-33.

that he was trying to fend off a sexual attack by her on him;¹⁸⁸ he denied being responsible for her sexual injuries.¹⁸⁹

The Maine authorities also told Lt. Nagle that Fortin had agreed to a plea deal on the Maine charges within a matter of days, but not before casts had been made of his teeth for comparison to Trooper Gardner's bite wounds just in case he went to trial.¹⁹⁰ These had been sent to Dr. Lowell Levine, a leading forensic odontologist and bitemark identification expert, but as a result of the plea he didn't really make any comparisons with the photographs of the Gardner wounds.¹⁹¹ Nagle directed that Levine be contacted by the New Jersey detectives and provided with the autopsy report and other details of the Padilla case, and photographs of Melissa Padilla's wounds. On April 19, Levine was asked to compare the casts of Fortin's teeth previously provided by the Maine authorities with Padilla's wounds.¹⁹² On May 3, Levine issued a report that positively identified Steven Fortin's teeth as the source of the wounds on Melissa Padilla's left breast.¹⁹³ However, such bitemark identification may not turn out to be enough for a conviction.¹⁹⁴

¹⁸⁸ See *id.* at 25-38. It seems appropriate to note that Fortin describes these claimed events in extreme and literally unbelievable detail. See *id.*

¹⁸⁹ See *id.* at 36.

¹⁹⁰ See TT, May 6, 2000, at 26 (testimony of Dr. Lowell Levine given May 6, 1998).

¹⁹¹ See *id.*

¹⁹² See Search Warrant Affidavit of Gerard Madden, April 26, 1995, §6 [hereinafter *Madden Affidavit*].

¹⁹³ See Report of Dr. Lowell J. Levine, May 3, 1995.

¹⁹⁴ This article is not about the weaknesses of the identification of marks on human flesh as human bitemarks, or the weaknesses of methods for establishing the origin of such bitemarks, if bitemarks they be. However, a number of points must be made here to put the "M.O. expertise" issue in context, and to show that its admission in Fortin's case is not so clearly harmless error. Bitemark identification evidence is extremely controversial. It gained admissibility as the result of unusual circumstances at a time when standards of admissibility of claimed expertise were generally more lax than they are today, and it has been grand-fathered by precedent in such a way as to be insulated from more searching post-*Daubert* examination. We use the term "post-*Daubert*" to refer to a general heightening of scrutiny after the Supreme Court's decision in *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579 (1993), even in jurisdictions that do not formally claim to "follow" *Daubert*. See generally D. Michael Risinger, *Navigating Expert Reliability*, *supra* note 8 (establishing the general influence of *Daubert*). On the odd history of bitemark evidence, see *id.* at 135-42; DAVID L. FAIGMAN, DAVID H. KAYE, MICHAEL J. SAKS & JOSEPH SANDERS, *MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY* §§ 30-1.3 & 30-1.3.1 (West, 2d ed. 2002). On the "grand-fathering," see *id.* § 30-1.3.2.

The *Fortin* case itself shows how difficult it can be to determine if a mark is in fact a bitemark. In the autopsy report, Dr. Shuster explains why he asked Dr. J. Kartagener, D.M.D. (whom he refers to as a "Forensic Odontologist") to view wounds on the body, noting, "A few of these lesions had a general, vague appearance of possible bite-mark-type." *Autopsy Report, supra* note 82, at 6, 5.

Dr. Kartagener came to view the body at noon on August 13, 1994, at which time he declared that except for the chin and breast wounds, "[o]ther markings on the body were not discernible [sic] at the time of examination and therefore were not useful for identifying purposes." Report of Dr. Jay Kartgener, Nov. 1, 1994, at 1. As to the chin marks, Dr. Kartagener reported, "[t]he two sets of markings on the lower chin area bear a strong resemblance to marks that could be made by a human dentition. These markings were studied extensively and compared with photographs of known bite marks. The patterns strongly resemble human bite marks and it is the

Meanwhile, Nagle and Detective O'Brien went up to Maine on April 24 to attempt to take a statement from Fortin.¹⁹⁵ Fortin did not request a lawyer, but he didn't give them much. The closest thing he came to an admission was, when he was told his teeth had been matched to the bitemarks,¹⁹⁶ Fortin replied "well, if the evidence says I did it, I must have done it, I don't remember."¹⁹⁷ He also asked if he was facing a manslaughter charge.¹⁹⁸

opinion of this Examiner that they are indeed the result of a human dentition. If that be the case, then the marks can be explained as follows." *Id.* He then went on to identify individual teeth, and even provided separation measurements and degrees of rotation. As to the breast marks, Dr. Kartaganer said, "[t]he markings on the breast may be the result of heavy and firm suckling with two teeth marks being observed, but with their shape and size too indistinct to measure." *Id.* at 2. Dr. Levine found the breast marks and the bitemarks on the left chin to be sufficient to yield a positive identification. See Report of Dr. Lowell J. Levine, D.D.S., May 3, 1995, at 2. The defense expert, the equally credentialed Dr. Sperber (both Levine and Sperber had decades of experience, both had law enforcement positions, Levine had been president of the American Board of Forensic Odontologists ("ABFO"), and Sperber had been chair of the ABFO Committee on Standards) asserted that it was unclear that either mark was a bitemark, and that if either was, they clearly did not match Fortin's dentition. See TT, Nov. 30, 2000, at 34, 47, 42-43, 49-51, 53. In his testimony, Sperber said Levine's conclusion was "totally inaccurate" and "an affront to science." *Id.* at 45, 79. In closing, the prosecutor called Sperber a liar. See TT, Dec. 5, 2000, at 103 (prosecutor's closing statement). Conflicts between Sperber and Levine have a long history. See generally *People v. Prante*, 498 N.E.2d 889 (Ill. Ct. App. 1986).

It is becoming increasingly clear that, as a general proposition, bitemark identification is shockingly untrustworthy. The results of the most recent ABFO-run blind proficiency tests showed that board-certified Forensic Odontologists given a line-up type problem were wrong about as often as they were right, and that their errors were skewed strongly toward false positives, declaring a "match" when there was none. See FAIGMAN ET AL., *supra* note 194, § 30-2.1.3. In regard to the claim that "experience" will lead to accuracy, this study, involving only experienced, board certified forensic odontologists, is one more piece of evidence that this proposition is a significant overstatement in many, if not most, contexts.

One of the reasons for such poor performance is the inherent high subjectivity of the process of judging what constitutes a bitemark, or a match. If there is one thing that seems to have been established clearly by modern cognitive psychology, it is that the more subjective an evaluative process is, the more it is subject to the inaccuracy-inducing effects of expectation and suggestion. See D. Michael Risinger, Michael J. Saks, William C. Thompson & Robert Rosenthal, *The Daubert/Kumho Implications of Observer Effects in Forensic Science: Hidden Problems of Expectation and Suggestion*, 90 CAL. L. REV. 1, 16 (2002) [hereinafter Risinger et al., *Daubert/Kumho Implications*]. In the *Fortin* case, Dr. Levine was provided with much information, irrelevant to his claimed expertise, that suggested Fortin's likely guilt, along with the specifically domain-relevant information, such as the photographs of the wounds and the casts of Fortin's teeth. Perhaps even worse, investigators traveled to Dr. Levine's office on April 19, 1995, and sat with him discussing the case while he did his preliminary comparisons and presented his initial conclusions. See *Madden Affidavit*, *supra* note 192, § 6. Such circumstances render results extremely suspect, even if arrived at with the purest of conscious intent.

¹⁹⁵ See TT, Nov. 8, 2000, at 139, 179 (Nagle testimony).

¹⁹⁶ While Dr. Levine had not yet issued a report, he had committed to this position during an April 19, 2000 meeting with detectives. See *Madden Affidavit*, *supra* note 192, § 6. This was known to Lt. Nagle when he went to Maine. See TT, Nov. 8, 2000, at 138 (Nagle testimony).

¹⁹⁷ See *id.* at 134.

¹⁹⁸ See *id.* at 144 (Nagle testimony). These statements were admitted at trial as admissions. They are quite problematical for a number of reasons. First, they were the product of a 2½ hour unrecorded interview which Officer Joraskie, in his testimony, later gave strong reason to believe utilized the standard "good cop/bad cop" methodology, with Nagle taking the role of "bad cop" and Joraskie the role of "good cop." See Hearing, Aug. 9, 2000, at 88 (suppression hearing tes-

Fortin's blood samples had been supplied to Cellmark in mid-April,¹⁹⁹ and on May 16 they issued their report.²⁰⁰ They resolved the ambiguities in regard to the left fingernail scrapings and the cigarette in favor of them being mixed samples, and assuming they were mixed samples, Fortin could not be excluded as the primary contributor of the dominant amount of the DNA on the cigarette butt, or a secondary contributor to the minority percentage of the DNA under the left fingernail.²⁰¹ The report also gave the incidence of Fortin's DNA profile for the tested markers as 1 in 3500 for the Caucasian population, and less than that for the black or Hispanic population.²⁰² However, the DNA evidence was not without its problems, and might be shown to mean a lot less than it appeared at first glance to mean.²⁰³

timony of Joraskie). Fortin declined to continue when the officers asked if they could begin to tape the interview. See TT, Aug. 9, 2000, at 22 (Nagle testimony). The effect of this interrogation technique on accuracy of admissions is debatable. Second, few notes were taken during the interview, the intent having been to obtain a taped statement after the lengthy "preliminary interview" had settled what would be in it, again, lamentably, near universal practice. In this case, the lack of tape recording and/or verbatim notes is critical, as the exact meaning of the sentence quoted from Fortin is dependant both on exactly what he was responding to, on exact phrasing, and on intonation. If, hypothetically, two words are added and an ironic tone is used, the statement becomes meaningless, or even a form of denial. The statement, "[w]ell, if the evidence says I did it, I must have done it, but I don't remember it" demonstrates this point. Third, some notorious cases of wrongful conviction have been the result of the admission of such qualified statements agreeing with police premises. A question such as "[t]he evidence says you did it; is it possible you blacked out and don't remember?" is another standard interrogation technique. See, for example, the notorious Peter Reilly case, detailed in JOAN BARTHEL, *A DEATH IN CANAAN* (1976) and DONALD S. CONNERY, *GUILTY UNTIL PROVEN INNOCENT* (1977). For the aforementioned reasons, these statements would seem to be strong candidates for exclusion under Rule 403 (New Jersey's version is identical to the Federal Rule of Evidence). In the event, no specific Rule 403 objection was made.

¹⁹⁹ See *Rule 104 Hearing*, *supra* note 108, at 126 (testimony of Paula Yates, Cellmark Diagnostics).

²⁰⁰ See *id.* at 133.

²⁰¹ See TT, Nov. 14, 2000, at 135-42 (testimony of Charlotte Word, Cellmark Diagnostics).

²⁰² See *id.* at 139. This is a misleading number when dealing with mixtures because it inaccurately suggests a random match probability that, in this case, is more than a whole order of magnitude too high.

²⁰³ As with bitmark evidence, the main focus of this article is not on the rather surprising weakness of DNA evidence under some conditions. However, once again, to establish the significance of the main point, it is necessary to understand those weaknesses as they apply to the *Fortin* case. The popular imagination considers DNA evidence to be nearly infallible. Such an assumption is misplaced. It is true that, with the exception of naturally occurring clones (i.e., identical twins), it is vanishingly unlikely that any two humans have exactly the same full sequence of DNA base-pairs in their full chromosomal DNA. And this statement, unlike similar statements for things like fingerprints, is actually based on good empirically-rooted science. However, forensic identifications are not made by comparing the entire base-pair sequence. That would be impossible with current technology in any practical way. Instead, only short sequences of base pairs are located and compared. The sequences that are used are selected for two characteristics. They are started and ended by sequences that are virtually unique and invariable, which means that they can be located easily with available technology that is specific to those locations, and that they are variable enough in the middle to yield meaningful data on identification because they come in multiple forms, called alleles. Some such "genetic markers" have only two or three alleles, while some have many more. To complicate things further, each person potentially has two forms of each allele, one from each parent (although sometimes both forms are the same

Fortin was indicted in early September of 1995, but his incarceration

form, if both parents happen to contribute the same form). Which two forms a person has of these alleles at a given marker site is his “allotype” for that site. The incidence of such alleles in the population, or in various sub-populations (since incidences do vary between groups such as racial, ethnic, or local sub-populations) is established by population studies, of which hundreds have been conducted in the last fifteen years. Obviously, the more sites you test and the more variable the alleles, the more likely a clear match will “identify” because it will exclude a higher percentage of the population of the world as possible sources. Sometimes, with enough such sites, the likelihood of a “random match,” that is, drawing such a matched contributor at random, is less than 1 in more than the population of the planet. When DNA is this good, it is very, very good indeed. See generally David H. Kaye & George F. Sensabaugh, Jr., *Reference Guide on DNA Evidence*, in FEDERAL JUDICIAL CTR., REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 485 (2d ed. 2000). However, it can only be at its best when the evidence sample (i.e., the questioned sample), such as the fingernail scraping in the *Fortin* case, is from a single source and exists in sufficient quantity to be sure that every allele will be identified and displayed by existing technology. Unfortunately, when one is dealing with small amounts of DNA material, with potentially mixed samples, and with a selection of tests with limited marker sites and small numbers of alleles at each site, results can become questionable. This is compounded by technologies that do not use instrumented perceptrons but instead use human subjective judgment to determine the presence and amount of particular alleles picked up by the tests. This is because many tests result only in visual dots or bars of similar or different intensity at or near a place where they are supposed to be on a “test strip” if a given allele is present in the mixture applied to the test strip.

The test used in the *Fortin* case was the “DQ alpha-Polymarker test,” which was the test then commonly used when small amounts of DNA had to be amplified by the use of the polymerase chain reaction process. Unfortunately, the DQ alpha-Polymarker test used only six sites of very limited allele variability. In addition, it was notoriously difficult at times to determine whether a given test result was or was not the product of a mixed sample. The only way to be sure was if dots for more than two alleles showed up at a given marker site. In the *Fortin* case, this did not happen at any of the six marker sites for either the fingernail scrapings or the cigarette. This does not, of course, eliminate the possibility of a mixture. Perhaps two contributors had only the manifested alleles, or perhaps one person’s DNA was present so weakly that one allele did not color a dot at all. Sometimes differences in intensity in allele dots at a specific marker site can indicate a mixture, but what constitutes such a diagnostic difference in intensity is both subjective and controversial, since some tests manifest differences in intensity with clearly unmixed samples. See generally W.C. Thompson, *Subjective Interpretation, Laboratory Error and the Value of Forensic DNA Evidence: Three Case Studies*, 96 GENETICA 153 (1995); William C. Thompson, *Accepting Lower Standards: The National Research Council’s Second Report on Forensic DNA Evidence*, 37 JURIMETRICS 405, 414 n.24 (1997).

Because of such problems, Cellmark no longer uses the DQ alpha—Polymarker test and, in the *Fortin* case, all of these problems were present to such a pronounced degree that there was disagreement among the first two Cellmark readers which had to be resolved by the third vote of a supervisor—an uncommon circumstance. In addition, because of the problems involved in excluding persons as contributors to mixed samples, the products rule, used to establish the rate of occurrence on the populace, so the 1 in 3500 figure given by Cellmark was essentially irrelevant. The true random match probability of all persons who “could not be excluded,” assuming the samples were mixtures, was actually 1 in 130. See TT, Nov. 30, 2000, at 170 (Shields’ testimony). And if the samples did not represent mixtures, Fortin was excluded as a source. Of course, as to the cigarette, he was pretty much excluded anyway because whoever was the “primary” smoker was a type A secretor, and Fortin was a type AB secretor. It was virtually certain that the “primary” source of the DNA, assuming a mixture, was also the person whose saliva was the source of the ABO group, and that was not Fortin. And it wasn’t clear the cigarette had anything to do with the case at all. Why it was not excluded under Rule 403 is a mystery, but no specific Rule 403 objection seems to have been made. Pages of examination and cross examination of both the prosecution’s DNA witness and the defense expert were devoted to the cigarette, but it was never even mentioned in the prosecution’s extensive closing. It appears to have served mostly to divert attention from the problems of the evidence derived from the fingernail scrapings.

tion in Maine caused his case to fall into a kind of limbo. It was not until a year and a half later, in mid-March of 1997, that things had progressed to the point of filing a notice of aggravating factors, which is the Rubicon for seeking the death penalty.²⁰⁴ In the interim, the prosecutor had a lot of time to mull his case.

There has been a lot of talk recently about the resurgence of pragmatism as a school of philosophy.²⁰⁵ Pragmatism assigns value to concepts and practices only as they are useful in the accomplishment of ends.²⁰⁶ In many ways, the most pragmatic of humans are litigators in general, and criminal litigators in particular. Once having determined that their role is to obtain a conviction (or acquittal) they will use whatever the system allows them to use which will help them to prevail, whether they themselves think it makes any sense or not. Consider the position of the prosecutor of Mr. Fortin in June of 1997. His most powerful evidence in the Padilla murder is the evidence of the details of Fortin's attack on Trooper Gardner in Maine. In a commonsense way, it seems pretty persuasive. Steve Fortin is not just a bad person, he is the kind of person who would do some pretty extreme things to a woman on the spur of the moment, things that seemed fairly similar to the Padilla facts in relatively unusual ways, involving beating, hand-strangling, biting, digital anal penetration on an apparently lifeless body, and no apparent use of his penis. If the Maine episode gets in front of the jury, Fortin will be convicted. If it is not admitted, he might not be convicted. Recall the points made about Rule 404 and the propensity rule in Part I. If the prosecutor is shrewd, and if the judge is on the right wavelength, he will have this evidence classified as admissible under 404(b). On these facts, he can probably depend on winning admission, at least enough admission,²⁰⁷ 8 or 9 out of 10 times. But judges are unpredictable and the standards are grey. Also, this kind of case may be one in which a judge would feel less comfortable being sure what is commonly present and what is not commonly present in a sex murder, which opens the door for effective argument by the defense to the judge and maybe to the jury. If there were some way to reduce these risks, the normal prosecutor would use it in a heartbeat, even if he believed it were voodoo. If phrenologists were effective tools for obtaining the admission of such prior crimes (and judges allowed them to

²⁰⁴ Notice of Aggravating Factors, *State v. Fortin*, March 18, 1997.

²⁰⁵ See *THE REVIVAL OF PRAGMATISM* (Morris Dickstein ed., 1998); Symposium, *The Revival of Pragmatism*, 18 *CARDOZO L. REV.* 1 (1996); Symposium, *The Renaissance of Pragmatism in American Legal Thought*, 63 *S. CAL. L. REV.* 1569 (1990).

²⁰⁶ That is the central tenet of the main branch of pragmatism associated with William James. See H.S. Thayer, *Pragmatism*, in 6 *THE ENCYCLOPEDIA OF PHILOSOPHY* 430, 433-34 (Paul Edwards ed., 1967).

²⁰⁷ Judges often rule in such a way as to appear judicious by excluding part of a problematical proffer, preserving the appearance of balance while letting in most of what is useful to the proponent. See Risinger, *Heartstrings and Gore*, *supra* note 7, at 430-31.

testify) he would call a phrenologist even if he believed phrenology was bunk. After all, it's not up to him to decide what is good enough and what is not. That's the judge's job. And the defense would, and often does, do the same thing. So he picks up the phone and dials the number of the Academy Group, Inc.

The Academy Group is a consulting firm giving post-retirement employment opportunities to, among others, retired criminal profilers from the FBI.²⁰⁸ One of these was Robert R. (Roy) Hazelwood. In August of 1997, Mr. Hazelwood agreed to review the Padilla and Gardner episodes in order to "form an opinion as to whether the two crimes were committed by the same offender."²⁰⁹

In order to understand what might or might not be involved in such a process, it is necessary to examine Mr. Hazelwood's career and the history of the profiling efforts of the FBI Behavioral Science Unit, of which he was a part. We will begin with the latter.

II. THE FBI BEHAVIORAL SCIENCES UNIT: RESEARCH PRACTICE AND PROFILING PRACTICE

Criminology is generally regarded as the social science that examines patterns of criminal behavior using the techniques of social science research. It may concern itself with macro-patterns of crime, such as the general rise and fall of the murder rate, or it may concentrate on more particular questions, right down to issues of individual behavior and motivation which might more comfortably be labeled criminal psychology.²¹⁰ Traditionally, criminology has been rooted in an academic setting, or in supported research done by academics.²¹¹ Rarely have law enforcement agencies themselves undertaken such research predominantly utilizing law enforcement personnel. This began to change at the FBI in the mid-1970s.²¹²

The FBI's entry into basic research was not driven by academic curiosity. From the beginning, the hope was that research could help in developing a reliable method for determining the likely characteristics of a crime's perpetrator by examining the details of the crime itself. In turn, it was hoped that such a reliable method would aid in the capture of the perpetrator by narrowing down and prioritizing the set of possible perpetrators to be investigated.²¹³ This approach to criminal investiga-

²⁰⁸ See <http://academy-group.com> (last visited June 18, 2002).

²⁰⁹ *Hazelwood Report*, *supra* note 121, at 1.

²¹⁰ See GEORGE B. VOLD & THOMAS J. BERNARD, *THEORETICAL CRIMINOLOGY* 1-2 (3d ed. 1986).

²¹¹ See *The Men Who Murdered*, in 54 FBI LAW ENFORCEMENT BULLETIN 8, Aug. 1985, at 2.

²¹² See *id.*

²¹³ "The goal of the profiler is to provide enough information to investigators to enable them to limit or better direct their investigations." Richard L. Ault, Jr. & James T. Reese, *A Psycho-*

tion has come to be known as “criminal profiling,” “offender profiling,” or more popularly, just plain “profiling.”²¹⁴

The pioneer of profiling at the FBI was Howard Teten, an FBI agent with an interest in applied psychology who had studied with Dr. James Brussel, a New York forensic psychiatrist whom many regard as the first successful profiler.²¹⁵ Teten was assigned to the FBI academy in 1969 and taught a course in “Applied Criminology,” but was not really free to follow his main interest until J. Edgar Hoover, who regarded psychology unfavorably, died in 1972.²¹⁶ Soon after Hoover’s death, the hostage negotiation training operation was expanded to become the Behavioral Science Unit, and Teten more or less apprenticed himself to Brussel to learn his approach to what would come to be called “profiling.”²¹⁷ Teten and his partner in the behavioral sciences unit, Patrick Mullaney, became, in the words of John Douglas, “the first wave of modern behavioral science” in the FBI²¹⁸ (how much actual

logical Assessment of Crime: Profiling, 49 FBI LAW ENFORCEMENT BULLETIN 3, Mar. 1980, at 22, 23.

²¹⁴ The terminology is not wholly standardized. It is referred to in various places simply as “profiling,” see *id.*, and as “criminal profiling,” “offender profiling,” “psychological profiling,” “specific profile analysis,” and “criminal personality profiling,” OFFENDER PROFILING: THEORY, RESEARCH AND PRACTICE 2 (Janet L. Jackson & Debra A. Bekerian eds., 1997) [hereinafter OFFENDER PROFILING], and “Criminal Investigative Analysis Profiling” in others. See Robert R. Hazelwood, Robert K. Ressler, Roger L. Depue and John C. Douglas, *Criminal Investigative Analysis: An Overview*, in PRACTICAL ASPECTS OF RAPE INVESTIGATION: A MULTIDISCIPLINARY APPROACH (Robert R. Hazelwood & Ann Wolbert Burgess eds., 2d ed. 1995). These terms differentiate offender profiling from the kind of profiling used in interdiction contexts, such as drug courier profiling or terrorist profiling, which has received much recent publicity because of the use of race as a factor in constructing such profiles. In this article, we will use the term “offender profiling” and, occasionally, “criminal profiling.” This is what is meant when the term “profiling” is used without qualification.

²¹⁵ See DUDLEY D. SHOENFELD, THE CRIME AND THE CRIMINAL: A PSYCHIATRIC STUDY OF THE LINDBERGH CASE (1936). This palm may actually belong to Dr. Dudley Schoenfeld, whose predictions about the personality of the Lindbergh Baby kidnapper given to the authorities soon after the kidnapping fit the description of Bruno Richard Hauptmann surprisingly well. See *id.* at 41-56. See JAMES A. BRUSSEL, CASEBOOK OF A CRIME PSYCHIATRIST 7 (1968). Brussel became famous for his role in the 1956 case of George Metesky, the “Mad Bomber.” See *id.* This case has taken on somewhat mythic proportions in the profiling community, despite the fact that Dr. Brussel, the psychiatrist who was asked to analyze the personality of the unknown bomber before he was caught, was not as accurate as most people believe. See *id.* For instance, the detail repeated in virtually every book or article on the development of profiling ever written concerns Brussel’s uncanny accuracy in predicting what the Bomber would be wearing when arrested. As Brussel himself phrased it: “When you catch him—and I have no doubt you will—he’ll be wearing a double-breasted suit . . . And it will be buttoned.” BRUSSEL, *supra* at 46. In fact, when he was arrested, the “Mad Bomber” was wearing pajamas. *Id.* at 69.

²¹⁶ See JOHN DOUGLAS & MARK OLSHAKER, MINDHUNTER: INSIDE THE FBI’S ELITE SERIAL CRIME UNIT 94 (1995) [hereinafter DOUGLAS & OLSHAKER, MINDHUNTER]. See also BRENT E. TURVEY, CRIMINAL PROFILING: AN INTRODUCTION TO BEHAVIORAL EVIDENCE ANALYSIS 9 (1999).

²¹⁷ TURVEY, *supra* note 216, at 9. See H. PAUL JEFFERS, WHO KILLED PRECIOUS?: HOW FBI SPECIAL AGENTS COMBINE HIGH TECHNOLOGY AND PSYCHOLOGY TO IDENTIFY VIOLENT CRIMINALS 32 (1991).

²¹⁸ DOUGLAS & OLSHAKER, MINDHUNTER, *supra* note 216, at 82. See ROBERT K. RESSLER & TOM SHACHTMAN, WHOEVER FIGHTS MONSTERS 33-34 (1992) [hereinafter RESSLER &

science this might represent is another question).

To their credit, it occurred fairly quickly to the early members of the Behavioral Sciences Unit that there were plenty of details about the patterns of real criminals that no one really knew, and that these details might be useful in the process of profiling if they could be procured.²¹⁹ And in the organized getting hold of that information, the FBI agents in the behavioral sciences unit were uniquely situated. Essentially, if they wished they could go into any prison in the country and, once there, they had a higher likelihood of cooperation from both the authorities and any prisoner they wished to interview than anyone else.²²⁰

The members of the BSU started to capitalize on this opportunity more or less informally in early 1978.²²¹ The interviews were conducted as a side activity whenever other agency business, usually local police training sessions, took the involved agents near the prison where an interesting prisoner was held.²²² The target group was predominantly what today would be called “serial killers.”²²³ Exactly who was responsible for the initial idea of such interviews, and even who actually was present for the first such interviews, is today a topic of some fairly ugly controversy. However, it is clear that it was either Robert Ressler or John Douglas or both.²²⁴ Before long, these men were conducting

SHACHTMAN, WHOEVER FIGHTS].

²¹⁹ See RESSLER & SHACHTMAN, WHOEVER FIGHTS, *supra* note 218, at 31.

²²⁰ See ROBERT K. RESSLER, ANN W. BURGESS & JOHN E. DOUGLAS, SEXUAL HOMICIDE: PATTERNS AND MOTIVES xi-xii (1988) [hereinafter RESSLER ET AL., SEXUAL HOMICIDE]; DOUGLAS & OLSHAKER, MINDHUNTER, *supra* note 216, at 119, 127; RESSLER & SHACHTMAN, WHOEVER FIGHTS, *supra* note 218, at 35-36.

²²¹ See DOUGLAS & OLSHAKER, MINDHUNTER, *supra* note 216, at 99; RESSLER & SHACHTMAN, WHOEVER FIGHTS, *supra* note 218, at 36.

²²² This is Douglas' version. See DOUGLAS & OLSHAKER, MINDHUNTER, *supra* note 216, at 111. Ressler's account is somewhat different. See RESSLER & SHACHTMAN, WHOEVER FIGHTS, *supra* note 218, at 36-39.

²²³ Various people have claimed credit for originating the term “serial killer.” For instance, Robert Ressler claims to have coined the term in the mid-1970s. See RESSLER & SHACHTMAN, WHOEVER FIGHTS, *supra* note 218, at 29. In fact, it was coined by British crime writer John Brophy in a 1966 book. See JOHN BROPHY, THE MEANING OF MURDER (1966). Brophy used the term “serial murderer,” but the dust jacket to the American edition used the term “serial killer.” See *id.*

²²⁴ In his 1991 memoir, Ressler claims that the idea was entirely his and mentions Douglas only in passing as a protégé who was present at the *second* interview of serial killer Edward Kemper, which occurred after a number of other killers had been interviewed. See RESSLER & SHACHTMAN, WHOEVER FIGHTS, *supra* note 218, at 36-43. In his 1995 memoir, Douglas does not mention Ressler's published version of events but claims the entire project was *his* idea, and that he was at the first interview done, which he claims was with Edward Kemper. See *id.* at 105-06, 109-10. In his subsequent memoir, Ressler did not address this collision but, in his single reference to Douglas in the book, he somewhat gratuitously hammers Douglas for claiming to have “gone ‘face-to-face’” with John Wayne Gacy when “the records showed that Douglas had never interviewed Gacy in person.” ROBERT K. RESSLER & TOM SHACHTMAN, I HAVE LIVED IN THE MONSTER 92 (1997) [hereinafter RESSLER & SHACHTMAN, MONSTER]. In subsequent books, Douglas has mentioned Ressler only in passing. See JOHN DOUGLAS AND MARK OLSHAKER, THE ANATOMY OF MOTIVE: THE FBI'S LEGENDARY MINDHUNTER EXPLORES THE KEY TO UNDERSTANDING AND CATCHING VIOLENT CRIMINALS 18 (1999) [hereinafter DOUGLAS &

prison interviews with the most notorious serial killers then in prison.

These early interviews appear to have been fairly ad hoc and free-form affairs. Neither Ressler nor Douglas appear to have had any formal training in research design or the standards which might be applied to the products of their research to determine whether any generally well-grounded or useful information might emerge from such interviews.²²⁵

At some point these interviews moved from a kind of hobby to an accepted part of their professional activities. The behavioral sciences unit management was becoming supportive of something that might be called “research” into criminals and the patterns of their behavior. In the Spring of 1978, a mandate was issued to the staff of the FBI training division (which housed the BSU) “to originate original in depth research as one method of increasing the Bureau’s knowledge base in area’s relevant to the law enforcement community.” In response to this directive, various research efforts would be undertaken by BSU personnel over the next ten to fifteen years.²²⁶ One of the earliest was under-

OLSHAKER, MOTIVE]; JOHN DOUGLAS & MARK OLSHAKER, OBSESSION 16, 18, 93 (1998) [hereinafter DOUGLAS & OLSHAKER, OBSESSION]; JOHN DOUGLAS & MARK OLSHAKER, JOURNEY INTO DARKNESS: FOLLOW THE FBI’S PREMIER INVESTIGATIVE PROFILER AS HE PENETRATES THE MINDS AND MOTIVES OF THE MOST TERRIFYING SERIAL KILLERS 20 (1997) [hereinafter DOUGLAS & OLSHAKER, JOURNEY INTO DARKNESS]. It is no wonder that Stephen Michaud observed in 1998 that “[t]here are certain present and former BSU agents it is best not to invite to the same function.” MICHAUD & HAZELWOOD, EVIL, *supra* note 31, at 8.

²²⁵ See DOUGLAS & OLSHAKER, MINDHUNTER, *supra* note 216, at 117.

²²⁶ A complete review of BSU research done in the 1980s and later is beyond the scope of this article. However, it is useful to have an overview in order to understand why we have concentrated on Ressler, Douglas, and Hazelwood to the exclusion of others in the BSU who were involved in research and publication. In this article, we are concerned primarily with a blitz attack homicide with a sexual overtone. Hence, it seems appropriate to deal only with information that speaks to the accuracy of profiling or profiling-inspired expertise (e.g., “linkage analysis”) in that general type of case. With that in mind, let us examine the formal research attempts by BSU members that resulted in definable data sets.

The development of formal research data sets tended to be an activity undertaken by a team led by one or two FBI agents who specialized in that type of research. What we might call “weird motivation murder” (e.g., serial murder, lust murder, sadistic murder) was the specialty of Ressler and Douglas, although the only formal research data set they constructed appears to be the 36 serial killer data set resulting from their interview program after Burgess was on board and had overseen the construction of a defensible interview instrument and procedure. Serial rape and sexually motivated crime was the specialty of Roy Hazelwood (who shared sexually motivated murders with Ressler and Douglas), and he and Burgess, and later Janet Warren, oversaw the creation of a formal data set comprising interview data on 41 serial rapists, a later casefile-based data set on 30 sexually sadistic criminals, a later case file based data set of 20 sexually sadistic murderers (which shared 16 individuals with the previous data set), and a case file based data set of another 108 serial rapists (later rising to 112). Kenneth Lanning specialized in pedophilia and child pornography and developed offender research databases documenting these crimes. Special Agent David J. Icove developed, with others, a research database involving 1016 arsonists. Other agents had additional research projects, see DOUGLAS & OLSHAKER, MINDHUNTER, *supra* note 216, at 381, but none conducted any research relevant to this paper which we have discovered which was not also co-authored by Ressler, Douglas, or Hazelwood. In addition, these three are generally recognized as the main FBI players in the generation of whatever theoretical structures are claimed to lie behind the profiling process as performed and taught by the FBI. Finally, they,

taken by BSU member Robert (Roy) Hazelwood, who began looking at the phenomenon of death by autoerotic misadventure, which was apparently more common than generally thought and often confused with suicide or homicide by virtue of the commonness of some form of ligature strangulation as the immediate cause of death. Hazelwood had no formal training in research methodology, but he had been involved in a couple of team research projects while in the military, and he apparently had an appreciation of the advantages of methodological and statistical expertise in conducting research of the kind he envisioned. To that end, he enlisted an experienced researcher as a co-participant in his project, Dr. Ann Wolbert Burgess.²²⁷

Dr. Burgess is a psychiatric nurse who had first come to prominence in 1974 as a result of her research concerning sexual violence against women, and indeed, she and a co-author had coined the term “rape trauma syndrome.” Though the “autoerotic death” project was not exactly central to her own research interests, she must have seen the potential inherent in a relationship with Hazelwood and the FBI. At any rate, she joined with Hazelwood, who as a member of the FBI behavioral sciences unit, was in a unique position to gather relevant case files from across the country for examination. The result was a series of articles and a monograph detailing their analysis of what ultimately was 150 case files sent to them by local authorities as relevant to an examination of the results of dangerous autoerotic practices.²²⁸ Either Ressler, Douglas, or both had figured out that the serial killer interview research was going to need some methodologically-trained input to make it acceptable to the world outside of law enforcement, and Burgess was contacted. This was at the same time that Hazelwood and Douglass were working on the only piece they would ever publish together, a four and a half page article in the FBI Law Enforcement Bulletin entitled “the Lust Murderer.”

From this point onward, if one examines the published output of the profiling pioneers of the BSU, it quickly becomes apparent that the publications can usefully be divided into three main groups: research reports dealing with the analysis of research generated data (which are generally co-authored with a person of academic training in social science research and often published in reputable journals) professional

especially Douglas and Hazelwood, are the persons who offer to testify most often to “linkage analysis.” Hence, it seems appropriate to concentrate on their work.

²²⁷ At about the same time, Hazelwood enlisted prominent forensic pathologist Dr. Park Dietz as part of the research team dealing with autoerotic death. See MICHAUD & HAZELWOOD, *EVIL*, *supra* note 31, at 50. Dr. Dietz would continue to play a role in various BSU research and publication over the next decade and beyond. Dietz also may have played a role in bringing in Burgess.

²²⁸ The results of this research program are of little relevance to claims concerning perpetrator profiling or claims of an ability to determine accurately if two crimes were committed by the same perpetrator.

publications (which generally have no such co-authorship and usually appear in the FBI Law Enforcement Bulletin), and memoirs.²²⁹ The interesting thing about these three forms of writing is that the formal research often seems to have little discernable impact on the professional publications or the memoirs. For those, the pre-research publication "The Lust Murderer" sets the tone which was to continue through the years in the dominantly professional writings in spite of the different tone of the research report publications.

An examination of the *FBI Law Enforcement Bulletin* from its beginnings in 1938 until 1980 would indicate that no serious social scientist would have paid much attention to it as a source of sound research or conclusions prior to 1980,²³⁰ and the first publications stemming from the BSU research mandate would not have changed their minds. In "The Lust Murderer," Hazelwood and Douglas identify a subset of "sadistic murderers" which they label "lust murderers."²³¹ These are defined by the presence of "a mutilating attack or displacement of the breasts, rectum or genitals."²³² They then declare that the vast majority of perpetrators of such lust murders fall into two types defined by two variables: organization and sociability. The two types are "Organized Non-social" and "Disorganized Asocial."²³³ They then proceed in a kind of stream-of-consciousness collection of unqualified declarative sentences to describe what they claim to be the characteristics of these two types of murderers: "The organized nonsocial . . . lust murderer exhibits complete indifference to the interests and welfare of society and displays an irresponsible and self-centered attitude. While disliking

²²⁹ These classifications provide a feel for the type of literature generated, but they are not leak-proof. For instance, the two part article, *The Serial Rapist: His Characteristics and Victims*, that appeared in the FBI LAW ENFORCEMENT BULLETIN in January and February of 1989 and was co-authored by Hazelwood and Janet Warren (a Doctor of Social Work who was on the University of Virginia Medical Faculty and who played the "Ann Burgess" role in some of the later BSU research), reads much more like an academic article than does the usual FBI Law Enforcement Bulletin production. The non-memoir books and book chapters also present something of a problem. Some book chapters are co-authored versions of research previously published in academic journals. Some are non-co-authored pieces akin to the professional publications referred to in the text. In general, most non-memoir books are collections of readings in which the BSU agent was listed as a co-editor and may have contributed a chapter or two. This generalization, however, has a major exception. The taxonomic parts of the *Crime Classification Manual* are purported to be the product of the three authors and a multi-person advisory panel for each major division. Who contributed what is not entirely clear.

²³⁰ The Bulletin then consisted mostly of common sense articles on issues of policing and police administration, with an occasional conclusory summary article concerning research relevant to policing from academic sources.

²³¹ Robert R. Hazelwood & John E. Douglas, *The Lust Murderer*, 42 R.C.M.P. GAZETTE 10, 10 (1990).

²³² *Id.*

²³³ The psychological descriptors later were dropped from the categories. See RESSLER ET AL., SEXUAL HOMICIDE, *supra* note 220, at 121-22 (1988). However, at least one psychologist, Ronald M. Holmes, who is both an academic and a practicing profiler, believes that they should have been retained. See RONALD M. HOLMES & STEPHEN T. HOLMES, PROFILING VIOLENT CRIMES 47 (2d ed. 1996).

people in general, he does not avoid them. Instead, he is capable of displaying an amiable façade”²³⁴

The point here is not the empirical inaccuracy of the claims being made, but the total lack of any means to evaluate their accuracy. The authors admit as much: “The data presented here have not been quantified, but are based upon the author’s examination of case reports, interviews with investigative personnel, and careful review of the literature.”²³⁵ The absence of any means of checking the authors’ claims is in stark contrast to the confident, authoritative, rarely qualified, and global descriptions, and this is typical of much of the literature later generated.²³⁶

This is not to say that there was nothing of value in the article. While the psychodynamic descriptors are largely the product of categorical dramatic overgeneralization, the fundamental distinction between organized and disorganized perpetrators (later, perhaps inevitably, qualified by the intermediate category of “mixed”²³⁷) and the kinds of crime scenes they leave behind, appears to have proved fruitful for investigatory purposes in some cases, providing usable rule-of-thumb guidance regarding the general characteristics of a likely perpetrator in those cases which clearly fall into one or the other of the polar groups in terms that are intelligible to the average detective.²³⁸ While even these fundamental taxonomic categories and their correlation with general perpetrator types have never been the subject of any rigorous published validation²³⁹ (though they are the stuff of numerous memoir anecdotes), this is perhaps less important for investigatory purposes than if they formed the basis for proffered courtroom testimony. These categories remain prominent in the teaching and practice of perpetrator profiling today, even in cases not involving “lust murder.” “The Lust Murderer” remains the single most reprinted article from the *FBI Law Enforcement Bulletin*.

²³⁴ Hazelwood & Douglas, *supra* note 231, at [MISSING]

²³⁵ Hazelwood & Douglas, *supra* note 231, at 1.

²³⁶ In the best review of the literature on the subject published to date, Fox and Levin observe: “The research literature, still in its infancy, is more speculative than definitive, based primarily on anecdotal evidence rather than hard data” James Alan Fox & Jack Levin, *Multiple Homicide: Patterns of Serial and Mass Murder*, 23 CRIME & JUST. 407 (1998).

²³⁷ “Mixed” appears as a category by 1985. See *Classifying Sexual Homicide Crime Scenes: Interrater Reliability*, in 54 FBI LAW ENFORCEMENT BULLETIN 8, Aug. 1985, at 13, 16 [hereinafter *Interrater Reliability*].

²³⁸ In 1998, Michaud, with Hazelwood, quoted Vernon J. Geberth, author of the standard PRACTICAL HOMICIDE INVESTIGATION: TACTICS PROCEDURES AND FORENSIC TECHNIQUES, as saying, “[t]he disorganized and organized classification of crimes was fantastic, a brainstorm. For a police officer to be able to define and describe behavior without using clinical terms was just fantastic.” MICHAUD & HAZELWOOD, *EVIL*, *supra* note 31, at 8.

²³⁹ However, there exists one very important FBI reliability study, reported in *Interrater Reliability*, *supra* note 237. The less than impressive levels of interrater reliability shown by this study have serious implications for validity, as validity is, in part, a function of reliability. See discussion *infra* notes 285-305.

Hazelwood's next research project after the autoerotic death project involved serial rapists, defined as offenders who had committed more than 10 rapes. This was the first project that Burgess actually helped design from the outset. It involved forty-one serial rapists from across the country who were responsible for over a thousand rapes and sexual assaults. All available casefiles relating to the forty-one were examined, and each of the forty-one was interviewed at length. At the conclusion of the interview, the interviewer filled out a 70-page protocol which served as the basis for further analysis, statistical and otherwise. This data set was then mined over the course of some years to generate a number of articles and book chapters by Hazelwood and Burgess, and later Hazelwood and Janet Warren.²⁴⁰ The tone of these publications is in general appropriately circumspect in its treatment of the data. Compare this to the general tone of Hazelwood writing alone in the *FBI Law Enforcement Bulletin* in 1983, prior to the completion of the project or the analysis of the data. The name of the article is *The Behavior-oriented Interview of Rape Victims: The Key to Profiling*.²⁴¹ There, Hazelwood starts with what has become a standard device: a case description, followed by an FBI profile given before apprehension of the perpetrator, followed by a claim of amazing accuracy.²⁴² In the particular case the victim could provide no visual description of the perpetrator because he had awakened the victim in her bed and placed a pillowcase over her head. We join the article at that point:

Needing additional information in order to complete a profile, the requesting agency was sent a set of questions specifically designed to elicit information from the victim concerning the rapist's behavior during the assault. The victim was reinterviewed, using the questions as a guide. As a result, a 9-page typewritten statement was obtained. Based on the new statement, a profile was prepared with opinion as to the offender's age, race, marital status, occupational level, arrest history, socioeconomic background, type and proximity of residence to victim, military history, approximate age and style of automobile, as well as certain personality characteristics. The rapist was subsequently arrested and confessed to a series of rapes. When the profile was compared to the offender, only the marital status was found to be incorrect.²⁴³

What follows thereafter in the article is an explanation to those

²⁴⁰ See, e.g., RESSLER ET AL., SEXUAL HOMICIDE, *supra* note 220; Anne W. Burgess et al., *Sexual Homicide: A Motivational Model*, 1 J. INTERPERSONAL VIOLENCE 251 (1986); Robert K. Ressler et al., Anne W. Burgess, John E. Douglas & Roger I. Depue, *Criminal Profiling: Research on Profiling*, in RAPE AND SEXUAL ASSAULT: A RESEARCH HANDBOOK (Anne W. Burgess ed., 1985) [hereinafter RAPE AND SEXUAL ASSAULT].

²⁴¹ Robert R. Hazelwood, *The Behavior-oriented Interview of Rape Victims: The Key to Profiling*, in 52 FBI LAW ENFORCEMENT BULLETIN 9, Sept. 1983, at 8 [hereinafter Hazelwood, *Key to Profiling*].

²⁴² See *id.* A version of this device was used to open the very first FBI publication on the profiling process. See Ault & Reese, *supra* note 213, at 22-23.

²⁴³ Hazelwood, *Key to Profiling*, *supra* note 241, at 8.

whose job it is to interview rape victims explaining what details of the victim's story will lay the foundation for such an astounding performance.

It is not as if anything that is said in the rest of the article is counterintuitive or obviously wrong. It is just that, once again, everything is stated authoritatively and positively, generally with little reference to any evidence supporting the assertions.²⁴⁴ A typical example among potentially dozens of similar examples is the following, discussing rapists who take items from the rape victim or her premises and later return them: "Some do so to maintain power over the victim by intimidation, while others wish to convince the victim they meant no harm to her life and wish to convince themselves that they are not bad persons."²⁴⁵

It seems plausible that such psychologizing might lead one to a feeling that one understands the perpetrator better,²⁴⁶ and even sometimes to specific predictions about the type of person involved in a rape that turned out to be right. However, there is nothing mystical involved in such predictions. Such predictive exercises are nothing more or less than playing the odds, and unless the predictions are linked to the reality in such a way that the correlation is virtually 100 percent, errors will occur which, over the run of cases, mirror the underlying probabilities. For example, in the book which, among other things, serves as the research report based on the data generated in the study of thirty-six "sexual murderers" which grew out of the very first interview research by the BSU,²⁴⁷ the highest correlation in the data between an offense characteristic and an offender characteristic (other than sex, they were all male) was for previous sexual assault conviction (94 percent).²⁴⁸ So a little over one in twenty profiles should get this wrong. And most common items that profilers opine about have much higher rates of exception. Of these murders, 80 percent had problems with jobs, but 20 percent did not.²⁴⁹ Around 60 percent of those who had been in the military had discipline related discharges, but 40 percent had honorable or general discharges.²⁵⁰ Assuming reasonable independence (and nothing in the report indicates otherwise), less than half of profiles dealing with all three variables should get all three items right. And even as-

²⁴⁴ There are a few uses of previously conducted studies. For example, A. N. Groh & Ann W. Burgess, *Sexual Dysfunction During Rape*, 297 NEW ENG. J. MED. 764 (1977), is cited on page twelve regarding the nature and prevalence of sexual dysfunction among rapists, and Holmstrom & Burgess, *Rapist's Talk: Linguistic Strategies to Control the Victim*, 9 ARCHIVES OF SEXUAL BEHAVIOR 437 (1980), is cited on page thirteen to highlight the importance of the exact content of the rapists' communications with the victim.

²⁴⁵ See Hazelwood, *supra* note 241.

²⁴⁶ The distinction between this feeling of explanatory understanding and the accuracy of predictions becomes important later in this piece. See *infra* note 280-284.

²⁴⁷ See RESSLER ET AL., SEXUAL HOMICIDE, *supra* note 220.

²⁴⁸ See *id.* at 66.

²⁴⁹ See *id.* at 31.

²⁵⁰ See *id.*

suming some robust dependence, no performance should exceed 60 percent.

The normal assumption would be that unless the profiler is simply lucky, to the degree that profilers' assumptions do not mirror the underlying probabilities reflected in objectively assembled data, their predictions will fall short of optimum performance. Like a gambler in the habit of drawing to inside straights, there will be even more errors than the true probabilities would generate. But it is in the interest of the investigators involved to appear to be better than the probabilities by some sort of magical or mystical process. Lest the reader think that we are being unduly cynical, consider the following claim to that magical mantel which begins Chapter 9, "Profiling from Crime Scene Analysis," in the very same volume in which the study data are reported:

"You wanted to mock yourself at me! . . . You do not know your Hercule Poirot." He thrust out his chest and twirled his moustache.

I looked at him and grinned . . . "All right then," I said. "Give us the answer to the problem—if you know it."

"But of course I know it"

Hardcastle stared at him incredulously. . . "Excuse me, Monsieur Poirot, you claim that you know who killed three people. And why? . . . All you mean is that you have a hunch."

"I will not quarrel with you over a word . . . Come now, Inspector, I know—really know . . . I perceive you are still sceptic. But first let me say this: To be sure means that when the right solution is reached, everything falls into place. You perceive that in no other way could things have happened."

The ability of Hercule Poirot to solve a crime by describing the perpetrator is shared by the expert investigative profiler. Evidence speaks its own language of patterns and sequences that can reveal the offender's behavioral characteristics. Like M. Poirot, the profiler can say, "I know who he must be."²⁵¹

How is one to account for this hubristic claim of perfection. First, even in the purest of sciences one of the ironies of the process of scientific advance is that it is a group process that depends in part on the "unscientific" commitment of those generating new theories to the validity of their hypotheses far in excess of what would be justified by extant data.²⁵² One might think of such persons as the intellectual entrepreneurs of science. The ideas of most don't pan out. In the end, and in a Darwinian process, they and their ideas fall by the wayside. Occasionally, the ideas of such enthusiasts do pan out, and they receive Nobel

²⁵¹ *Id.* at 135 (quoting and commenting on AGATHA CHRISTIE, *THE CLOCKS* 227-28 (1963)).

²⁵² See D. Michael Risinger et al., *Brave New "Post-Daubert World": A Reply to Professor Moenssens*, 29 SETON HALL L. REV. 405, 438 (1998) (discussing this phenomenon in the practice of science as a social enterprise).

Prizes for their pains. Our point here is that it is not only law enforcement officers who run the risk of developing belief in their own ideas and powers beyond what is warranted by the data, and in some contexts that is a benign condition. However, in other contexts, such as when the life or liberty is dependent on the accuracy of a process, it can be destructive.

Second, people involved in profiling have a large personal and professional stake in fostering the mystique of their own accuracy independent of its truth. Aside from the obvious awe in which all shamans are held by those who believe in them, and both the professional and popular iconic status it can generate, there are investigatory uses for processes with no objective validity, as the famous blue chicken case long ago proved.²⁵³

Finally, it must be pointed out that the official professional commitment of law enforcement officers to strict truth-telling is less than that of scientists. This is not intended to be inflammatory, and a moment's reflection will show that it is a virtual inevitability. This is because, in a significant percentage of things that law enforcement officers do professionally, lying is a necessary professional activity and skill and is not only tolerated but encouraged and respected. All undercover investigations or investigations relying on informants are based on lying in the service of an assumed greater good. It would hardly be surprising to find that one of the soul risks of operating as a law enforcement officer, especially one in a primarily investigatory agency, is difficulty in controlling the line between acceptable lying and exaggeration in effective investigation and unacceptable lying and exaggeration in both investigative and other capacities.²⁵⁴

Or perhaps we are being too harsh. While the literature of profil-

²⁵³ See *State v. Goldstein*, reported in New York legal newspaper and recounted in RICHARDSON ON EVIDENCE § 167, at 135, 136 (Jerome Prince ed., 10th ed. 1973). In *Goldstein*, a larceny was committed and circumstances indicated that the guilty party was one of a small group of individuals. The investigator on the case told the group that he owned a specially bred chicken which would squawk when touched by a person guilty of a crime. He then sent each person into a darkened room with instructions to touch the chicken, which was fastened to a table in the middle of the room. Unbeknownst to the group, he had placed a blue dye on the chicken. Only one man emerged without dye on his hands. *Id.*

²⁵⁴ Those engaged in law enforcement sometimes appear as pragmatic with the truth as litigators are in regard to the kind of expertise that they will proffer. Recall Ressler's charge in *I HAVE LIVED IN THE MONSTER* against Douglas that he had untruthfully claimed to have interviewed John Wayne Gacy face to face when he had not. See RESSLER & SHACHTMAN, *MONSTER*, *supra* note 224. Consider further that, in the same memoir, Ressler rather proudly recounted filing false and backdated reports in order to protect his research program from bureaucratic higher-ups in the FBI, and Hazelwood similarly recounted in his first book that, during his service as an officer in the military police, he rid himself of a problem prisoner he could not otherwise get transferred away from him by framing him for a petty offense. See MICHAUD & HAZELWOOD, *EVIL*, *supra* note 31, at 39-40. Though these episodes are old, their proud recitation is recent. The point is not that these actions are not understandable, but that they show a certain pragmatic willingness to sacrifice truth to an assumed personally-held higher goal, which is dangerous in an expert witness.

ing does contain such claims to virtual infallibility as the one set out above, it also contains concessions of the possibility of error (though indeed it must be said that it often leaves the impression that error is uncommon).²⁵⁵ What circumstances lie behind its claims to general accuracy, and could the claims be true in fact?

The universe of offender profilers seems to contain two groups representing two schools of thought and two approaches to process: "statistical profilers" and "clinical profilers."²⁵⁶ Statistical profilers stick closely to correlations revealed by formally gathered and at least semi-publicly²⁵⁷ available data sets, and generally view as desirable efforts to develop computerized profiling systems. Clinical profilers do not rely very specifically on formal statistical correlations (though they may be knowledgeable about them), but instead rely heavily on "experience" and "intuition." While there have been efforts directed toward computerized profiling in some areas at the FBI,²⁵⁸ the main orthodox approach of the BSU deriving from Ressler, Douglas and Hazelwood is the clinical approach.²⁵⁹ In fact, Douglas has gone so far as to assert that the ability to accurately profile is personality dependent and only partly teachable,²⁶⁰ and has even embraced the possibility that it involves psychic powers.²⁶¹ In the right hands, the claim seems to be, that

²⁵⁵ Compare DOUGLAS & OLSHAKER, *MINDHUNTER*, *supra* note 216, at 154 (possibility of partial error conceded), with *A Word of Caution about Profiling*, in *RAPE AND SEXUAL ASSAULT*, *supra* note 240, at 124-25 (warning against over-reliance on profiles to the exclusion of other investigatory tools without mentioning possibility of error).

²⁵⁶ G.H. Gudjonsson & G. Copson, *The Role of the Expert in Criminal Investigation*, in *OFFENDER PROFILING*, *supra* note 214, at 69.

²⁵⁷ Some data analyses, such as the ones cited earlier in regard to the FBI formal research efforts of the 1980s, have been published. However, the raw data for the FBI projects, as well as raw data gathered by other police agencies, is generally available only with the permission of the police agency that holds it, which, at least in the case of the FBI, does not appear to be given often, and then only under conditions which would deter serious objective researchers. For instance, in the early 1990s, Dr. William C. Thompson, a professor at the University of California at Irvine, applied to the FBI for access to their DNA research data and was told that it would be granted without a court order only if he agreed that any publication resulting from his examination of the data would be co-authored by a member of the FBI. See Personal Telephone Communication from Dr. William C. Thompson (Jan. 2002). In another instance, Dr. Michael Saks, on behalf of himself and one of us (Risinger), asked the FBI for the raw data involved in Dr. Moishe Kam's research on handwriting identification, pursuant to the clear directive of federal regulations. Initially, he was promised the data but, later, he was told it would be released only after a successful Freedom of Information Act suit was brought. See *Handwriting Identification*, in FAIGMAN ET AL., *supra* note 194, § 28-2.3.6 n.161; Telephonic Communication from Michael J. Saks (Nov. 2001). The reluctance to share data (which seems to be related to a general tendency toward secrecy in law enforcement and bureaucratic culture in general) is one reason to doubt that much of sustained scientific value can come out of research embedded in law enforcement agencies, as valuable as some of the published results of the FBI efforts of the 1980s may have been.

²⁵⁸ See David J. Icove, *Automated Crime Profiling*, 85 FBI LAW ENFORCEMENT BULLETIN 12, Dec. 1986, at 27.

²⁵⁹ See DOUGLAS & OLSHAKER, *MINDHUNTER*, *supra* note 216, at 169; DOUGLAS & OLSHAKER, *JOURNEY INTO DARKNESS*, *supra* note 224, at 20.

²⁶⁰ See DOUGLAS & OLSHAKER, *JOURNEY INTO DARKNESS*, *supra* note 224, at 19-22.

²⁶¹ "If there is a psychic component to this, I won't run away from it" DOUGLAS &

a good clinical profiler can actually beat the formal odds, and perform better than the data in databases because their experience and “intuition”²⁶² allow them to sense subsets applicable to the individual case with different and higher probabilities than those revealed in the formal data, and accurately assign probabilities better than those revealed in the formal data. They can, in a sense, beat the formally known odds by intuiting accurately when to draw to an apparent inside straight.²⁶³

The problems of dealing with such claims of “experience based,” “clinical” expertise are a continuing significant topic both in the courts and in legal scholarship.²⁶⁴ What underlies such claims, from diagnostics to dowsing, is an assertion that, after appropriate training or experience, all or some humans can internalize a set of only partly conscious algorithms or responses which allow them to accurately convert data of a certain type into accurate judgments about some other non-obvious fact without full conscious access to all of the processes that go into the result.²⁶⁵ Such claimed expertises frankly use the subjective responses of human beings as an integral part of their methodology, despite the well-known vulnerability of such subjective processes to gross distortion under many conditions.²⁶⁶ This use of human processing is justified, it is claimed, because humans are better than any available non-subjective technology at accurately processing the non-quantifiable complexity of the variables presented.²⁶⁷ Such claims are undoubtedly sometimes true. Judge McKenna’s famous harbor pilots²⁶⁸ do learn to arrive at the right dock, and perhaps at least a few people who claim to be able to recognize wine by vineyard upon tasting actually can. However, such claims are also undoubtedly sometimes false, even though the practitioner sincerely believes in those abilities, as some hundreds of studies of the predictions of astrology ought by now to have established.²⁶⁹ The problem for the law is how to distinguish which claims

OLSHAKER, MINDHUNTER, *supra* note 216, at 147.

²⁶² “Intuition” might better be called “tacit knowledge.” See Debra A. Bekerian & Janet L. Jackson, *Critical Issues in Offender Profiling*, in OFFENDER PROFILING, *supra* note 214, at 211-12 (discussing tacit knowledge).

²⁶³ See John Douglas’ explanation of this effect, as given to Janet Reno, recounted in DOUGLAS & OLSHAKER, JOURNEY INTO DARKNESS, *supra* note 224, at 18.

²⁶⁴ For a full exploration of this topic, see FAIGMAN ET AL., *supra* note 194, § 1-3.5 and authorities cited therein.

²⁶⁵ See D. Michael Risinger, *Preliminary Thoughts on a Functional Taxonomy of Expertise for the Post-Kumho World*, 31 SETON HALL L. REV. 508, 522-23 (2001) [hereinafter Risinger, *Functional Taxonomy*].

²⁶⁶ See Risinger et al., *Daubert/Kumho Implications*, *supra* note 194 (cataloguing such circumstances).

²⁶⁷ See this explicit claim in regard to offender profiling in DOUGLAS & OLSHAKER, JOURNEY INTO DARKNESS, *supra* note 224, at 32.

²⁶⁸ An example of reliable experiential expertise given in *United States v. Starzecpyzel*, 880 F. Supp. 1027, 1043 (S.D.N.Y. 1995), discussed in D. Michael Risinger & Michael J. Saks, *Science and Nonscience in the Courts: Daubert meets Handwriting Identification Expertise*, 82 IOWA L. REV. 21, 30-31 (1996) [hereinafter Risinger & Saks, *Science and Nonscience*].

²⁶⁹ See generally H. EYSENCK & D. NIAS, *ASTROLOGY, SCIENCE OR SUPERSTITION* (1982)

are which.

One thing that does not provide a strong basis for a conclusion of accuracy, as the Supreme Court has recognized, is the self belief of the expert by itself.²⁷⁰ So what is available beyond this? For claimed expertise which, unlike harbor piloting, does not yield clear unambiguous and public evidence of success or failure every time it is undertaken, some form of objective testing would seem to be a minimum requirement.²⁷¹ For purposes of evidentiary admissibility, at any rate (as opposed to use of such claimed expertise for investigatory purposes²⁷²), such a testing requirement would appear most consistent with the Supreme Court's emphasis on testability, testing and error rates in *Daubert v. Merrell Dow Pharmaceuticals Inc.*²⁷³

As applied to profilers, the point of all this is simple. As we have already pointed out, the research data available to profilers, either from their own research or that of others, neither specifically validate the assumptions of the process nor provide the information from which to construct the profiles they produce. Thus, the profiles themselves must be generated in great part by a subjective experience-based process not unlike others, such as handwriting identification, which we have examined at great length and depth in the past.²⁷⁴ Such "black box" processes may or may not be accurate, but their accuracy is not guaranteed by the self-belief of those involved in the process. Some form of external validation of the products of the process is necessary.²⁷⁵ And one should always bear in mind that accuracy alone is not the fundamental issue if testimony based on such claimed expertise is offered in court, but some marginal advantage of performance over the ordinary juror.

III. PROFILING RELIABILITY AND VALIDITY—THE EMPIRICAL RECORD

In the case of profiling, two such sources of validation suggest themselves immediately: proficiency testing and statistically valid accuracy review of the universe of profiles actually performed.²⁷⁶ Thus, the

(reviewing studies).

²⁷⁰ "[N]othing in either *Daubert* or the Federal Rules of Evidence requires a district court to admit opinion evidence that is connected to existing data only by the *ipse dixit* of the expert." *Kumho Tire Co., Ltd. v. Carmichael*, 526 U.S. 137, 157 (1999) (quoting *General Electric Co. v. Joiner*, 522 U.S. 136, 146 (1997)).

²⁷¹ See Risinger, *Functional Taxonomy*, *supra* note 265, at 535-36.

²⁷² For a discussion of the different warrants applying to testifying experts for considering information compared with detectives using the same information for "investigatory purposes," see Risinger et al., *Daubert/Kumho Implications*, *supra* note 194, at 27-30.

²⁷³ 509 U.S. 579 (1993).

²⁷⁴ See generally D. Michael Risinger, *Handwriting Identification*, in FAIGMAN ET AL., *supra* note 194, § 28-2.3.6 n.175.

²⁷⁵ See Risinger & Saks, *Science and Nonscience*, *supra* note 268, at 40-41.

²⁷⁶ A third source of information, so called "consumer satisfaction surveys," is sometimes suggested as another source of such information. Finally, one recent study compared witness de-

profilers themselves, especially the leading profilers who also claim to be scientific researchers, control access to the validation data and processes. This makes it even more suggestive to observe that no such validation effort has ever been undertaken by the FBI or the profiling community at large, at least so far as we know.²⁷⁷

Only two studies, which have any bearing at all on the issue have been done under FBI auspices: a 1981 “consumer satisfaction survey” directed to local police officers who had solicited and received profiling services from the FBI in regard to difficult cases,²⁷⁸ and a study of “interrater reliability” in the profiling process.²⁷⁹

The exact relationship of consumer satisfaction surveys to accuracy is itself not clear, and interpreting the FBI survey is not made easier by the fact that it has never been published. The only way to glean its likely result is from bits and pieces of description given by authors who have managed to obtain copies. From these sources, it appears that the results were consistent with the results of a similar survey recently undertaken in Great Britain: profiles do not often lead to the identification of the unknown perpetrator,²⁸⁰ but local police like them because

scriptions of crime scene actions by rapists with the characteristics of rapists actually convicted to determine if similar described actions correlated with any rapist characteristics in an attempt to explore whether profiling claims are even plausible. The author’s concluded that that no significant correlations existed. See Andreas Mokros & Laurence Alison, *Is Offender Profiling Possible? Testing the Predicted Homology of Crime Scene Actions and Background Characteristics in a Sample of Rapists*, 7 LEGAL & CRIMINOLOGICAL PSYCHOL. 25 (2002).

²⁷⁷ Recent events in regard to the recent “fingerprint validity” case, *United States v. Llera-Plaza*, 188 F. Supp. 2d. 549 (E.D. Pa. 2002), require the “so far as we know” to be inserted. The FBI had done some proficiency tests of fingerprint identification but, apparently, since they were not perfect, kept them secret until they lost a *Daubert* challenge. Only then did the FBI reveal the existence of the tests and its results in an attempt to obtain reconsideration of Judge Pollak’s decision. In the event, they were successful. See *id.* As a side effect, the results of the tests are now available for public evaluation.

²⁷⁸ See Douglas, J., Evaluation of the psychological profiling program: Institutional Research and Development Unit, FBI Academy, in G.H. Gudjonsson & G. Copson, *The Role of the Expert in Criminal Investigation*, in OFFENDER PROFILING, *supra* note 256, at 61.

²⁷⁹ See *Interrater Reliability*, *supra* note 237, at 13. A piece of anonymous research is unusual, but this piece appears in an issue of the FBI Law Enforcement Bulletin which was, according to the message of FBI Director William Webster located on the first page of the Bulletin, intended as an update on the latest developments at the National Center for the Analysis of Violent Crime (the “NCAVC”). The entire issue is made up of unsigned articles. The responsibility is given, on the title page, to virtually every member of the center under the general editorship of Robert K. Ressler. See Fox & Levin, *supra* note 236. Perhaps based on private knowledge, authorship may be attributed to Ressler and Burgess. See *id.* at 428.

²⁸⁰ The profile was credited with bringing about apprehension in only 2.7 percent of the cases in the British study. See Gudjonsson & Copson, *supra* note 256, at 75. The FBI study often is cited as claiming a 17 percent apprehension rate, but that was 17 percent of solved cases, which constituted less than half the sample (i.e., 88 cases out of 192). See HOLMES & HOLMES, *supra* note 233, at 44. Accordingly, their real rate is reduced to 7.8 percent (i.e., 15 cases out of 192). However, the wording in the FBI survey was apparently broader than in the British case (i.e., “helped in the identification” as opposed to “led to the identification”). Compare Gudjonsson & Copson, *supra* note 256, at 74 with HOLMES & HOLMES, *supra* note 233, at 44. For an example of a case in which the profile apparently led to the identification of the culprit, see DOUGLAS & OLSHAKER, *MINDHUNTER*, *supra* note 216, at 164-66 (describing identification of Carmine

they feel they help them better understand the person who committed the crime.²⁸¹ Note that neither of these circumstances settles the issue of the investigatory value of profiles, and neither is linked very directly to accuracy. Regardless of profile accuracy, a high percentage of these cases are never solved. And accurate profiles may still not play a role in catching a perpetrator even in a case which is solved, in the common situation where the real break in the case turns out to be the normal kind of circumstance that would lead to apprehension, profile or no: an acquaintance informs on the perpetrator, or other more specific investigatory or forensic processes identify him.²⁸² Thus, the lack of impact for profiles does not establish their inaccuracy.

Similarly, high satisfaction among users does not establish accuracy. The feelings of understanding commonly cited as reasons for satisfaction appear to result from the psychodynamic aspects of the profile (which in the FBI's case emphasize childhood abuse, fantasy, and ritual).²⁸³ Whatever the validity of this model, it may lead to feelings of understanding (have high apparent explanatory power), without giving rise to accurate predictions, as anyone watching economists and stock market pundits explain today very precisely and satisfyingly why they were wrong yesterday can attest. Indeed, this has been a powerful criticism of psychodynamic models ever since Freud. Finally, satisfaction might simply be the product of the increase in energy, or focus which can result from sharing frustrations with, or being taken seriously by, a reputed outside authority, a variant of the well known Hawthorne effect.²⁸⁴

The results of the single FBI reliability study are not reassuring to the claims of programming. "Reliability," as a technical term of art, refers to the extent to which two runs of a given test can be expected to come to the same result. It is thus synonymous with "consistency."²⁸⁵ It is often contrasted with "validity," which refers to the actual accuracy of a test. An unreliable test cannot have high validity, but a reliable test can have low validity because it is not testing what it is assumed to be

Calabro).

²⁸¹ See Gudjonsson & Copson, *supra* note 256, at 74.

²⁸² For example, David Carpenter (the "Trailside Killer") was caught as a result of surviving victim and witness details about his age, appearance, and car. See DOUGLAS & OLSHAKER, *MINDHUNTER*, *supra* note 216, at 155-56. Also, John Prante was caught on account of an acquaintance tip followed by dental evidence. See *id.* at 276-78.

²⁸³ The FBI's approach to its psychological theories, particularly the version of trait theory manifested in FBI writings, can be criticized as primitive and outmoded. For instance, Julian C.W. Boon identified 23 theories of personality, each which he claims may be useful in profiling under some circumstances. See J.C.W. Boon, *The Contribution of Personality Theories to Psychological Profiling*, in OFFENDER PROFILING, *supra* note 214, at 43.

²⁸⁴ See Risinger et al., *Daubert/Kumho Implications*, *supra* note 194, at 20.

²⁸⁵ See D. Michel Risinger et al., *Exorcism of Ignorance as a Proxy for Rational Knowledge: The Lessons of Handwriting Identification "Expertise,"* 137 U. PA. L. REV. 731, 737 n.22 (1989), for a discussion of the differences and interconnection between reliability and validity.

testing.²⁸⁶ The reader will recall that one of the foundational and threshold classifications of the FBI's profiling approach to rape and sexually related homicide is the distinction between organized and disorganized offenders, and the crime scenes that reflect them.²⁸⁷ One would therefore expect profilers, or at least the crème de la crème profilers in the BSU, to all come to the same conclusion when looking at a murder scene and asking "is it sexual murder" and "is it organized, disorganized or mixed."²⁸⁸ In 1985, somebody at the FBI decided to test whether this was true.²⁸⁹

Six BSU profilers took part in the study.²⁹⁰ Sixty-four cases were selected.²⁹¹ For each case, there was already one BSU agent who had been involved in profiling the case originally.²⁹² That agent made a presentation to the other five agents explaining the details of the crime scene with slides,²⁹³ but not intentionally (to the extent humanly possible) communicating his own conclusions about the sexual nature of the homicide or whether the scene showed the perpetrator to have been organized, disorganized or mixed, or did not contain enough information to draw conclusions.²⁹⁴ The other five then made their classifications.²⁹⁵ None of the six took part in classifying every case.²⁹⁶ Participation ranged from 62 cases to only 27 cases.²⁹⁷ One agent made 89 percent of the presentations, so his performance was omitted, since the main object of the test was not to determine interrater agreement or disagreement directly, but agreement with the presenter.²⁹⁸

Agreement scores on "sexual murder" ranged from a high of 93 percent to a low of 77 percent (actual scores in percent: 93, 88, 82, 77, 77),²⁹⁹ but the person with 93 percent agreement with the presenter only did 27 cases.³⁰⁰ 81.4 percent of all classifications made agreed with the presenter.³⁰¹ What this means is that, under test conditions procedurally subject to non-blind suggestion effects (intended or unintended) which

²⁸⁶ See *id.*

²⁸⁷ See *supra* discussion accompanying notes 230-234.

²⁸⁸ See *supra* note 237 (noting that the "mixed" category was added sometime in the 1980s).

²⁸⁹ See *Interrater Reliability*, *supra* note 237, at 103. As noted above, the exact persons who conceived and ran the test are not identified.

²⁹⁰ See *id.* at 13.

²⁹¹ See *id.*

²⁹² See *id.*

²⁹³ See *id.*

²⁹⁴ See *id.* at 13-14.

²⁹⁵ See *id.*

²⁹⁶ See *id.* at 17, Figure 4.

²⁹⁷ See *id.* at 17.

²⁹⁸ See *id.* One suspects that the test might better be called "how much do you agree with John Douglas" since he was the chief profiler doing the majority of cases at the time. See MICHAUD & HAZELWOOD, *EVIL*, *supra* note 31, at 98. However, the "chief presenter" also may have been Hazelwood. In one year, Douglas did 80 profiles and Hazelwood 60. See *id.*

²⁹⁹ See *id.* at 17, Figure 3.

³⁰⁰ See *id.*

³⁰¹ See *id.* at 17.

would be expected to skew the results to raise agreement, the net overall agreement was only 81.4 percent. For most serious tasks, 81.4 percent is not good reliability³⁰² (think of a lab test for cancer where two runs of the test only agree on the binary choice “cancer/no cancer” only 81.4 percent of the time).

The performance for agreement in regard to organization and disorganization was even worse. The overall agreement rate was 74 percent. The agreement scores ranged from a high of 85 percent to a low of 52 percent (actual scores in percent: 85, 77, 76, 70, 52). And while the highest and the lowest scores were posted by the same subjects as on the sexual crime classification test, the person who was second on the first test came in fourth on the last test.

While the BSU attempted to put a happy face on the results, blaming bad performance on the inexperience of some test subjects, and asserting that they “demonstrated that there is reliability in the classification of crime types and scenes by the BSU”³⁰³ the numbers have showed otherwise.³⁰⁴ These reliability data can give no comfort to those claiming the high accuracy of the FBI profiling practice.³⁰⁵

Nor can they look to the results of those studies that have sought to measure both profiler accuracy and profiler marginal advantage over other groups, such as detectives, psychologists, psychics, and average people. Such comparative proficiency studies create a test case, usually from a closed case file, and administer it to profilers and various control groups to gauge the performance of the profilers absolutely and against the other groups. There have been three such studies, two published and one semi-published. The two published studies are Pinnizzoto and Finkel³⁰⁶ in 1990, and Kocsis, Irwin, Hayes and Nunn³⁰⁷ in 2000. The

³⁰² This is only an indirect reliability score. A full reliability would look at the correspondences among all judgments given for each decision category. Agreement rates between individual scorers were figured, and a range was given (77 to 100 percent for sex relation), though this is problematic, as no two scorers scored all of the tests. Presumably, this is the agreement rate on common tests, which theoretically could have been as low as $n=3$, or even $n=0$ (since two raters scored less than half the tests). No full reliability score for each judgment category is given; on the raw data it obviously could be figured, but it cannot be derived from the summary data published. However, these scores must be assumed to be at least as low as the indirect measure of agreement with the presenter, or they would likely have been given explicitly.

³⁰³ *Interrater Reliability*, *supra* note 237, at 17.

³⁰⁴ This has not been entirely lost on those outside of the FBI. For example, Fox and Levin point out that since the sample of cases was weighted heavily toward cases displaying “organized” behavior, the 74 percent aggregate agreement rate was “not much better than a fixed ‘organized’ response.” Fox & Levin, *supra* note 236, at 428.

³⁰⁵ Even though the study is from 1985, there is no reason to believe that reliability has improved since then. There have been no real advances in theory and no claimed advances in practice.

³⁰⁶ See Anthony J. Pinnizzotto & Norman J. Finkel, *Criminal Personality Profiling: An Outcome and Process Study*, 14 LAW & HUM. BEHAV. 215 (1990).

³⁰⁷ See Richard N. Kocsis, Harvey J. Irwin, Andrew F. Hayes & Ronald Nunn, *Expertise in Psychological Profiling: A Comparative Assessment*, 15 J. INTERPERSONAL VIOLENCE 311 (2000).

semi-published study (because the results have been previewed in a book chapter but the actual study has not yet been published) is by Copson and Holloway.³⁰⁸ We now turn to these three studies in detail.

Pinnizzotto and Finkel obtained the partial cooperation of the FBI in their study, that is to say, the profilers tested were from the BSU, but BSU profilers who were contacted to participate were free to refuse.³⁰⁹ The researchers reported that they had difficulty obtaining agreement to participate from then-active profilers at the BSU, only two of whom consented, and they then completed the “expert/teacher” test group (a small group, n=4) with two persons who had apparently been BSU profilers but were no longer engaged in the practice actively.³¹⁰ The second profiler test group consisted of six detectives from police agencies across the country who had been trained by the BSU as profilers,³¹¹ for a total of ten profilers (the reported results are combined into a single profiler score).³¹² The other groups consisted of six detectives with experience in both homicide and rape investigations from a large urban police force,³¹³ six clinical psychologists “naïve to both criminal profiling and criminal investigation,”³¹⁴ and six university undergraduate students.³¹⁵

The test material consisted of the details of two actual criminal cases, a rape and a homicide, which had concluded with both apprehensions and convictions.³¹⁶ Half the members of each test group were given the rape file, and half the homicide file.³¹⁷ The respondents then answered a questionnaire about the inferred characteristics of the perpetrator, and also gave a narrative profile.³¹⁸ On the results of the objective questionnaire, though the profilers were significantly more accurate than the other groups in the rape case, they were not in the murder case.³¹⁹ And most importantly, the profilers got one-third of the questions wrong even in the rape case, and two thirds wrong in the homicide case.³²⁰

³⁰⁸ See Gary Copson & K. Holloway, *Coals to Newcastle?, Pt. 2: An Analysis of Offender Profiling Advice, Methods and Results*, as described in Gudjonsson & Copson, *supra* note 256, at 72-75.

³⁰⁹ See Pinizzotto & Finkel, *supra* note 306, at 218.

³¹⁰ See *id.*

³¹¹ See *id.*

³¹² See *id.* at 223 (Table 1), 224 (Table 2), 225 (Table 3), 226 (Table 4).

³¹³ See *id.* at 218-19.

³¹⁴ See *id.* at 219.

³¹⁵ See *id.*

³¹⁶ See *id.*

³¹⁷ See *id.* at 220.

³¹⁸ See *id.* at 220-21.

³¹⁹ See *id.* at 224 (Table 2).

³²⁰ See *id.* at 223-24. Pinizzotto and Finkel reanalyzed the results giving half credit for some of the inaccurate multiple choice answers based on the judgment of the “expert” profiler subgroup that some wrong answers were less wrong than others; however, they never set out the results of that reanalysis, simply asserting that for both cases the only significant differences that emerged

The narratives were also analyzed. The profilers wrote a lot more than anyone else, but a higher percentage of their statements were non-confirmable (well over half in both cases) than those of any other group.³²¹ They made the highest number of confirmable statements in both cases by a factor of two, which meant that they made the highest number of accurate statements, but also the highest number of inaccurate statements,³²² in each case. Their mean accuracy was 82 percent in the rape case³²³ (compared with 81 percent for the homicide detectives, 82 percent for the psychologists and 91 percent for the students, who only made a fifth as many predictions as profilers, however) and 76 percent for the homicide case (compared with 83 percent for the detectives, 76 percent for the psychologists and 84 percent for the students). The most that can be said is that the profilers seemed to have been able to sustain the 75-80 percent accuracy rate which was typical of other groups over a wider range of details, which were accompanied by a large number of unconfirmable propositions such as statements about the offender's mental processes.

Kocsis *et al.* contacted forty active profilers in several countries and asked them to participate in a study of profiler skills.³²⁴ Only five agreed, and no further information is given on them except that they were four men and one woman, ranging in age from twenty-seven to forty-five years, and that they had been "consulted by a law enforcement agency for the purposes of constructing a psychological profile."³²⁵ The other test groups were a group of thirty-five Australian police detectives, a group of thirty Australian psychologists not involved in "forensic or criminal psychology," a group of thirty-one university undergraduates, and a group of twenty self-described psychics.³²⁶ A single set of test materials composed of the details of a closed case was given to all five test groups: profilers, policemen, psychologists, students and "psychics."³²⁷ The participants were then asked to write a narrative description of the perpetrator,³²⁸ fill out a forty-five item multiple-choice questionnaire on offender characteristics (of which only

were an advantage of the profiler group compared to the student group. *See id.* at 224.

³²¹ *See id.* at 223 (Table 1).

³²² *See id.* Pinizzotto and Finkel only set out the number of accurate predictions. They did not add the next obvious line for the number of inaccurate predictions, which is easily derived by subtracting the number of accurate predictions from the total number of confirmable predictions.

³²³ *See id.* Pinizzotto and Finkel did not give these percentages, which are derived by dividing the number of accurate predictions by the total number of predictions. Pinizzotto & Finkel spent the latter part of the article making excuses for the profilers' performance. Pinizzotto later went to work for the FBI.

³²⁴ *See* Kocsis *et al.*, *supra* note 307, at 316.

³²⁵ *Id.*

³²⁶ *See id.*

³²⁷ *See id.* at 317.

³²⁸ Unlike Pinizzotto and Finkel, Kocsis *et al.* do not appear to have scored the written descriptions.

thirty items were used to score results) and a second questionnaire comprising “agree/disagree” choices on 300 personality descriptors for the offender. The results as to the offender characteristic questionnaire show that: (1) No group got as many as half of the objective items correct (the percentages ranged from 46 percent right for profilers to 38 percent right for psychics);³²⁹ (2) There were no significant differences among any of the groups in total performance or performance on any of the sub-measures;³³⁰ except: (3) Psychologists accurately identified significantly more physical characteristics of the offender than did police officers or psychics, and more offense behaviors than police officers.³³¹ The profilers were not significantly better than any other group on any index.³³² These results lend no support to the claims of profilers.

Kocsis *et al.* also noted one weakness in their own study (a weakness partially shared by Pinnizzotto & Finkel, as already noted): most profilers refused to take part in the tests, leaving the results open to the interpretation that better profilers might have performed better.³³³ Kocsis quotes British profiler Paul Britten as saying that “psychological profilers tend to exhibit exceptionally strong professional rivalry and jealousy, and thus they hesitate to expose any shortcomings in their profiling expertise where there is no personal gain in their doing so . . .”³³⁴ and they conclude by observing that any such tendency “is clearly a major impediment to the conduct of scientific investigation into the skills involved in psychological profiling.”³³⁵

We have seen this pattern before in forensic science fields.³³⁶ Nevertheless, we should bear in mind where the burden of persuasion lies, in science and in law. No one should benefit from their own failure to aid the generation of defensible data. This is especially true of the

³²⁹ See *id.* at 320 (Table 1). The percentages are not given in the table but are derived by dividing the total correct by the total number of questions scored, which is derived by adding sub-test number of questions given. See *id.* at 319.

³³⁰ See *id.* at 320 (Table 1).

³³¹ See *id.* at 321. As to the personality assessments, psychologists were in the same position as profilers in that they checked off many more boxes than anyone else and got both the most right and the most wrong. When this was accounted for, no group was significantly more accurate than another. See *id.* at 323.

³³² The authors undertake some *post-hoc* consolidated data analyses that combine some near significant profiler performances into significance. Perhaps the most questionable is the comparison of profiler performance to total non-profiler performance when the non-profiler aggregation is made up of 17 percent practicing psychics (not a very typical group for any purposes). Of course, combining groups raises the number of the non-profiler group so high that apparent significance may emerge as an artifact. However, even as a result of this questionable process, the only significant advantage in accuracy for profilers over the combined group was not for total accuracy but for accuracy on the twelve questions shared between their test and the Pinnizzotto and Finkel test. See *id.* at 322.

³³³ See *id.* at 324.

³³⁴ *Id.*

³³⁵ *Id.*

³³⁶ This was noted as long ago as 1939 in regard to handwriting identification experts. See Fred Inbau, *Lay Witness Identification of Handwriting*, 34 ILL. L. REV. 433, 440 n.11 (1939).

profilers who have been part of the FBI's operation, for a simple reason. They have access to data that would settle the issue of raw accuracy (if not relative advantage over lay persons) effectively and efficiently. They presumably have files containing every written profile generated by members of their organization in open cases for the purposes of giving direction to an investigation. In a not insignificant number of those cases the perpetrator was caught and convicted. It would not be an impossible research effort to develop a defensible protocol for objectively scoring the accuracy of those profiles and generating statistical accuracy and error rates for every variable commonly included, and for global accuracy. Until such an effort is undertaken, their claims to some mystical level of accuracy ought to be regarded more as a form of self-promoting science fiction than as fact.

The reason we can say that it is possible to do such research is because it has been done in Britain, which brings us to the last study, Copson and Holloway.³³⁷ This study deals with the examination of fifty solved cases in which profiles were created before the perpetrators were identified, divided between cases involving statistical profilers and cases involving clinical profilers.³³⁸ First, all the profiles had a high number of items which could not be verified, either because they were not empirically verifiable (descriptions of subjective processes, normative statements) or because the case file did not contain information concerning them.³³⁹ As might be expected from the Pinizzotto & Finkel study, the clinical profilers said a lot more, but over 50 percent of what they said could not be scored for accuracy, while 80 percent of the statistical profiler's statements could be scored.³⁴⁰ On the scorable statements, statistical profilers were accurate 69 percent of the time, while the clinical profilers were accurate 74 percent of the time.³⁴¹ This may give some mild support to the claim that clinical profilers can beat the probabilities reflected in formal data sometimes, but still they were wrong more than a quarter of the time, on the aggregate. While still a valuable investigatory tool perhaps, the existing data does not indicate that process of offender profiling results in sufficiently reliable information to support evidentiary admissibility.³⁴²

³³⁷ See Copson & Holloway, in Gudjonsson & Copson, *supra* note 256, at 72-75.

³³⁸ See *id.* at 73.

³³⁹ See *id.*

³⁴⁰ See *id.*

³⁴¹ See *id.* This error rate is consistent with the range revealed in Pinizzotto & Finkel, *supra* note 306, and less than suggested by Kocsis *et al.* and the implications of the FBI reliability study. Copson and Holloway did identify a subset of profilers whose accuracy rate was around 79 percent (a dangerous course post-test, for dredging reasons) but, as Gudjonsson & Copson, *supra* note 256, at 74, state, "even taking the very best result on offer, a detective must expect more than one-fifth of his advice [from a profiler] to be misleading." *Id.*

³⁴² In a Note, Scott Ingram urges the admission of profiles, but seems unaware of most of the then-extant data concerning their reliability and validity. See Scott Ingram, Note, *If the Profile Fits: Admitting Criminal Psychological Profiles into Evidence in Criminal Trials*, 54 WASH. U. J.

IV. PROFILING AND LINKAGE/SIGNATURE ANALYSIS: OF JOHN
DOUGLAS AND STEVEN PENNELL

Why the lengthy *discursus* above on the history and validity of perpetrator profiling when that is not what was directly at issue in the *Fortin* case, or any case in which “linkage expertise” is being proffered? Though courts have generally rejected testimony concerning profiling frankly so offered,³⁴³ they have often bent over backwards to admit profiling-based testimony, or testimony by profilers, when it could be labeled differently.³⁴⁴ The weight of the proposed expert’s claims are added to substantially when he can invoke a background in “research and publication” and extensive experience in the practice of offender profiling as a member of the FBI Behavioral Science Unit, which research and experience is often given as the ground out of which the expertise claimed in court has grown. In considering what claims are being made in the individual case, therefore, it is important to know the realities of both the research and the profiling practice being invoked as the precursors of the claimed expertise.

Of course, when the prosecutor in the *Fortin* case called the Academy Group and was put in touch with Roy Hazelwood, he wasn’t asking Hazelwood for a profile of the killer of Melissa Padilla. Instead, he was interested to know if Hazelwood thought he could offer an expert opinion on whether the Maine crime (indisputably committed by Fortin) was committed by the same person as the Padilla murder. Ultimately, Hazelwood thought that he could.³⁴⁵

It cannot be overemphasized that what Hazelwood was offering to do was only indirectly related to either his own research, the FBI research in general, or his claimed skill as a profiler. In essence, as previously noted, he was claiming the ability to examine the details of the two crimes and to determine accurately by virtue of their shared characteristics that they had been committed by one and the same person, and to do so more reliably than a jury. But what, if anything, lay behind this claim? To support his claim, Hazelwood invoked a theory, to which he attached the name “linkage analysis.” And whence came the theory?

Linkage analysis, theory and practice, appears to have been invented by John Douglas especially for the purpose of justifying his own

URB. & CONTEMP. L. 239 (1998). The only study he cites is Pinizzotto & Finkel, which he mischaracterizes in a single line but he never deals with the actual data at all. See *id.* at 264.

³⁴³ See *id.* at 259.

³⁴⁴ See, for example, the contrasting statements about profiling testimony and John Douglas’ “signature” testimony in *Pennell v. State*, 602 A.2d 48 (Del. 1991), discussed *infra* text accompanying note 400-401.

³⁴⁵ See Robert R. Hazelwood, Report to Middlesex (NJ) Court Prosecutor, Oct. 17, 1997.

acceptability (and, by extension, those of other BSU profilers such as Hazelwood who might follow him) as an expert witness offering testimony in criminal cases that separate crimes were all committed by the same perpetrator. And when we say “invented” we do so advisedly. Whatever one may think of the validity of the claims made for perpetrator profiling, at least one can be sure of the essential bona fides of the underlying effort to develop some helpful skill in describing perpetrators from the details of their crimes. This is because profiling was looked upon as almost exclusively an investigatory aid from the beginning. And while personal vanity or institutional power and status might lead its practitioners to exaggerate its accuracy or usefulness, there is no doubt that the main goal was maximal accuracy. No such statement can be made in regard to “linkage analysis” because it appears to have been developed, not as an investigatory aid, but primarily as a means of obtaining either the admission of other crimes evidence which might not otherwise be admitted, or a means to convince the jury that the other crimes evidence was more meaningful than they otherwise might believe, or both. In sum, it was not a way to identify unknown perpetrators, but a tool to help build a case against defendants already believed to be guilty.³⁴⁶ As Douglas has expressed it:

[F]rom our work in behavioral profiling from crime scenes and signature analysis, there is another arrow in the police's and prosecution's quiver. In and of itself, it's not usually enough to convict. But taken together with one or more of the other elements, it can often link various crimes together and be just what is needed to put a case over the top.³⁴⁷

Though no publicly available written *precis* of Douglas's linkage theory was published until 1992, Douglas first made claims of such linkage expertise in connection with the 1989 trial of Steven B. Pennell. Pennell was charged in a single indictment with the murders of three prostitutes in three separate incidents. The indictment grew out of the

³⁴⁶ Some process of linkage analysis can be significant in tying crimes together for investigatory purposes, of course, but such a process need not display the over-the-top claims of specificity and uniqueness adopted by the Douglas theory. Indeed, there is reason to believe that the theory may be applied more circumspectly when admission of testimony is not the goal. Compare the qualified results of the linkage analysis by Hazelwood in the case of Werner Ferrari, the Swiss child murderer described in MICHAUD & HAZELWOOD, *EVIL*, *supra* note 31, at 178-87 (where administrative closing of open case files was the only purpose), with his performance in the *Fortin* case. And while Douglas must get the credit (or blame) for the published formulation of the theory, it is likely the case that Hazelwood contributed to it in consultation through the years. The need for an investigatory approach to linkage analysis has been pointed out publicly by Stephen Egger in 1984. See Steven A. Egger, *A Working Definition of Serial Murder and the Reduction of Linkage Blindness*, 12 J. POL. SCI. & ADMIN. 348 (1984). Hazelwood and Douglas worked closely in the years leading up to the theory. The only discernable difference in approach between them is that Hazelwood sometimes uses the term “ritual” as a synonym for “signature.” See MICHAUD & HAZELWOOD, *EVIL*, *supra* note 31, at 178.

³⁴⁷ DOUGLAS & OLSHAKER, *MINDHUNTER*, *supra* note 216, at 259-60.

following facts:³⁴⁸

On November 29, 1987, in northern Delaware, the body of a known prostitute, Shirley Ellis, was found³⁴⁹ in a construction site in Newark, Delaware.³⁵⁰ She had been savagely beaten and tortured. There were multiple skull injuries consistent with having been struck by a hammer,³⁵¹ ligature strangulation marks around her neck,³⁵² marks of bindings around her wrists,³⁵³ pattern bruising on the left breast and nipple³⁵⁴ consistent with having been tortured with a pair of pliers.³⁵⁵ The bindings and ligatures had been removed. There was black duct tape in her hair of a special type used by electricians.³⁵⁶ She was wearing a pair of aqua blue pants.³⁵⁷

On June 29, 1988 the nude body of Catherine DiMauro was found at another construction site in the same general northern Delaware area.³⁵⁸ She too had had her hands bound, been beaten on the head with what appeared to be a hammer, was strangled with a ligature, and had her breasts subject to severe bruising consistent with an attack with pliers.³⁵⁹ Once again, the ligatures had been removed and taken away. There was also gray duct tape in her hair. In addition, her body had numerous blue textile fibers³⁶⁰ later identified as probable carpet fibers³⁶¹ on various points of its surface, and two red textile fibers were found on her face.³⁶²

Because Ellis was last seen walking along U.S. Highway 40, and because both of the deceased women were known to hitchhike and cruise for customers along a stretch of U.S. Highway 40, in July of 1988

³⁴⁸ The *Pennell* case was the subject of a number of opinions, five of which contain significant recitations of the details of the case: State v. Pennell, 1989 WL 112555, at *1 (Del. Sup. Ct., Sept. 12, 1989) (Gebelein, J.); State v. Pennell, 587 A.2d 513 (Del. Sup. Ct. 1989) (Gebelein, J.); State v. Pennell, 583 A.2d 1348 (Del. Sup. Ct. 1990) (Gebelein, J.); Pennell v. State, 602 A.2d 49 (Del. 1991). The facts are drawn, as indicated, from those opinions.

³⁴⁹ See *Pennell*, 602 A.2d at 49.

³⁵⁰ See ROBERT D. KEPPEL, SIGNATURE KILLERS 195 (1997). In his section heading, Keppel repeats the name given these murders by John Douglas—the “I-40 murders.” See DOUGLAS & OLSHAKER, MINDHUNTER, *supra* note 216, at 247; DOUGLAS & OLSHAKER, JOURNEY INTO DARKNESS, *supra* note 224, at 26, 51-52. It is a mystery why Douglas would give the murders this title, since Interstate 40 does not run within 200 miles of Delaware. Keppel, however, correctly identifies the road as U.S. Highway 40 in the text, seems reliable on the basic facts of the case (though he confuses Pennell’s first and second trial), and is used as a source whenever the facts in opinions are not specific enough.

³⁵¹ See *Pennell*, 602 A.2d at 49.

³⁵² See *id.*

³⁵³ See *id.*

³⁵⁴ See *id.*

³⁵⁵ See *Pennell*, 1989 WL 112555, at *1.

³⁵⁶ See *Pennell*, 602 A.2d at 52.

³⁵⁷ See *id.* at 49.

³⁵⁸ See KEPPEL, *supra* note 350, at 40.

³⁵⁹ See *Pennell*, 602 A.2d at 50.

³⁶⁰ See *id.*

³⁶¹ See *Pennell*, 1989 WL 112555, at *1.

³⁶² See *id.*

the police began an undercover operation using female officers dressed as prostitutes in an attempt to develop leads in connection with anyone who approached them who might be the perpetrator.³⁶³ They were not allowed to enter any john's vehicle.³⁶⁴

During the period of the undercover operation, another prostitute, Margaret Finner, disappeared.³⁶⁵ She was last seen getting into a blue van, which her pimp said was driven by a large white male and had no side windows and round headlights.³⁶⁶

On September 14, 1988, Officer Renee Lano was working as an undercover decoy along Route 40.³⁶⁷ After having driven past her several times, a blue van with no side windows and round headlights pulled up on the shoulder ahead of her, and the driver motioned for her to get in.³⁶⁸ She opened the passenger side door but did not enter, engaging the driver in conversation as she observed that the truck was carpeted in blue.³⁶⁹ She surreptitiously picked up a sample of a few fibers from the area around the open door and then concluded her negotiations with the driver.³⁷⁰ The truck was registered to a Steven Pennell and his wife jointly.³⁷¹ The fibers were analyzed at the FBI Laboratory and found to be of the same type, material and color as those found on DiMauro's body.³⁷² Search warrants were obtained which resulted, among other things, in the discovery of blood in the back of the truck which was DNA-matched to DiMauro.³⁷³ In addition, fibers from a red cloth in the van were found to be indistinguishable from those recovered from DiMauro's face, and a fiber from Pennell's buck hunting knife was found to match the fibers of Shirley Ellis's pants.³⁷⁴ Besides this, they seized plastic handcuffs and other items of evidence.³⁷⁵

On September 20, 1988, before the lab results on the rug fibers from the van were known, another body was discovered. Michelle Gordon, another prostitute known to work Route 40, was found having washed up on some rocks by the Chesapeake and Delaware canal,³⁷⁶ which runs parallel to Route 40 about a mile to the south.³⁷⁷ Because of her submersion, an exact cause of death was never determined.³⁷⁸ She

³⁶³ See KEPPEL, *supra* note 350, at 201.

³⁶⁴ See *id.*

³⁶⁵ See Pennell, 602 A.2d at 50.

³⁶⁶ See Pennell, 1989 WL 112555, at *2.

³⁶⁷ See *id.* at *3.

³⁶⁸ See *id.* at *3-4.

³⁶⁹ See *id.* at *4.

³⁷⁰ See *id.*

³⁷¹ See *id.*

³⁷² See *id.*

³⁷³ See Pennell, 587 A.2d at 515.

³⁷⁴ See Pennell, 602 A.2d at 52.

³⁷⁵ See *id.* at 48.

³⁷⁶ See *id.* at 50.

³⁷⁷ See THE ROAD ATLAS 24 (2002) (map of Delaware).

³⁷⁸ See Pennell, 602 A.2d at 50.

had not been strangled or beaten on the head. There was evidence that her arms and legs had been bound, however, she had been beaten on the buttocks and hips, and one of her nipples had been cut off.³⁷⁹

Pennell was indicted for the murders of DiMauro, Ellis and Gordon. As to DiMauro, the prosecutor had to feel pretty confident, given the DNA evidence and the fiber evidence. As to Ellis, the case was still pretty good, since the DiMauro and Ellis crime scenes and injuries were almost photocopies of each other, and there was the single fiber on the buck hunting knife consistent with Ellis's pants.³⁸⁰ However, Gordon was a problem. The details of the other two victims had been in the press for a long time, and the possibility that Michelle Gordon was the victim of a copycat, or someone with a personal grudge using the other killings as cover, was obvious. What to do? Get someone with intimidating credentials from the FBI to say that the same person killed all three, since it was a lock that Pennell killed DiMauro. Enter John Douglas.

We have not obtained a transcript of Douglas's Pennell testimony. However, his claims in his various memoirs concerning the Pennell case make it clear that he regards it as the maiden voyage of his theory of linkage, with its two elements, "Modus Operandi" and "Signature." Since these are the elements that were invoked by Hazelwood to justify his proposed testimony in *Fortin*, we should examine the claims made for them in some detail. In his memoir, *Mindhunter*, Douglas describes his thinking thus:

[W]hen I started research into the minds and motivations of serial murderers, then, when I began analyzing crime scenes for behavioral clues, I would look for the one element or set of elements that made the crime and the criminal stand out, *that represented what he was*. . . . [E]ventually, I would come up with the term *signature* to describe this unique element and personal compulsion, which remained static. And I would use it as distinguishable from the traditional concept of *modus operandi*, which is fluid and can change.³⁸¹

Later Douglas has this to say about "Modus Operandi" and "Signature":

Both of these are extremely important concepts in criminal investigative analysis, and I have spent many hours on the witness stand trying to get judges and juries to understand the distinction between them.

Modus operandi—MO—is learned behavior. It's what the perpetrator does to commit the crime. It is dynamic—that is, it can change. Signature, a term I coined to distinguish it from MO, is *what the perpetrator has to do to fulfill himself*. It is static; it does not change.

³⁷⁹ See *Pennell*, 583 A.2d at 49.

³⁸⁰ See *Pennell*, 602 A.2d at 55.

³⁸¹ DOUGLAS & OLSHAKER, *MINDHUNTER*, *supra* note 216, at 58-9.

For example, you wouldn't expect a juvenile to keep committing crimes the same way as he grows up unless he gets it perfect the first time. But if he gets away with one, he'll learn from it and get better and better at it. That's why we say that MO is dynamic. On the other hand, if this guy is committing crimes so that, say, he can dominate or inflict pain on or provoke begging and pleading from a victim, that's a signature. It's something that expresses the killer's personality. It's something he needs to do

In many states, the only way prosecutors can link crimes is by MO, which I believe we've shown is an archaic method. In the Christopher case, a defense attorney could easily make the argument that the Buffalo .22-caliber shootings and the Manhattan midtown slashings showed a markedly different *modus operandi*. And he'd be right. But the signature is similar—a propensity to randomly assassinate black men fueled by racial hatred.

The shootings and the (cabdriver) eviscerations, on the other hand show me a markedly different signature. The individual who cut out the hearts, while still possessing a related underlying motivation, has a ritualize obsessive-compulsive signature. Each type needs something out of the crime, but each one needs something different.

The differences between MO and signature can be subtle. Take the case of a bank robber in Texas who made all of his captives undress, posed them in sexual positions, and took photographs of them. That's his signature. It was not necessary or helpful to the commission of a bank robbery. In fact, it kept him there longer and therefore placed him in greater jeopardy of being caught. Yet it was something he clearly felt a need to do.

Then there was a bank robber in Grand Rapids, Michigan. I flew out to provide on-site consultation in the case. This guy also made everyone in the bank undress, but he didn't take pictures. He did it so the witnesses would be so preoccupied and embarrassed that they wouldn't be looking at him and so couldn't make a positive ID later on. This was a means toward successfully robbing the bank. This was MO.³⁸²

Douglas's first published exposition of his theory came in 1992, in an article in the *FBI Law Enforcement Bulletin*,³⁸³ which he later revised and included as a chapter in the back of the *Crime Classification Manual*.³⁸⁴ That chapter is the fully expanded exposition of the theory, intended for specialist audiences, and

³⁸² *Id.* at 252.

³⁸³ John E. Douglas & Corinne Munn, *Violent Crime Scene Analysis: Modus Operandi, Signature and Staging*, in 61 FBI LAW ENFORCEMENT BULLETIN 2, Feb. 1992, at 1. Hazelwood has testified that this was the first published exposition of the theory that he follows. See *State v. Fortin*, 318 N.J. Super. 577, 591 (N.J. Super. Ct. App. Div. 1999).

³⁸⁴ John E. Douglas & Corinne M. Munn, *Modus Operandi and the Signature Aspects of Violent Crime*, in JOHN E. DOUGLAS ET AL., *CRIME CLASSIFICATION MANUAL* 259-68 (1992). Given the topic and the small amount of research involved, the role of Munn, an "Honors Intern" at the FBI academy, is unclear, but the content is chargeable entirely to Douglas. Calling this work the "Crime Classification Manual" is something of a misnomer, as it applies only to Homicide, Arson, and Rape—the crimes with which the National Center for the Analysis of Violent Crime is most concerned.

says as follows:

The MO has great significance when investigators attempt to link cases. An appropriate step of crime analysis and correlation includes connecting cases due to similarities in MO. However, an investigator who rejects an offense as the work of a serial offender solely on the basis of disparities in MO . . . has made a mistake. What causes an offender to use a certain MO? What influences shape a modus operandi? Is it static or dynamic? By answering these questions, one sees the error of attributing too much significance to the MO when linking crimes.

A novice prowler prepared to enter a house through a basement window to burglarize it. Although the window was closed and locked, the prowler shattered the window and gained access to the house. He had to rush his search for valuables because he feared the breaking window had attracted attention. During a later crime, he burglarized another residence, but this time he brought tools with him to force the lock and keep the noise minimal. This allowed him time to commit the crime and to obtain a more profitable haul.

This example demonstrates that MO is a learned behavior. The offender's actions during the perpetration of a crime form the MO. The offender develops and uses an MO over time because it works, but it also continuously evolves. The modus operandi is very dynamic and malleable. During his criminal career, an offender usually modifies the MO as he gains experience. The burglar refines his breaking and entering techniques to lower his risk of apprehension and to increase his profit. Experience and confidence will reshape an offender's MO. Incarceration usually impacts on the future MO of an offender, especially the career criminal. He refines the MO as he learns from the mistakes that led to his arrest.

The victim's response can also significantly influence the evolution of an MO. If the rapist has problems controlling a victim, he will modify his MO to accommodate resistance. He may bring duct tape or other ligatures, he may use a weapon, or he may blitz-attack the victim and immediately incapacitate her. If such measures are ineffective, he may resort to greater violence or kill the victim. Thus, the MO will evolve to meet the demands of the crime.

The violent, repetitive offender often exhibits another element of criminal behavior during an offense: the signature aspect, or calling card. This criminal conduct goes beyond the actions necessary to perpetrate the crime. It composes a unique and integral part of the offender's behavior while he is committing the offense.

An offender's fantasies often give birth to violent crime. As the offender broods and daydreams, he develops a need to express these violent fantasies. When he finally acts out, some aspect of the crime will demonstrate a unique, personal expression or ritual based on these fantasies. Committing the crime does not satisfy the offender's needs. This insufficiency compels him to go beyond the scope of perpetration and perform his ritual. When the subject displays this ritual at the crime scene, he has left his calling card.

How does the crime scene manifest this calling card, or signature aspect? The subject introduces an aspect of his personality into the scene through this ritual. The crime scene displays this aspect by peculiar crime scene characteristics or unusual offender input during the perpetration of the crime. A rapist demonstrates his signature by engaging in acts of domination, manipulation, or control during the verbal, physical, and/or sexual phase of the assault. Exceptionally vulgar and/or abusive language or scripting represents a verbal signature. When the offender scripts a victim, he dictates a particular verbal response from her (e.g., "Tell me how much you enjoy sex with me" or "Tell me how good I am"). The use of excessive physical force exemplifies another aspect of a subject's signature. One example of signature sexual behavior involves the offender who repeatedly engages in a specific order of sexual activity with different victims.

The core of the offender's ritual will never change. Unlike the MO, it remains a constant and enduring part of the offender. However, signature aspects may evolve (e.g., the lust murderer, who performs greater postmortem mutilation as he progresses from crime to crime). Elements of the original ritual become more fully developed. In addition, the signature does not always show up in every crime because of contingencies that might arise, such as interruptions or an unexpected victim response.

The investigator cannot always identify the signature aspect. Violent offenses often involve high-risk victims or decomposition from outdoor body disposal, both of which interfere with recognition of signature.

A rapist entered a residence and captured a woman and her husband. The offender ordered the husband to lie on his stomach on the floor. He then placed a cup and saucer on the husband's back. "If I hear that cup move or hit the floor, your wife dies," he told the husband. He then took the wife into the next room and raped her.

In another situation, a rapist entered a house and ordered the woman to phone her husband and use some ploy to get him home. Once the husband arrived, the offender tied him to a chair and forced him to witness the rape of his wife.

The rapist who used the cup and saucer had developed an effective modus operandi to control the husband. The second rapist, however, had gone beyond the simple commission of rape. The full satisfaction of his fantasies not only required raping the wife, but also humiliating and dominating the husband. His personal needs compelled him to perform the signature aspect of crime.

....

When investigators attempt to link cases, the modus operandi plays an important role. However, as stated previously, MO should not be the only criterion used to connect crimes, especially with the repeat offender who alters the MO through experience and learning. The first offenses may differ considerably from later offenses. However, the signature aspect remains the same, whether it is the first of-

fense or one committed ten years later. The ritual may evolve, but the theme persists.³⁸⁵

We have set out these quotations *in extenso* because it is not very often that one is able to actually quote virtually the entire *corpus* of theory that lies behind a claimed area of expertise, but this is pretty much it. If, after having read it, you have been satisfied that what is set out is internally coherent and empirically justified, your powers of perception exceed ours. We have read it many times, and even typed it through in writing this article, and we are still not sure that we are clear on what is being said. However, from what we do understand, we believe it is fair to say that this asserted theory has some demonstrable and serious problems.

First, from a historical evidence law perspective, Douglas is simply wrong to assert that “in many states, the only way that prosecutors can link crimes is by MO,” at least as he uses the term. He is also wrong to assert that he “coined” or “came up with” the term “signature” in the context of asserted proof that two crimes were committed by one perpetrator. As to the first point, in traditional evidence theory, the emphasis has always been on characteristics of the two crimes that were so unusual as to raise a strong inference of common perpetration, whether Douglas would classify those characteristics in his scheme as belonging to “MO” or to “signature.”³⁸⁶ Indeed the term “signature” has been commonly used to refer to such a detail or combination of details since it was first applied in this context by Charles T. McCormick in the first edition of his *Handbook on Evidence* in 1954,³⁸⁷ and there are literally hundreds of examples of the usage in the case law prior to Douglas’s arrival on the scene.³⁸⁸

Second, what the Douglas theory does create is a simple two category taxonomy of the reconstructed actions of a criminal at a crime scene, which assigns each action either to the “MO” category, or to the “Signature (or “calling card”)” category. In order to be useful, such a taxonomy would, as a minimum initial condition, have to be reasonably reliable. But there are serious reasons to doubt its likely reliability (and of course there has never been any empirical testing of the reliability of the classification system). At a glance it seems less likely to be reliable than the “organized/disorganized” dichotomy (which, you will recall, did not turn out to be very reliable under test). The first reason to have reservations about reliability lies in the expression of the criteria for assigning something to the category “MO”: “It’s what the perpetrator does to commit the crime” or, “[t]he offender’s actions during the per-

³⁸⁵ *Id.* at 260-63 (section heads omitted).

³⁸⁶ See IMWINKELRIED, *supra* note 21.

³⁸⁷ CHARLES T. MCCORMICK, A HANDBOOK ON THE LAW OF EVIDENCE 328 (1st ed. 1954).

³⁸⁸ Go to Westlaw’s “ALLCASES” database (i.e., all cases since 1945), type in “signature crime*” and observe the result.

petration of a crime form the MO.” Taken literally, this would cover every action, making the things labeled “signature” a special subset of MO behaviors. That might have been a good way to go, since Douglas’ notion of “signature” applies only to a tiny subset of crimes, those committed by “violent repetitive offenders,” or at least those who are violent and may be repetitive. Hence, most crimes will not have “signatures,” so a system which made signature a specialized subset of MO would make sense. But that is clearly not what was intended, since the entire thrust of the theory as set out is that MO is one thing and signature is something else entirely.

So, though the above statements about MO being “what the offender does to commit the crime” are made, they cannot be taken literally. So how are they to be taken? A clue comes from the discussion of the bank robbers who made the witnesses undress, and the rapist with the teacup. In both examples, Douglas attempts to assign behaviors depending on whether they are “necessary or helpful to the commission” of the crime on the one hand, or are “not necessary or helpful to the commission” of the crime or go “beyond the actions necessary to perpetrate the crime” on the other. The problem is that, taking these statements literally, what is MO and what is signature may vary from jurisdiction to jurisdiction depending on the legal definition of the crime. In a case of homicide, if the crime were “murder” then any torture would perhaps be unnecessary to the commission of the crime. If however, the jurisdiction had a separate crime for “sadistic murder,” then the torture would be necessary to the commission of the crime. Douglas seems to have something else in mind, but what that something else is remains unclear. A blitz attack to efficiently disable a victim would be MO, but the same blitz attack to scare and impose pain on the victim would be signature. In a world of mixed motives, the distinction is unlikely to be reliable.

Even if the taxonomy were reliable, it does not correlate directly with details that raise weaker or stronger inferences of common perpetration. For instance, assuming the teacup to be properly classed as an MO detail, the inference of common perpetration in a series of “teacup rapes” would be strong regardless of the fact that it was not a “signature” aspect in the Douglas system. Douglas admits as much when he concedes that linkage through MO can be strong.

Ironically, linkage through the category with the labels that appear to make a strong claim of just such particularization, (“signature” or “calling card”), may, on the average, be weak. For while the labels appear to have been selected to communicate a claim of individual particularization, they are, in a sense, a fraud. To borrow a phrase from other asserted forensic sciences, close examination of what is said about “signature” reveals that it is usually not an individual characteristic, but

a class characteristic.³⁸⁹ Thus, Douglas identifies the common signature in the Buffalo .22 caliber shootings and the Manhattan “Midtown Slasher” killings as “a propensity to randomly assassinate black men fueled by racial hatred.”³⁹⁰ But unless every random act of race-motivated murder in the U.S. in the last few decades was committed by the same person, this is a “signature” characteristic for a class of offenders. The same can be said of the “acts of domination, manipulation, or control during the verbal, physical, and/or sexual phase of the assault,” and the “[e]xceptionally vulgar and/or abusive language” and the “use of excessive physical force” whereby the perpetrator “demonstrates his signature.”³⁹¹ While in a particular case, such variables may show a pattern rare enough to infer particular common perpetration, they can fit the category of “signature” even if generic and variable. Indeed, the variability is built into the system by assertions that sometimes signature aspects may be omitted due to interfering circumstances, and sometimes they may differ in details because of “evolution” (though not change—no, no, they don’t change). “The theme will remain the same.” But a theme is too generic to be a signature in the McCormick sense, and in the sense implied by the rhetoric adopted for the theory, both by Douglas and by the courts.³⁹² And it is not just we who have concluded that operationally Douglas’ “signature” is a dominantly a class characteristic. Even Robert D. Keppel, one of Douglas’ strongest disciples, the only one ever to have written extensively about applications of this “linkage analysis” theory, adopts descriptive language which generally applies “signature” to types of perpetrators, not to individual perpetrators.³⁹³

³⁸⁹ See the discussion of the distinction in John I. Thornton and Joseph L. Peterson, *The General Assumptions and Rationale of Forensic Identification*, in FAIGMAN ET AL., *supra* note 194, § 24-2.1. Professors Thornton and Peterson observe:

[A document] examiner may note an unusual letter formation, which in the experience of that examiner seems to be unique. . . . But it may be that every schoolchild in a Bulgarian town was taught to execute that particular letter formation. The characteristic may be obscure, but it is still a class characteristic, not an individual characteristic, and should be given only the weight that a class characteristic deserves and not the additional weight that ordinarily would be given to an individual characteristic.

Id. Our use of the distinction between class characteristics and individual characteristics for illustrative purposes in this article should not be taken as an indication that we believe that the distinction actually represents a distinction in kind rather than degree. “Individual Characteristics” are in fact generally class characteristics of an extremely small class.

³⁹⁰ DOUGLAS & OLSHAKER, *MINDHUNTER*, *supra* note 216, at 252.

³⁹¹ Douglas & Munn, *Violent Crime Scene Analysis*, *supra* note 383, at 261.

³⁹² See, for example, the formula invoked by the Supreme Court of New Jersey in *State v. Fortin* “the prior criminal activity . . . must be so nearly identical in method as to earmark the crime as defendant’s handiwork. The conduct in question must be unusual and distinctive so as to be like a signature” *State v. Fortin*, 162 N.J. 517, 532 (2000).

³⁹³ See KEPPEL, *supra* note 350. Keppel, who was an investigator in both the Theodore Bundy case and the Green River Killer case, is a longtime associate of Douglas from common involvement in a number of cases over the years. Keppel begins his book by setting out Douglas’ theory pretty much, *see id.* at 1-7, and he makes it sound as if the term “signature” is going to refer to something highly individualized. Throughout the rest of the book, however, he repeatedly refers

One good illustration is Douglas's own characterization of his Pennell "signature" testimony, given in his 1995 memoir:

I made it clear that regardless of the MO, the common denominator in each of the murders was physical, sexual and emotional torture. In some cases the murderer had used pliers to squeeze his victims' breasts and cut their nipples. He had bound others at the wrists and ankles, cut them on the legs, whipped or beaten their buttocks, or hit them with a hammer. So, though the methods of torture varied—the MO, if you will—the signature was the pleasure he received out of inflicting pain and hearing his victims' anguished screams. This wasn't necessary to accomplish the murder. It was necessary for him to get what he wanted to out of the crime.³⁹⁴

Of course, since the details of the torture could be and were different (at least as between the Gordon murder and the other two), Pennell shared this "signature" on this level of generality with every sadistic murderer of women, of which there have been depressingly many.

Douglas is aware of this weakness, for in describing the same case in the *Crime Classification Manual* chapter, he attempts to tie the Gordon murder to the other two more specifically. In so doing, he reports many of the facts erroneously. He claims that all victims died of blunt force trauma to the head,³⁹⁵ whereas the trial court says in its opinion that no specific cause of death was determined in regard to Gordon, nor was there evidence of the kind of head trauma present in the other two cases.³⁹⁶ Douglas claims that all the victims had "ligature marks around their necks,"³⁹⁷ though the court says there was no evidence of strangulation in Gordon's case.³⁹⁸ He also says that "[b]ody disposal was similar; Pennel [sic] left the bodies in full view, dumped with cold indifference by roadsides."³⁹⁹ However, as previously noted, Gordon was apparently thrown into the Delaware and Chesapeake canal and

to "signature" as a class characteristic. For example, Keppel states, "Frampton was a classic example of a signature killer who was a sexual sadist. He didn't just kill his victims, he 'overkilled' them," and "[d]ifferent killers have different signatures. Killers who pulverize their victims leave one type of calling card, while killers who torture living victims or who play with corpses leave another." *Id.* at 23, 26. Speaking of William Hierens, the Chicago serial killer of the late 1940s, Keppel notes, "It's not the actual wording of the notes or the writing medium, but his compulsion to leave notes that was the signature." *Id.* at 41. He also notes: "Without much question, Timothy Spencer's murder scenes were classical signatures of the anger-retaliatory type of rape-murderer who kills in response to a perceived injury or threat to his self-image from a target victim." *Id.* at 117-18. Other examples of this use of the term "signature" could be found. Additionally, the chapter titles of the book include "The Anger-Retaliatory Signature" and "The Picquerism Signature." *Id.* at 87, 124.

³⁹⁴ DOUGLAS & OLSHAKER, *MINDHUNTER*, *supra* note 216, at 254-55.

³⁹⁵ See Douglas & Munn, *supra* note 384, at 264.

³⁹⁶ See *State v. Pennell*, 583 A.2d 1348, 1349 (Del. Sup. Ct. 1990). In fact, the medical examiner later opined that Gordon had died of a heart attack while being tortured. See *Pennell v. State*, 604 A.2d 1368, 1372 (Del. 1992).

³⁹⁷ Douglas & Munn, *supra* note 384, at 265.

³⁹⁸ See *Pennell*, 583 A.2d at 1349.

³⁹⁹ Douglas & Munn, *supra* note 384, at 265.

later washed up on some rocks. And while Douglas admits that the injury done to Gordon's nipple was post-mortem (unlike the other two cases), he says that the ante-mortem injury actually escalated in Gordon's case, if you properly evaluate the injuries done to the buttocks "[t]herefore, the signature aspect of torturing a live victim was present."⁴⁰⁰ Ultimately, he accounts for the differences with a cop-out: "interference with the ritual due to contingencies arising will alter that ritual. This victim probably died too soon for Pennel [sic] to complete his signature."⁴⁰¹

When the propriety of admitting Douglas' linkage testimony was challenged before the Delaware Supreme Court, the court disposed of the challenge in two paragraphs which essentially held that the testimony was not the product of science but experience, and, therefore the *Frye*⁴⁰² test did not apply, and further, Douglas' experience had given him specialized knowledge that could assist the trier of fact under Delaware's version of Rule 702.⁴⁰³ And that, as they say, was that. While the court declared itself strongly opposed to profile evidence, which it declared "is of little probative value and extremely prejudicial to the defendant since he is, in a sense, being accused by a witness who was not present at any of the crimes,"⁴⁰⁴ testimony regarding the "'signature' aspects of the crime" was just fine.⁴⁰⁵ There was no consideration of accuracy or validity at all.⁴⁰⁶

The *Pennell* decision set the tone for the judicial handling of "signature" or "linkage analysis" testimony until the New Jersey Supreme Court's decision on the admissibility of Hazelwood's proposed testimony in the *Fortin* case itself. In 1993, Douglas was allowed to testify to linkage in the California case of Cleophus Prince, who was charged with six murders of young women, three of which had taken place at the

⁴⁰⁰ *Id.*

⁴⁰¹ *Id.* Douglas inexplicably misspells Pennell's name throughout.

⁴⁰² *Frye v. United States*, 293 F. 1013 (D.C. 1923) (requiring proof of general acceptance in the relevant scientific community for novel scientific evidence).

⁴⁰³ *Pennell v. State*, 602 A.2d 48, 55 (Del. 1991).

⁴⁰⁴ *Id.*

⁴⁰⁵ *Id.* The court also grossly mischaracterizes *United States v. Rogers*, 769 F.2d 1418 (9th Cir. 1985), as precedent for its ruling. In *Rogers*, an FBI agent testified to a single point—the rarity of the use of bandannas as masks in armed bank robberies in Los Angeles, so his testimony did not involve "signature" or "linkage analysis." Moreover, his testimony was not objected to, which was the stated basis for the affirmation. See *id.* at 1425.

⁴⁰⁶ The subsequent unusual history of *Pennell* is worth a note. The jury convicted Pennell on the DiMauro and Ellis murders but was hung as to the Gordon case. Death was not recommended, so Pennell was sentenced to life without parole on the DiMauro and Ellis murders. While those two convictions were on appeal, Pennell was re-indicted for the Gordon murder and the murder of another woman, Kathleen Meyer, who disappeared after being seen entering Pennell's van. Pennell then demanded to represent himself *pro se* and, in that capacity, pled *nolo contendere* to the Gordon and Meyer murders. Accordingly, he was sentenced to death at his own request. Those cases were on mandatory appeal, with Pennell arguing to uphold his own death penalty, when the original appeal was heard (rendering it functionally moot). See generally *Pennell v. State*, 604 A.2d 1368 (Del. 1992). Pennell was executed in March 1992.

Buena Vista Garden Apartments in San Diego in a short period of time, and three of which had taken place somewhat later at other apartment complexes in San Diego.⁴⁰⁷ As Douglas explains his role:

The key to the case was the DNA match between semen found on the clothing of the second Buena Vista victim, twenty-one-year old Janene Weinhold, and blood and saliva samples they got from Prince. But what about the other five murders?

San Diego police asked us to reexamine the six cases to see if it was reasonable to conclude that one individual had committed all the murders. Several people, including prosecutors Dan Lamborn and Woody Clark and Sergeant Ed Petrick of the task force, came to meet with us at Quantico. If the prosecution could prove that the defendant had committed all six murders, rather than only that of Janene Weinhold, the number and nature of the crimes would qualify as "special circumstances" under California law, which would make it a capital case. They didn't want this guy getting out again?⁴⁰⁸

Another good illustration of prosecutorial pragmatism, perhaps, but not the kind of neutral and dispassionate conditions likely to yield an objective and reliable assessment.⁴⁰⁹ Douglas was allowed to testify, but we do not know the court's reasoning, since there was never an opinion on the issue published at any level of the California court system.

A similar story lies behind Hazelwood's participation in the 1995 trial of Kenneth Bogard, who was charged with the rapes of six women in San Diego. As Michaud and Hazelwood tell it:

In all, Lamborn [the prosecutor] had DNA evidence that directly implicated Bogard in only one of the assaults. Bogard vehemently denied all guilt. And after viewing Bogard in a lineup, only one of the victims, Dana Holly, was able to identify him as her attacker, and then only indirectly, by his voice.⁴¹⁰

....

There were no eyewitnesses to any of the assaults. What is more, there was no fingerprint evidence.

It was going to be an uphill prosecution

....

Lacking any solid eyewitness identifications, the prosecutor needed to tie the crimes together in a way the jury could follow. So as he prepared for trial, Lamborn contacted Hazelwood, who had just retired, and hired Roy to conduct a linkage analysis

⁴⁰⁷ See DOUGLAS & OLSHAKER, *JOURNEY INTO DARKNESS*, *supra* note 224, at 39.

⁴⁰⁸ *Id.* at 39, 65.

⁴⁰⁹ See Risinger et al., *Daubert/Kumho Implications*, *supra* note 194, at 1.

⁴¹⁰ For a discussion of the reliability problems of this kind of identification, see Lawrence M. Solan & Peter M. Tiersma, *Hearing Voices: Speaker Identification in Court* (draft submitted for publication, 2002) (on file with authors).

....

Inside superior court judge John Thompson's windowless, fluorescent-lit third-floor courtroom, Hazelwood took the stand, turned to the jury, and began to testify. Dan Lamborn remembers he had little to do except to occasionally interject a question or ask for amplification.

"Roy," says Lamborn, "was the star of the show."⁴¹¹

A familiar theme. Once again, we do not know the rationale for the admission of Hazelwood's "linkage analysis," since there are no reported decisions dealing which deal with the issue.

The post-*Pennell*, pre-*Fortin* reported decisions in other cases which have dealt with the issue are all very cursory. In *State v. Code*,⁴¹² a 1994 Louisiana decision, Douglas had been allowed at a Rule 404(b) admissibility hearing to testify to the linkage of eight charged murders (he had to share the spotlight with the local coroner, who was allowed to testify to the same thing).⁴¹³ The Supreme Court of Louisiana never even addressed the admissibility of such evidence at trial, and merely assumed its admissibility at the 404(b) hearing during its discussion of whether the three crimes are admissible under a 404(b) identity rationale.⁴¹⁴ On that issue, they seemed most impressed with the fact that the hands of all the victims were tied with a very unusual type of binding and knot. Douglas did not testify at trial.

The Oregon case of *State v. Russell*⁴¹⁵ requires a bit more exposition. Charles Russell was charged with the murders of three women, all of whom had been killed in Bellevue, Washington, within a mile or two of each other in a little over a two month period in the summer of 1990. Both John Douglas and his co-enthusiast for "signature analysis" Robert D. Keppel, testified at trial that all three of the crimes were committed by the same person, so we will let Keppel describe the problem.

[T]hey connected Pohlreich to Russell through a DNA analysis of Russell's semen that was found inside Pohlreich's body In the Beethe and Levine cases, the physical evidence was not as conclusive as in the Pohlreich case, which made the signature testimony linking all the cases that much more crucial.⁴¹⁶

The same, familiar theme again.

Actually, *Russell* is interesting because it turned both Douglas and Keppel partially into statistical profilers, at least in regard to that part of their testimony that drew the attention of the Washington Supreme Court. There is no doubt that each of the three murders involved what

⁴¹¹ MICHAUD & HAZELWOOD, *EVIL*, *supra* note 31, at 195.

⁴¹² 627 So. 2d 1373 (La. 1994).

⁴¹³ See *id.* at 1382. See the analysis of *Code* in the New Jersey intermediate appellate opinion in *State v. Fortin*, 318 N.J. Super. 577, 605-06 (N.J. Super. Ct. App. Div. 1999).

⁴¹⁴ See *Code*, 627 So. 2d at 1382-83.

⁴¹⁵ 882 P.2d 747 (Wash. 1994).

⁴¹⁶ KEPPEL, *supra* note 350, at 256.

is called in the profiling business “posing with props,” that is, arranging the body to be discovered in some sort of pose where foreign items are an integral part of the pose. And these were not just cases of the insertion of sticks into orifices, which can raise issues concerning whether such counts as posing. Each was different, but each was fairly elaborate. The first was the simplest: the body was found outside near a dumpster:

Pohlreich’s body was unclothed, but she was wearing two pieces of jewelry. There was a Frito Lay dip container lid over her right eye and forehead, her arms were folded over her stomach, her legs were extended and crossed at the ankles, and she had a pine cone in one of her hands.⁴¹⁷

The second murder seven weeks later involved an apartment break-in:

Beethe was on her back on the bed. The bedspread was pulled down to the foot of the bed. Her body was unclothed except for a pair of red high-heeled shoes. Her feet were together with legs spread and knees bent. Blood had been smeared on her legs in a manner that resembled “finger painting”. . . . A rifle had been placed resting symmetrically between Beethe’s legs, resting on her shoes. The firearm penetrated approximately five or six inches into her vagina.⁴¹⁸

Finally, the last murder three weeks later contained posing as elaborate as the second:

Levine was on her back, on the bed. Her face was turned toward her left shoulder. Her legs were spread with knees straight. Her right arm extended above her shoulder while her left arm rested by her side. Under Levine’s left forearm was the book *More Joy of Sex*. A plastic dildo was partially inserted into Levine’s mouth.⁴¹⁹

At trial both Douglas and Keppel testified to the general signature material.⁴²⁰ However, their most powerful argument concerned the likelihood of more than one killer who indulged in posing with props operating in such a small area in a sixty-seven day period.⁴²¹ This, in turn, depended on the commonness of “posing with props” as an element of sexual murders. Keppel testified not merely to his impression from experience, but to the results of a search of Washington’s Homicide Information and Tracking System (“HITS”) database, which he said showed that posing was rare.⁴²² Douglas testified similarly regarding

⁴¹⁷ Russell, 882 P.2d at 756.

⁴¹⁸ *Id.* at 757 (citation omitted).

⁴¹⁹ *Id.* at 758.

⁴²⁰ See KEPPEL, *supra* note 350, at 258; DOUGLAS & OLSHAKER, MINDHUNTER, *supra* note 216, at 256. Keppel’s extensive account of the case can be found in Robert D. Keppel, *Signature Murders: A Report of Several Related Cases*, 40 J. FORENSIC SCI. 670 (1994), and in *State v. Russell*, 882 P.2d 747 (Wash. 1994).

⁴²¹ See Russell, 882 P.2d at 776.

⁴²² See *id.* at 777.

the results of an FBI Violent Criminal Apprehension Program (“VICAP”) search.⁴²³ The Washington Supreme Court concentrated mostly on the propriety of this use of databases (which it accepted).⁴²⁴ The threshold dependability of the “signature” testimony was disposed of even more summarily than in *Pennell*, and in much the same terms.⁴²⁵ So that was where the courts stood when the prosecutor in *Fortin* picked up the phone to dial the Academy Group.

V. FORTIN IN THE COURTS

When Roy Hazelwood was contacted by the prosecutor in *Fortin*, he set out to do a “linkage analysis” *a la* Douglas.⁴²⁶ The process undertaken by Hazelwood was perhaps the ultimate example of a radically non-blind process, with all that that entails in terms of biasing results.⁴²⁷ Hazelwood knew from the outset what was desired and set about to see if he could deliver it. He also knew from the information that was irrelevant to his asserted function and claimed expertise which he examined,⁴²⁸ such as the report of the forensic odontologist, that there was evidence pointing to Fortin’s guilt independent of the details of the two crimes he was to examine for linkage.⁴²⁹ He then proceeded to compare the two cases, find the correspondences, list them, and then declare that he was confident based on the similarities that the same person had committed both crimes. If this appears to partake of the “Two Room” process, this appearance will not be dispelled by his list of significant correspondences. You will also recognize the framework of Hazelwood’s analysis as straight Douglas,⁴³⁰ for he writes as follows:

⁴²³ See *id.* at 778.

⁴²⁴ See *id.* at 779.

⁴²⁵ See *id.* at 775.

⁴²⁶ As previously noted, Hazelwood testified that the Douglas work was the first published exposition of the theory that he follows. See *State v. Fortin*, 318 N.J. Super. 577, 591 (N.J. Super. Ct. App. Div. 1999).

⁴²⁷ See generally Risinger et al., *Daubert/Kumho Implications*, *supra* note 194.

⁴²⁸ Simply importing the conclusions of another expert to be rubberstamped with your own ostensibly independent conclusion is not part of the theory of “linkage analysis” or “signature” and forms no part of the claim of expertise made by Hazelwood to justify his conclusions in the case. For the central importance to valid results of shielding experts from non domain-specific information, see *supra* at 27-30.

⁴²⁹ Hazelwood’s report makes clear that he had the forensic odontologist’s report and all of the other case information prior to construction of his “linkage analysis.” See Robert R. Hazelwood, Report to Middlesex (N.J.) Count Prosecutor, Oct. 17, 1997, at 2 [hereinafter *Hazelwood Report*].

⁴³⁰ The only thing Hazelwood ever wrote on linkage analysis, outside of his memoir writings, is three paragraphs in a book he co-authored with Janet Warren. See Robert R. Hazelwood & Janet I. Warren, *The Relevance of Fantasy in Serial Sexual Crime Investigations*, in PRACTICAL ASPECTS OF RAPE INVESTIGATION: A MULTIDISCIPLINARY APPROACH 134-35 (Robert R. Hazelwood & Ann Wolbert Burgess eds., 2d ed. 1995). The paragraphs describe the Douglas theory, with the addition of Hazelwood’s three purposes of M.O.—to protect identity, to ensure success, and to facilitate escape. See *id.*

When examining crimes for linkage, one must study the offender's behavior for similarities over the crimes. This behavior is referred to as "M.O." (modus operandi) and "Ritualistic" ("Signature") behavior. The M.O. is learned behavior and is developed by the criminal to accomplish three things: Ensure success; Protect identity and; Facilitate escape. Because it is learned behavior, the M.O. is in a constant state of evolution which allows it to meet the demands of the crimes. Therefore the M.O. is subject to change over time and the primary causes of such change are: Experience; Maturity; and Education. It is my opinion that the M.O. of the crimes involving Ms. Padilla and Ms. Gardner demonstrate the following similarities:

<u>Melissa Padilla</u>	<u>Vicki Gardner</u>
High-risk crime	High-risk crime
Crime committed impulsively	Crime committed impulsively
Female victim	Female victim
25 years-old	34 years-old
Victim crossed path of offender	Victim cross[ed] path of offender
Victim was alone	Victim was alone
Assault at confrontation point	Assault at confrontation point
Adjacent to well-traveled roadway	On well-traveled roadway
Occurred during darkness (11:30p.m)	Occurred during darkness (8:40p.m)
No weapons involved in assault	No weapons involved in assault
Blunt force (fists) injuries	Blunt force (fists) injuries
Trauma primarily to upper face with no damage to teeth	Trauma primarily to upper face with no damage to teeth
Lower garments totally removed	Lower garments totally removed
Shirt left on victim and breasts free	Shirt left on victim and breasts free
No seminal fluid found on/in victim	No seminal fluid found on/in victim
No theft of valuables	No theft of valuables.

The violent offender who repeats his offenses typically demonstrates a second type of behavior and that is termed "Ritualistic" behavior. Such behavior is frequently referred to as the "Signature" of a crimi-

nal. This behavior goes beyond what is necessary to commit the crime. Its sole purpose is to provide the offender with mental and/or emotional gratification. The “Ritualistic” aspects of a crime remain constant over time, although there may be improvements as the ritual becomes more fully developed. It is my opinion that both of the crimes were anger-motivated and that the offender demonstrated that anger through the following Ritualistic or “Signature” behaviors:

Melissa Padilla	Vicki Gardner
Bite to lower left chin	Bite to lower left chin
Bite to lateral left breast	Bite to lateral left breast
Injurious anal penetration	Injurious anal penetration
Brutal facial beating	Brutal facial beating
Manual (frontal) strangulation	Manual (frontal) strangulation ⁴³¹

Hazelwood then concludes:

In my 35 years of experience with a variety of violent crimes committed in the U.S., Europe, Canada, and the Caribbean, I have never observed this combination of behaviors in a single crime of violence. The likelihood of different offenders committing two such extremely unique crimes is highly improbable. Based upon a comparison of the M.O. and the Ritualistic behaviors of the two crimes, it is my opinion that the same person was responsible for the murder of Ms. Melissa Padilla and the subsequent attempted murder of Ms. Vicki Gardner.⁴³²

The above manifests indeterminacy, accuracy, base-rate, Monty Hall and Two Rooms problems, and a misleading form of expression to boot. We will address each problem in turn.

The indeterminacy problems spring from the vague descriptors used in some of the items, particularly packaging a Maine rural interstate and an urban artery together under the label “well traveled roadway,” the times of the two assaults under the label “darkness” the implied assertion that the ages of the victims, 25 and 34, are significantly similar in some way and the characterization of both incidents as being “high risk crimes.” As to accuracy problems, there is insufficient known detail regarding the Padilla murder to conclude that no weapon was at least displayed, or that the perpetrator had not stalked Padilla for

⁴³¹ *Hazelwood Report*, *supra* note 429, at 3-5.

⁴³² *Id.* at 5.

a short or long period, unless one assumes the conclusion that Fortin committed the crime, and that it happened consistently with the Maine crime. Finally, as to “no theft of valuables,” there *was* in fact theft in the Padilla case.

The Monty Hall problems are of the subtype represented by the well known case of *People v. Collins*,⁴³³ where variables are held out as presenting largely independent information which, in fact, is entailed to a large degree in other variables listed, either alone or in combination. This problem may be seen in the first two characteristics listed: “crime committed impulsively” and “high risk crime.” One need not have 35 years of law enforcement experience to be pretty confident that the set of crimes committed on impulse is a set richer in crimes with higher risks of interruption and apprehension than the set of crimes committed with reflection and planning. Or consider the “female victim” item. This information is largely entailed in the “lower garments totally removed” and “shirt left on victim and breasts free” items listed later. “Assault at confrontation point,” “trauma primarily to upper face” and “no weapons involved in assault” incorporate “blunt force (fist) injuries” and most of “brutal facial beating.”

Note that one ought not necessarily be critical of such a list generated to describe a crime for investigatory purposes. Such a listing, even of overlapping variables, could be helpful because each slightly different way of expressing the details might stimulate some helpful reflection or suggest some potentially fruitful investigatory action. However, in the context of claiming similarities between two crimes, such a list, intentionally or unintentionally, merely gives an illusion of correspondence greater than that which actually exists. And, of course, this illusion is compounded terribly by “Two Room problem” considerations.

Recall the central lesson of the Two Room problem. In any two human episodes of even moderate informational complexity, a fairly large number of apparently surprising paired correspondences of fairly low base-rate occurrences can be found by *post hoc* dredging. This can be controlled methodologically by identifying the variables believed to be significant in the first crime before knowing anything about the second crime (an analogue to the statistical and research design principle that says you must identify test variables in advance of running the test). However, nothing like this was done in this case, or indeed is ever formally done so far as we know. The second best (by a large margin) approach to such a problem is to limit analysis to variables one believes

⁴³³ 438 P.2d 33 (Cal. 1968). *Collins* is a famous case primarily because it illustrates so many of the ways in which the forms of formal statistical proof can be abused. However, one of the abuses was not limited to formal statistical proof. The prosecutor in the case dredged the correspondences between Collins and the witnesses' accounts of the perpetrator in the case, and many of the correspondences to which the prosecutor asked the statistician witness to apply the products rule clearly were not independent variables.

one would have concluded were in some way diagnostic of perpetrator similarity in advance of knowing that there even was a second crime. In this case, it is unlikely that Hazelwood or anyone else would have pointed to most of the variables on the list as identity related without the benefit of data dredging. Take “no weapons involved in assault.” It would have been highly risky to claim that this meant much after only examining the Padilla case, because it would run the risk that if a stick was used with fists in Maine the claimed similarity could blow up. The same applies for “bite to lower left chin” in that it runs the risk of no bite, no bite to chin, or no bite to face or no bite above the neck. Same for—well, most of the variables as they are set out.

This problem is substantially compounded by Hazelwood’s form of expression that, in his 35 years of experience, he had never seen “this particular combination of behaviors in a crime of violence.” This is essentially a trick phrasing meant to increase the appearance of rarity without actually telling a formal untruth. In any list of characteristics that contains one unique item, the witness can say that he has never seen that particular combination even though all of the other items are common and commonly observed together. In the Padilla killing, that seems to be true of everything on Hazelwood’s list with the exception of a bite to the chin, as a reference to the description of Disorganized Sexual Homicide in the *Crime Classification Manual* indicates.⁴³⁴

⁴³⁴ See DOUGLAS ET AL., CRIME CLASSIFICATION MANUAL § 132, *supra* note 384. Lest the reader doubt this, it is necessary to quote at length from the description in the *Crime Classification Manual* under section 129-130: *Disorganized Sexual Homicide*:

[T]he victim is often from his own geographic area because this offender acts impulsively under stress and also because he derives confidence from familiar surroundings. . . . The risk factor of a disorganized sexual homicide victim is situational in the sense that by crossing the path of the offender, the risk of becoming a victim is greatly elevated. The victim essentially becomes a casualty because he or she was in the wrong place at the wrong time. . . . The crime scene of a disorganized sexual homicide reflects the spontaneous and, in some cases, symbolic quality of the killing. It is random and sloppy with great disarray. The death scene and the crime scene are often the same.

The victim/location is known because it usually is where he or she was going about daily activities when suddenly attacked. There is evidence of sudden violence to the victim, a blitz style of attack. Depersonalization may be present. . . .

There is no set plan of action by the offender for deterring detection. The weapon is one of opportunity. . . . There is little or no effort to remove evidence. . . . The body is left at the death scene, often in the position in which the victim was killed. There is no attempt or only minimal attempt to conceal the body.

. . . .

Another example of the disorganized offender’s personation of his ritual sexual fantasies is excessive mutilation of the breasts, genitals, or other areas of sexual association (thighs, abdomen, buttocks, and neck). . . .

. . . .

There may be depersonalization that entails mutilation to the face and overkill (excessive amount or severity of wounds or injury) to specific body parts. The face, genitals, and breasts are most often targeted for overkill. . . .

The blitz style of attack common to this homicide is often manifested by focused blunt trauma to the head and face and by the lack of defensive wounds. . . .

We have already observed that humans in general do not do well at accurately assessing the meaning of dredged correspondences. This can only be exacerbated when such a presentation is made through claimed expertise. When faced with the problem in the *Fortin* case itself, how have the courts to whom the case has been presented handled it?

The issue of the admissibility of the Maine case under the *modus operandi* rationale, and also the admissibility of Hazelwood's proposed testimony, was raised by the defense by motion *in limine*. The trial court ruled that the details of the Maine case were sufficiently relevant to be admissible on the issue of Fortin's identity as Padilla's killer. As to Hazelwood's proposed testimony, the trial court simply accepted everything Hazelwood claimed, on the authority of his credentials, and the decisions in *Pennell*, *Code* and *Bogard*.⁴³⁵

The Appellate Division (New Jersey's intermediate appellate court) adopted a different tack. Without much analysis beyond pointing out that the determination was to be reviewed by an abuse of discretion standard, it affirmed the trial court's ruling on the admissibility of the Maine episode.⁴³⁶ It then undertook a lengthy exposition of the claimed principles of profiling and linkage analysis, which it appeared to accept at face value, though it was clearly troubled by the impact the testimony would have on the jury if it were admitted.⁴³⁷ However, having found a quotation from Hazelwood, which the court took to say that proper determination of signature characteristics required at least three crimes, it took that at face value also and, pointing out that the *Fortin* case involved only two crimes, ruled Hazelwood's testimony inadmissible.⁴³⁸

The New Jersey Supreme Court then spoke,⁴³⁹ in an opinion that manifests a kind of split personality altogether too common when courts deal with issues of expert reliability in criminal cases. After setting out the facts and the issues in Part I of the opinion, in Part II of the opinion the court takes up the issue of the reliability of linkage analysis, concluding that "the proposed expert testimony of Hazelwood concerning linkage analysis lacks sufficient scientific reliability to establish that the

Sexual acts occur after the victim's death and often involve insertions of foreign objects into body orifices (insertional necrophilia). This is often combined with acts of mutilation (e.g., slashing, stabbing, and biting of the buttocks and breasts). Because these acts often do not coincide with completed acts of sexual penetration, evidence of semen may be found on the victim's clothing and (less frequently) in the victim's wounds. Most often death results from asphyxia, strangulation, blunt force, or the use of a pointed, sharp instrument.

Id. This description sounds pretty much like the death of Melissa Padilla. doesn't it? Most of the items on Hazelwood's list are class characteristics by his own profession's standard description.

⁴³⁵ Letter Opinion of Judge Barnett E. Hoffman, New Jersey Superior Court, Law Division, June 3, 1998, at 3-5.

⁴³⁶ See *State v. Fortin*, 318 N.J. Super. 577, 595 (N.J. Super. Ct. App. Div. 1999).

⁴³⁷ See *id.* at 601-09.

⁴³⁸ See *id.* at 609-10.

⁴³⁹ See *State v. Fortin*, 745 A.2d 509 (N.J. 2000).

same perpetrator committed the Maine and New Jersey crimes.”⁴⁴⁰ In support of that conclusion, the opinion of the court contrasts the claims of near perfection for the technique found in Hazelwood’s 1998 memoir with the dearth of data in the literature.⁴⁴¹

At this point, one would think that Hazelwood was not going to be able to testify at trial. However, in Part III of the opinion, the court begins, rather naively:

In all fairness, Hazelwood did not purport to cloak his testimony with a mantra of scientific reliability. He candidly acknowledged that linkage analysis is not a science, but rather is based on years of training, education, research, and experience in working on thousands of violent crimes over an extended period of time.⁴⁴²

What the court seems to be unaware of, as it bestows its compliments on Mr. Hazelwood, is the history of forensic science claims in federal court after the decisions in *Daubert v. Merrell Dow Pharmaceuticals*,⁴⁴³ and *United States v. Starzecpyzel*.⁴⁴⁴ *Daubert* directed the federal courts to determine threshold reliability for proffered expertise under Federal Rule of Evidence 702. Because *Daubert* itself dealt with the frank products of science, some courts limited its requirements to scientific evidence. In *U.S. v. Starzecpyzel*, an attack was mounted upon the “forensic science” of handwriting identification. In a well known opinion, Judge McKenna first eviscerated the claims of handwriting expertise as a science, saying at one point that if it had to meet the *Daubert* standards it would have to be excluded.⁴⁴⁵ Having said this, however, he concluded that since it was not a science, it did not have to meet the *Daubert* standards, and allowed the testimony. After *Starzecpyzel*, there was a stampede among former practitioners of “forensic science” to repackage their testimony as not based on “science” but on “experience.” In that way, they managed to avoid substantial scrutiny until the Supreme Court closed the door on this loophole in *Kumho Tire v. Carmichael*.⁴⁴⁶ Even before *Daubert*, eschewing scientific status helped “linkage experts” avoid the *Frye* test, as the court’s opinions in *Pennell v. State* and *State v. Russell* show. Hazelwood deserved no gold stars for frankness in his embrace of the mantle of experience-based expertise.

Nevertheless, apparently dazzled by his frankness, the court continues: “Such methods have great value for purposes of criminal inves-

⁴⁴⁰ *Id.* at 513.

⁴⁴¹ *See id.* at 514.

⁴⁴² *Id.* at 515.

⁴⁴³ 509 U.S. 579 (1993).

⁴⁴⁴ 880 F. Supp. 1027 (S.D.N.Y. 1995).

⁴⁴⁵ *See id.* at 1036.

⁴⁴⁶ 526 U.S. 137 (1999) (holding that *Daubert*’s gatekeeping requirements apply to all proffered expertise).

tigation.”⁴⁴⁷ No serious controversy here, if such methods are limited to providing investigative leads. But the court goes on:

We therefore believe that one such as Hazelwood has a proper role in a criminal trial based on his experience as an expert in criminal investigative techniques. Such a witness is qualified to discuss similarities between crimes without drawing conclusions about the guilt or innocence of the defendant. Within that ambit, his testimony can be of assistance to the court and perhaps a jury on the issue of admission of other-crime evidence. Of course, Hazelwood would not be permitted to testify on the ultimate issue of whether the person that assaulted Trooper Gardner is the same person that murdered Melissa Padilla.⁴⁴⁸

What “proper role in a criminal trial” does the court envision? Apparently explaining things to the court about similarities and differences between crimes in a Rule 104 hearing concerning admissibility of the supposedly similar crimes, as was done in *State v. Code*.⁴⁴⁹ “And perhaps a jury . . . ?” Well, which is it? Does Hazelwood get to testify to his views on similarities in front of a jury, or not? The court does not actually say. However, wherever he gets to testify, he can’t give an opinion that the same person who assaulted Trooper Gardner murdered Melissa Padilla, which the court characterizes as an “ultimate issue.”⁴⁵⁰

In Part IV of the opinion, the court goes over the rhetoric of the standards of similarity required to render an uncharged crime admissible on an identity or signature theory, concluding:

In order for evidence of a prior crime to be admissible on the issue of identity, “the prior criminal activity with which defendant is identified must be so nearly identical in method as to earmark the crime as defendant’s handiwork. The conduct in question must be unusual and distinctive so as to be like a signature, and there must be proof of sufficient facts in both crimes to establish an unusual pattern.”⁴⁵¹

The court then proceeds to Part V of the opinion, which it opens thus:

To state the law, however, is easier than to apply the law. The meaning of such words is not self-revealing. We are not so certain that the M.O. factors cited by Hazelwood, such as that both victims were mature females and were attacked while alone and at night time, demonstrate an “unusual pattern.” (Defendant argues that there

⁴⁴⁷ *State v. Fortin*, 745 A.2d 509, 515 (N.J. 2000).

⁴⁴⁸ *Id.*

⁴⁴⁹ See *supra* note 31 and accompanying text.

⁴⁵⁰ This characterization should not, by itself, make a difference, given the explicit terms of New Jersey Rule of Evidence 704, which holds that “testimony in the form of an opinion or inference otherwise admissible is not objectionable because it embraces an ultimate issue to be decided by the trier of fact.” N.J. R. EVID. 704.

⁴⁵¹ *Fortin*, 745 A.2d at 517 (quoting *State v. Reldan*, 185 N.J. Super. 494, 502 (N.J. Super. Ct. App. Div. 1982)).

are sixteen differences between the crimes.)⁴⁵²

At this point, one might think things looked bad for Hazelwood on remand, at least as to his testimony regarding the M.O. factors on his list, or at least the one's just recited skeptically by the court. But don't be too quick to judge, as the court continues: "It is on this question of an 'unusual pattern' that the testimony of Hazelwood would be helpful."⁴⁵³ This is a curious statement since the court just said that many of Hazelwood's variables were not too convincing. Well, the court continues:

For example, if the witness can from a reliable data base offer evidence that a combination of bite marks on the breast, bite marks on the chin, and rectal tearing inflicted during a sexual attack is unique in his experience of investigating sexual assault crimes, that evidence could help to establish an "unusual pattern." Such expert testimony would help a court make an initial determination of whether to admit the other-crime evidence and would, if presented at trial, better enable a jury to understand whether the crimes were "unusual and distinctive so as to be like a signature" such that an inference could be drawn to "earmark the crimes as the handiwork of the same person."⁴⁵⁴

Does such testimony, in the courts view, have to be based on a reliable objective database available to the defendant for review (as was perhaps in part the case in *Russell*) or, since the witness is merely going to testify that the combination of factors are "unique in his experience of investigating sexual assault crimes," is a private, experience-based "data base" not subject to review sufficient. Though the opinion seems to emphasize the desirability of an objective and reviewable data compilation, in the end it is ambiguous.

The court continues:

It is initially for the court, and ultimately for the jury, however, to determine whether that inference concerning the ultimate issue of guilt may be drawn. In point of fact, the trial court did incorporate Hazelwood's testimony in its 404(b) ruling, stating that Hazelwood's testimony was persuasive in that Hazelwood had not seen in reviewing 4000 cases this combination of bite marks, anal tears, and brutal facial beatings to a victim. If there is such a database of cases, the witness' premise can be fairly tested and the use of the testimony invokes none of the concerns that we have expressed about the improper use of expert testimony.⁴⁵⁵

Of course, there is no such database by reference to which the witness' premises can be fairly tested. Hazelwood was referring to his subjective experience "database." Does that put him out of court? One might think so at this point, and the court goes on as follows: "We are

⁴⁵² *Id.*

⁴⁵³ *Id.*

⁴⁵⁴ *Id.* at 517-18 (citations omitted).

⁴⁵⁵ *Id.* at 518.

especially concerned about the use of expert testimony ‘to interpret matters that could be considered commonplace or conduct that could be accounted for commonsensically.’ Our concern is that a factfinder’s ‘uncritical acceptance of expert testimony can becloud the issues.’”⁴⁵⁶ Looking bad for Hazelwood on remand, right? Not so fast. The court continues:

We have no sense that Hazelwood’s suggestions are counterintuitive⁴⁵⁷ or will receive uncritical acceptance. Stripped of its scientific mantra, the testimony is nothing more than a description of the physical circumstances present, somewhat similar to the description of the knots used to tie the victims in *State v. Code*. We allowed similar testimony in *Zola*, when the testimony involved “common sense” deductions on subjects about which jurors may not have much familiarity, and such testimony did not infringe on the jury’s capacity to determine the ultimate fact at issue. Arguably at least, the questions here do not relate “to a subject-matter beyond the understanding of persons of ordinary experience, intelligence, and knowledge.” Still we doubt that most jurors will have much familiarity with the pattern of injuries inflicted in rape cases

. . . .

. . . [With proper limiting instructions on the use of the other crime evidence for identification only] we believe that a crime-scene analyst such as Hazelwood, with broad experience in investigating similar crimes, can assist a court and a jury in understanding whether the crimes bear such a unique signature that an inference may be drawn that the perpetrator of the two crimes was the same person.⁴⁵⁸

In dissent, Justice Long said in part:

[I]t is my view that the reliability defects that, according to the Appellate Division and the majority, preclude Hazelwood from testifying as a scientific expert on linkage, are equally applicable to his proffer of uniqueness testimony. Linkage analysis is the procedure used by criminal investigators when the concentration of modus operandi and ritualistic characteristics in crimes is high, such that the investigator can conclude that the perpetrator is the same person. Uniqueness testimony is linkage analysis under another name. It is no more reliable when Hazelwood testifies as a crime investigator than when he does so as an “expert” in ritualistic behavior.⁴⁵⁹

On remand, at trial, Hazelwood was allowed to testify to virtually everything he had originally proposed to testify about: BSU credentials, research credentials, publications, “six step methodology,” MO, ritual

⁴⁵⁶ *Id.* (quoting *State v. Zola*, 112 N.J. 384, 415 (1988)).

⁴⁵⁷ What about the suggestions that the M.O. list constituted an unusual pattern, which the court seemed to find counterintuitive earlier in the same section?

⁴⁵⁸ *Fortin*, 745 A.2d at 518-19 (citations omitted). We are somewhat perplexed as to what the court meant by “[s]tripped of its scientific mantra.” See *id.* at 518. We thought the court had previously said that linkage analysis made no claim to scientific status.

⁴⁵⁹ *Id.* at 523 (Long, J., dissenting).

(his term for “signature”), psychodynamic theory, and all the items on his lists but “no theft of valuables” (since he had by then been informed that a locket had been taken from Padilla). The single mandated exception was the explicit final conclusion that whoever committed the Maine assault committed the Padilla homicide.⁴⁶⁰ Fortin was convicted and sentenced to death.

So what’s wrong with that? Well, it does seem somewhat inconsistent with the narrow thrust of the opinion of the New Jersey Supreme Court, as unclear as it was. It seemed more as if the New Jersey Supreme Court were envisioning an educational expert role for Hazelwood, not the role of a *Hines* witness.

Some exposition of this distinction is undoubtedly necessary. An educational expert is one who is not called to translate the meaning of other information before the jury.⁴⁶¹ The expert function we most often think of is this translational function, where the witness claims to have a method of taking information available to both the witness and the jury, and accurately translate it better than the jury into some non-obvious other fact.⁴⁶² That is, in fact, what “linkage” witnesses claim to be able to do, to take the details of a crime available to both jury and expert, and translate them better than the jury into the conclusion “same perpetrator.” This would also seem to be what the New Jersey Supreme Court did not think linkage witnesses had shown they could do better than juries. However, because of their experience, these investigators of violent crime do know some case-relevant facts juries *don’t* know, such as the relative commonness or rarity of certain characteristics for types of crime with which juries are not familiar. It is perhaps appropriate that they be allowed to testify to those facts, to summarize their experience and educate the jury concerning relevant facts known to them, as long as their testimony is controlled and limited to that educational role.⁴⁶³ In this role, the expert actually does not address inferences in regard to a particular case, but merely provides so called “major premise” information about the world outside of the case.⁴⁶⁴ Such frank “summarizational” or “educational” experts have become increasingly common,⁴⁶⁵ and this appears to be the role the New Jersey Supreme Court had in mind for Hazelwood in *Fortin*. Which emphatically was not the role he played at trial on remand. Instead, Hazelwood was allowed to testify as

⁴⁶⁰ See TT, Nov. 9, 2000, at 58-96 (direct testimony of Robert R. Hazelwood).

⁴⁶¹ See Risinger, *Functional Taxonomy*, *supra* note 265, at 520.

⁴⁶² See *id.* at 520-21.

⁴⁶³ For observations on how difficult it sometimes is to control police witnesses playing this role, see *id.* at 516 n.16.

⁴⁶⁴ See *id.* at 511-18. See generally Ronald J. Allen & Joseph S. Miller, *The Expert as Educator: Enhancing the Rationality of Verdicts in Child Sex Abuse Prosecutions*, 1 PSYCHOL. PUB. POL’Y & L. 323 (1995); Ronald J. Allen & Joseph S. Miller, *The Common Law Theory of Experts: Deference or Education?*, 87 NW. U. L. REV. 1131 (1993).

⁴⁶⁵ See Risinger, *Navigating Expert Reliability*, *supra* note 8.

a *Hines* witness, or even beyond.

A “*Hines* witness” is named for the recent case of *United States v. Hines*.⁴⁶⁶ That case involved a challenge to the validity of handwriting identification expertise pursuant to *Daubert*.⁴⁶⁷ The court found that such expertise was not well validated, but instead of excluding the government’s proffered witness, the judge ruled that the witness could testify to, and point out to the jury, similarities between the questioned document and the defendant’s handwriting but could not give his opinion that the defendant wrote the document, an approach which the judge herself labeled a “compromise solution.”⁴⁶⁸ The *Hines* approach raises the spectre of unreliable witnesses being allowed to parade their credentials, point out similarities, and by obvious implication, leave the jury perfectly clear about what their conclusion in fact is. And this is the role Hazelwood was allowed to play on remand, in about as extreme a manner as possible.⁴⁶⁹ Whether the New Jersey Supreme Court will decide that this is what they meant must await events.

CONCLUSION

It may surprise the reader to find that even we believe Hazelwood, or a witness like him, might be given a legitimate role in the trial, but only on the shortest and most carefully constructed judicial leash designed to eliminate or substantially reduce the dangers of his testimony.

And from whence spring those dangers? From his unjustified shamanistic, unrealistically accurate image, carefully fostered by himself, the FBI as an institution, and the popular media, which is virtually certain to be shared by most if not all of the jury, and from his non-blind role as a person constructing evidence with an eye to “putting the prosecution over the top” in an area so subject to the Monty Hall/Two Rooms problems, a role fully manifest already in the Fortin case itself by his

⁴⁶⁶ In *United States v. Hines*, 55 F. Supp. 2d 62, 70 (Mass. Dist. Ct. 1999), Judge Gertner adopted the approach first suggested by Judge Matsch in an unreported opinion in Pre-Trial Transcript, *United States v. McVeigh*, No. 96-CR-68, 1997 WL 47724, at *1 (D. Colo. Feb. 5, 1997) (the Oklahoma City bombing trial). For a more detailed discussion of this approach, see D. Michael Risinger, *Defining the “Task at Hand”: Non-Science Forensic Science after Kumho Tire Co. v. Carmichael*, 57 WASH. & LEE L. REV. 767, 793-95 (2000) [hereinafter Risinger, *Defining the “Task at Hand”*].

⁴⁶⁷ See the extended discussion of *Hines* in Risinger, *Defining the “Task at Hand,”* *supra* note 467, at 793-95.

⁴⁶⁸ *Id.* at 795.

⁴⁶⁹ See *id.* at 795. The *Hines* approach has become fashionable, having been adopted in a number of courts, including *United States v. Santillan*, 1999 WL 1201765 (N.D. Cal. 1999), *United States v. Brown*, No. 99-184, slip op. (C.D. Cal. 1999) (on file with authors), *United States v. Rutherford*, 104 F. Supp. 2d 1190 (D. Neb. 2000), *United States v. Hernandez*, 2002 WL 1335595 (10th Cir. 2002), *United States v. Van Wick*, 83 Supp.2d 515 (D.N.J. 2001), and, most recently, by Judge Pollok in the famous fingerprint reliability case, *United States v. Llera-Plaza*, 2002 WL 32697 (E.D. Pa. 2002).

willingness to generate the “*People v. Collins* list” he has already generated. Under these circumstances, Justice Long’s position has much to recommend it, and perhaps Hazelwood himself is too tainted by circumstance to be respectably allowed to testify in this particular case. But if from the beginning everyone knew that, in every jurisdiction, the rules were:

1. That unless the witness made reference to some objectively maintained database for estimates of the rarity of characteristics of the crimes claimed to be linked in this case, such a witness could only testify as a subjective experience-based educational witness;

2. That such a witness would not be allowed to reveal his “profiler credentials” to the jury beyond saying that he had worked for many years for the FBI (or other organization) as a specialist in the investigation of sexually driven crimes like rape and sexual homicide, and that in the course of his career, both through research and through involvement in actual cases, he or she had seen the details of many cases;

3. Such a witness would only testify in regard to characteristics which in his experience were truly rare in the type of crime involved. In the case of the disorganized blitz attacks involved in *Fortin*, this would seem to be limited to the facial or chin biting, since the literature indicates that the biting of breasts and insertion of objects in the anus is relatively common in this kind of attack;⁴⁷⁰

4. In general the witness’s means of expression should make it clear that any assertion of rarity makes no claim to statistical precision; and

5. Most importantly, conclusions must be the result of a controlled and masked process by which the witness is presented with no information not relevant to the claimed experiential expertise (such as the forensic odontologists report in *Fortin*). Any episodes presented for the witness’s consideration must be arranged so that the witness must look at the charged crime first, and while looking only at the details of the charged crime, identify whatever characteristics are in fact unusual for that type of crime. Only when the witness has committed to this list should he or she be allowed to see assertedly related crimes.

Information derived in this manner and presented to the jury with these limitations would actually add specialized knowledge of base-rates in a way that might be reliable and helpful. But the way Hazelwood was actually allowed to testify in *Fortin* is more a mystical ritual not calculated to enhance reliability or accuracy of result, but merely “to put the prosecution over the top.” *Fortin* may very well have murdered Mellissa Padilla, and he may deserve his cell on Death Row. But, for the law of evidence, that is not the point.

⁴⁷⁰ See *supra* note 267.