The Probative Values and Pitfalls of Drug Courier Profiles as Probabilistic Evidence

Tung Yin*

I. Introduction

Suppose that Jones arrives at the Los Angeles airport in a taxicab. He pays entirely in cash for a one-way ticket to Detroit, from where he arrived earlier in the day. He does not check any luggage, opting to carry two shoulder bags with him. Before he boards the plane, Drug Enforcement Agency operatives detain him for an investigative stop,[1] suspecting that he is a narcotics courier because he fits a "so-called 'drug courier profile,' a somewhat informal compilation of characteristics believed to be typical of persons unlawfully carrying narcotics."[2] In Jones' case, the suspicious factors are that he: (1) is traveling from a "source city" where narcotics are known to be distributed; (2) paid entirely in cash for his ticket; (3) carried no luggage beyond his shoulder bags; and (4) stayed very briefly in Los Angeles despite having made a long flight from Detroit.[3]

While Jones is detained, the federal agents obtain a warrant to search Jones's bags, and sure enough, he is carrying controlled substances in an amount far greater than would be consistent with personal use. He is indicted for possession of controlled substances with intent to distribute.[4] If Jones brings a motion to suppress on the ground that the DEA agents lacked reasonable suspicion to detain him, should the prosecution be allowed to argue that Jones fit the drug courier profile? If so, should the prosecution be able to introduce testimony

* Associate, Munger, Tolles & Olson LLP. J.D., 1995, University of California, Berkeley (Boalt Hall); B.S., 1988, California Institute of Technology. e-mail: <yint@mto.com> Thanks to Elizabeth Clark, Michael Isgur, Kevin J. Heller, Mark J. Kadish, Joseph D. Lee, Amy C. Liu, and Jan Vetter. The views expressed in this article are entirely mine and should not be imputed to anyone else, including Munger Tolles.

1. *See generally* Terry v. Ohio, 392 U.S. 1, 27 (1968) (allowing a search for weapons during questioning based on reasonable suspicion); United States v. Sokolow, 490 U.S. 1, 10 & n.6 (1989) (allowing a search for narcotics at an airport based on reasonable suspicion).

2. Reid v. Georgia, 448 U.S. 438, 440 (1980) (per curiam); *see generally* Mark J. Kadish, *The Drug Courier Profile: In Planes, Trains, and Automobiles; and Now in the Jury Box*, 46 AM. U.L. REV. 747 (1997).

3. *Sokolow*, 490 U.S. at 8-9 (paying for a ticket in cash and brief stay); Kadish, *supra* note 2, at 748 (traveling to a source city—Los Angeles—and checking no luggage).

4. 21 U.S.C. § 841(a)(1).

regarding the profile in Jones's trial in an attempt to show that Jones, being a courier, was aware of the presence of the narcotics in his bags?[5]

Under current law, the first answer is "Yes." With regard to the second question, there is a split in authority, with the majority of courts answering "No," but a minority answering "Yes."

Unfortunately, the courts in the majority have not articulated a detailed analysis of why the drug courier profile can be used to justify probable cause/reasonable suspicion, but is inadmissible as substantive evidence except in limited circumstances. They have generally provided minimalist analyses, concluding that profile evidence is unfairly prejudicial and only barely probative.

In this article, I suggest that this distinction between the use of profiles as investigatory tools versus substantive evidence makes legal and mathematical sense and that the majority of courts have it right, though for different and more detailed reasons. The drug courier profile is essentially probabilistic evidence: if a suspect fits the profile, the probability that the person is a drug courier is significantly higher than if the person had simply been chosen from the sample population at random. Although it may seem questionable to use historical patterns to predict present criminal behavior, examples from modern business practices suggest that past actions can indeed have predictive value.

The inquiry does not end there, however. That the drug courier profile may have predictive value only justifies its use as a law enforcement tool. As with any diagnostic tool, the profile may falsely identify some innocent travelers as couriers and may fail to identify other couriers. The first kind of misidentification, though burdensome to the traveler who is erroneously stopped, is not particularly troublesome to society at the investigatory stage.[6] Any potential inaccuracies can be tolerated so long as the profile is relatively effective. At the trial stage, on the other hand, society's values reflect a judgment that the false identification of an innocent person is a grave injustice. As a result, a trial may demand a higher quality of evidence than an investigation does.

Admittedly, many evidentiary problems at trial consist of determining whether a particular challenge to a piece of evidence goes to its admissibility or its weight,[7] and this may seem to be no different. As it turns out, however, providing sufficient information to the jury to assess the weight that should be afforded the drug courier profile would require an inordinate amount of collateral testimony to validate the hidden probabilistic assumptions of the profile. As I shall demonstrate, these

5. As intent is an element of the crime, a defendant might be able to win an acquittal by persuading the jury that he did not know he was carrying controlled substances.

6. For now, I am glossing over complaints that such stops are disproportionately inflicted on minorities, particularly African-Americans. *See e.g.*, Elizabeth Mehren, *Judge Cites Man's Record of 'Driving While Black,' Eases His Sentence*, L.A. TIMES, Dec. 17, 1998, at A34.

7. *See* Delaware v. Fensterer, 474 U.S. 15, 22 (1985); Barefoot v. Estelle, 463 U.S. 880, 902-03 (1983).

hidden assumptions can be complex and detailed and may be far from the fact finder's or witness's expertise. It is for this reason, and not simple "unfair prejudice," that the exclusion of drug courier profiles as substantive evidence is justified.

II. Drug Courier Profiles

The ferrying of controlled substances by an airline passenger is a common way for drugs to enter the United States from a foreign country. To interdict such drug trafficking, law enforcement agents could, in the name of thoroughness, search every deplaning passenger. Such an approach, however, is unlikely to be socially acceptable. Nor would it be constitutional.

For one thing, warrantless searches "are per se unreasonable under the Fourth Amendment—subject only to a few specifically established and well-delineated exceptions."[8] Those exceptions can generally (but non-exhaustively) be categorized as (1) exigent circumstances such as the risk of destruction of evidence,[9] danger to individuals,[10] or fire safety;[11] (2) automobiles and movable objects inside automobiles;[12] (3) evidence in plain view;[13] (4) searches incident to lawful arrests;[14] and (5) border searches involving persons and property entering the country.[15] In the typical case involving a deplaning drug courier, none of the exceptions would appear to apply, with the arguable exception of the border search where the courier is traveling from a foreign country. In addition, searching every deplaning passenger is grossly inefficient, particularly if the number of anticipated couriers, relative to the entire group to be searched, is small. Imagine the public outrage, for example, if highway patrol officers, in an effort to detect drunk drivers, set up a roadblock in the middle of a heavily trafficked freeway and stopped every driver!

8. Katz v. United States, 389 U.S. 347, 357 (1967).
9. *See, e.g.*, United States v. MacDonald, 916 F.2d 766, 770 (2d Cir. 1990) (en banc).
10. *See, e.g.*, United States v. Riccio, 726 F.2d 638, 643 (10th Cir. 1984).
11. *See* Michigan v. Tyler, 436 U.S. 499, 509 (1978).
12. California v. Acevedo, 500 S. Ct. 565, 574 (1991) (containers inside cars); Chambers v. Maroney, 399 U.S. 42, 48 (1970) (cars seized and taken to police stations).
13. Horton v. California, 496 U.S. 128, 141 (1990).
14. Chimel v. California, 395 U.S. 752, 763 (1969) (objects within reach of arrested suspect).
15. United States v. Ramsey, 431 U.S. 606, 616 (1976).

A. Use of the Profile

The problem for agents, therefore, is determining which subset of passengers is likely to consist of couriers. The drug courier profile evolved as agents began compiling lists of characteristics common to couriers.[16] In *Elmore v. United States*,[17] the Fifth Circuit described a commonly used profile consisting of seven primary characteristics:

> (1) arrival from or departure to an identified source city; (2) carrying little or no luggage; (3) unusual itinerary, such as rapid turnaround time for a very lengthy airplane trip; (4) use of an alias; (5) carrying unusually large amounts of currency in the many thousands of dollars, usually on their person, in briefcases or bags; (6) purchasing airline tickets with a large amount of small denomination currency; and (7) unusual nervousness beyond that ordinarily exhibited by passengers.[18]

These primary characteristics are supported by four secondary characteristics:

> (1) the almost exclusive use of public transportation, particularly taxicabs, in departing from the airport; (2) immediately making a phone call after deplaning; (3) leaving a false or fictitious call-back telephone number with the airline being utilized; and (4) excessively frequent travel to source or distribution cities.[19]

In this article, I focus on the drug courier profile as used in airports. It is worth noting, however, that law enforcement officers have developed profiles for a host of other criminal suspects, including illegal alien smugglers,[20] child molesters,[21] drug dealers,[22] car thieves,[23] child abusers,[24] child batterers,[25] and sexual abusers.[26] Because these other

16. *See* Charles L. Becton, *The Drug Courier Profile: 'All Seems Infected That Th' Infected Spy, As All Looks Yellow to the Jaundic'd Eye,'* 65 N.C. L. REV. 417, 427 (1987). The profile is generally credited to Special Agent Paul Markonni of the DEA. *See* United States v. Ehlebracht, 693 F.2d 333, 335 n.3 (5th Cir. 1982).

17. 595 F.2d 1036 (5th Cir. 1979).

18. *Id.* at 1039 n.3.

19. *Id.*

20. United States v. Brignoni-Ponce, 422 U.S. 873, 884-85 (1975).

21. United States v. Gillespie, 852 F.2d 475, 479-80 (9th Cir. 1988).

22. Wilson v. State, 871 P.2d 46, 48 (Okla. Crim. App. 1994).

23. People v. Martinez, 12 Cal. Rptr. 2d 838, 840 (Cal. Ct. App. 1992).

24. Flanagan v. State, 586 So.2d 1085, 1099 (Fla. Dist. Ct. App. 1991).

25. Commonwealth v. Day, 569 N.E.2d 397, 399 (Mass. 1991).

26. State v. McMillan, 590 N.E.2d 23, 31 (Ohio Ct. App. 1990).

profiles also consist of groupings of characteristics, the analysis developed herein would apply to them as well.

If agents observe a person fitting a number of the characteristics, they might approach that person for further questioning. If the answers do not allay the agents' suspicions of the person, they might detain the person so that they can obtain a search warrant.[27]

B. Current Law

Currently, federal law impliedly recognizes the probative value of profile evidence. This, however, has not always been the case. In *Reid v. Georgia*,[28] a divided Court reversed Reid's conviction on the ground that police had "seized" him within the meaning of the Fourth and Fourteenth Amendments without having articulated a reasonable suspicion to justify his limited detention.[29] Much like Jones in my opening hypothetical, Reid arrived from a source city and in the early morning "when law enforcement activity is diminished." He also tried to conceal the fact that he was traveling with another person, and he had no luggage other than what he had carried on the flight.[30] The Court noted that most of the "circumstances describe a very large category of presumably innocent travelers, who would be subject to virtually random seizures were the Court to conclude that as little foundation as there was in this case could justify a seizure."[31]

Reid appeared to consider the factors it identified in isolation. That is, many people travel to or from source cities, and presumably, the overwhelming majority of them are innocent. Similarly, many innocent people travel in the early morning. And many people do not check any luggage, and presumably, most of those people are innocent as well. However, *Reid* did not address whether persons who travel to or from a source city in the morning, *and* who do not check any luggage *and* who conceal their association with other travelers are significantly less likely to be "innocent travelers" than persons fitting none of those factors.

Three years later, in *Illinois v. Gates*,[32] the Court replaced the more rigid framework that had governed probable cause analysis up to that point with a flexible, "totality of the circumstances" approach. Drawing on older cases, the Court noted that, "probable cause requires only a probability or substantial chance of criminal activity, not an actual

27. Morgan Cloud, *Search and Seizure by the Numbers: The Drug Courier Profile and Judicial Review of Investigative Formulas*, 65 B.U. L. REV. 843, 848 (1985).
28. 446 U.S. 438 (1980) (per curiam).
29. *Id.* at 440-41.
30. *Id.*
31. *Id.* at 441.
32. 462 U.S. 213, 238 (1983).

showing of such activity."[33] The Court acknowledged that probable cause would often be based on a collection of innocent activities, and that evaluating each individual activity as "innocent" or "guilty" was improper.[34] Indeed, the actions that the Gates' took were individually not indicative of guilt, although in sum, they did appear suspicious.[35]

By making clear that probable cause determinations should be based on the totality of the circumstances, *Gates* implicitly undermined the objection to drug courier profiles that had been articulated in *Reid.* Subsequently, in *United States v. Sokolow*,[36] the Court, while not overtly approving of profiles, rejected the argument that the factors giving rise to reasonable suspicion were per se invalid because they came from a courier profile: "A court sitting to determine the existence of reasonable suspicion must require the agent to articulate the factors leading to that conclusion, but the fact that these factors may be set forth in a 'profile' does not somehow detract from their evidentiary significance as seen by a trained agent."[37]

The Supreme Court, however, has yet to resolve whether profiles of the sort that it has thus far approved for law enforcement use are admissible as substantive evidence at trial. A majority of lower courts have addressed this question and have concluded that profile evidence is generally not admissible.[38] For example, in *United States v. Lui*, a U.S. Customs Inspector at the Los Angeles International Airport stopped Lui, primarily because he "appeared nervous."[39] His answers to questions by the customs agent only raised more questions, ultimately resulting in a search of his suitcases. The suitcases contained false panels housing more than twelve kilograms of nearly pure heroin.[40] Prior to his trial for importation of heroin and possession of heroin with intent to distribute, Lui moved to exclude any testimony that he fit a drug courier profile.[41] The trial court denied the motion, and the customs agent testified that Lui's behavior fit that of a drug courier in five different ways:

---

33. *Id.* at 244 n.13.

34. *Id.*

35. Some of the factors cited by the Court were the fact that Gates flew to Florida, "well-known as a source of narcotics and other illegal drugs;" stayed overnight in a motel; and immediately drove back to Illinois in the family car, which had been left "conveniently" in West Palm Beach. *Id.* at 243.

36. 490 U.S. 1 (1989).

37. *Id.* at 10.

38. *See Generally,* United States v. Williams, 957 F.2d 1238, 1242 (5th Cir. 1992); United States v. Lui, 941 F.2d 844, 847 (9th Cir. 1991); United States v. Jones, 913 F.2d 174, 177 (4th Cir. 1990); United States v. Carter, 901 F.2d 683, 684 (8th Cir. 1990); United States v. Hernandez-Cuartas, 717 F.2d 552, 555 (11th Cir. 1983); State v. Lee, 959 P.2d 799, 802 (Ariz. 1998); People v. Salcedo, 985 P.2d 7, 10 (Colo. Ct. App. 1998) *aff'd* 999 P.2d 833 (Colo. 2000); Dean v. State, 690 So.2d 720, 722 (Fla. Dist. Ct. App. 1997); People v. Hubbard, 530 N.W.2d 130, 133 (Mich. App. 1995). *But see* United States v. Foster, 939 F.2d 445, 451-52 (7th Cir. 1991).

39. *Lui*, 941 F.2d at 845.

40. *Id.* at 846.

41. *Id.*

> 1) although "heroin couriers have a hundred ways to smuggle heroin," typically they wrap it around their bodies or place it in false bottoms and tops of suitcases; 2) in approximately eighty percent of smuggling cases, couriers use hard-sided suitcases; 3) couriers often use the excuse of conducting business or visiting a relative in the United States; 4) couriers create itineraries with multiple stops for short periods of time so as to enter the United States from a "non-source" country to avoid detection; and 5) couriers often use paging devices.[42]

The Ninth Circuit concluded that the admission of the courier profile testimony was erroneous under the facts of the case. The court rejected the government's argument that the profile evidence should be admissible to show modus operandi, because this case was not a "complex drug-smuggling conspirac[y]."[43] Thus, the danger of prejudice greatly outweighed the probative value of the evidence. In light of the overwhelming evidence of guilt, however, the error was harmless.[44]

Other courts have used similar reasoning: "Drug courier profiles are inherently prejudicial because of the potential they have for including innocent citizens as profiled drug couriers."[45] Moreover, "the fact that an individual fits the profile does not necessarily mean that the evidence in a particular case will show that the person was carrying drugs."[46]

Despite the general ban on profiles as substantive evidence, however, courts have allowed them to be used for rebuttal purposes under certain circumstances, where the defendant "opens the door" by testifying that he did not fit the courier profile. For example, in *United States v. Beltran-Rios*,[47] defense counsel's cross-examination of the government agents consisted of asking whether the defendant, at the time of the arrest, wore expensive jewelry, had large amounts of cash on his person, or had recently made large purchases.[48] The Ninth Circuit held that Beltran-Rios had opened the door to the admission of profile evidence to rebut the inference that he was not a courier because he did not fit the drug courier profile.[49]

The Seventh Circuit, on the other hand, has held that drug courier profiles are admissible as non-scientific expert testimony used to assist the jury in determining whether the defendant was aware that he was

---

42. *Id.*

43. *Id.* at 848; *cf.* United States v. Klimavicius-Viloria, 144 F.3d 1249, 1259-60 (9th Cir. 1998) (allowing expert testimony of routes used by drug couriers).

44. United States v. Lui, 941 F.2d 844, 848 (9th Cir. 1991).

45. United States v. Hernandez-Cuartas, 717 F.2d 552, 555 (11th Cir. 1983).

46. United States v. Williams, 957 F.2d 1238, 1242 (5th Cir. 1992).

47. 878 F.2d 1208 (9th Cir. 1989).

48. *Id.* at 1211-12.

49. *Id.* at 1212-13.

carrying narcotics at the time of his arrest.[50] The linchpin of the court's holding was its assumption "that jurors are not well versed in the behavior of drug dealers."[51] Thus, an expert's testimony regarding drug couriers would be helpful to the jury by providing an " 'understanding of the intricate patterns and modus operandi' of those involved in narcotics trafficking."[52] Unlike the Ninth Circuit, however, the Seventh Circuit did not limit its holding to cases involving complex drug smuggling routes; rather, it seemed to presume that drug smuggling was itself a complex operation.

C. Traditional Arguments Against Profiles

Drug courier profiles have not been without controversy. Opponents charge that the profiles possess a " 'chameleon-like way of adapting to any particular set of observations,' "[53] and consist of "anything that happens to be suspicious in a particular case."[54] This protean nature, critics contend, allows law enforcement agents to come up with ad hoc rationalizations to justify what essentially amounts to unrestrained discretion.[55] Thus, this criticism strikes directly at the heart of the profile, at its predictive value.

Admittedly, the use of subjective factors in the profile can lead to abuse by law enforcement officers. For instance, the characterization of whether a passenger exhibited extreme nervousness is dependent on the observer's interpretation of physical symptoms.[56] Yet, that characterization may not be easily challenged or cross-examined. Moreover in other probable cause settings, courts have been skeptical of subjective characterizations of suspicious behavior. For example, one court rejected as "suspicious" an agent's characterization of a driver's failure to make eye contact:

> A driver's failure to look at the Border Patrol car [cannot be used to justify the agent's suspicion] since the opposite reaction, a driver's repeated glancing at a Border Patrol car,

50. United States v. Foster, 939 F.2d 445, 451-52 (7th Cir. 1991) (citing FED. R. EVID. 702).

51. *Id.* at 452.

52. *Id.* (quoting United States v. Gonzalez, 933 F.2d 417, 429 (7th Cir. 1991)).

53. United States v. Sokolow, 490 U.S. 1, 13 (1989) (Marshall, J., dissenting) (quoting 831 F.2d 1413, 1418 (9th Cir. 1987)).

54. Cloud, *supra* note 27, at 843.

55. Becton, *supra* note 16, at 438. For an in-depth analysis of the problems with unrestrained government discretion, *see generally* KENNETH CULP DAVIS, DISCRETIONARY JUSTICE (1969).

56. By that, I do not mean whether a symptom of nervousness is correctly ascribed to consciousness of guilt—for example, whether a person's hands tremble due to nervousness about carrying drugs or simply a fear of flying. Rather, I write here about whether a slight tremor constitutes extreme nervousness. In other words, the degree of nervousness is my concern, as opposed to the cause of the nervousness.

> can also be used to justify the agent's suspicion. To give weight to this type of justification "would put the officers in a classic 'heads I win, tails you lose' position [and] the driver, of course, can only lose."[57]

This criticism is directed at profiles as they are currently constructed. Practical concerns should not be brushed aside lightly, but these criticisms have been voiced elsewhere in detail.[58] I am interested in a different aspect of the question, and that is whether it is theoretically possible to use a profile properly, either as an investigative tool or as substantive evidence. Consider the so-called *Elmore* profile described by the Fifth Circuit.[59] Although this profile does contain one subjective characteristic (unusual nervousness), it is generally objective in nature. The question I address is whether the use of an objective profile, based upon observations of past events, can be justified for any purpose.

A different argument being used increasingly to attack profile evidence is that it constitutes "group character evidence."[60] Under the Federal Rules of Evidence, character evidence is generally "not admissible for the purpose of proving action in conformity therewith on a particular occasion."[61] Thus, as the Ninth Circuit has noted, "[T]he government may not . . . prove that the defendant is a bad person, simply to show that in all likelihood he acted criminally on the occasion at issue."[62] Since the probative value of profile evidence "depends on the jury drawing the forbidden inference that the defendant has a propensity to commit the crime with which he is charged," opponents argue that it should be excluded for the same reason.[63]

The problem with this argument is that character evidence can be admitted to demonstrate purposes such as intent, knowledge, or absence of mistake, etc.[64] Thus, the profile is not offered to prove the lack of propensity to transport controlled substances, but knowledge by the defendant that he or she was doing so. Opponents counter that profile evidence "does not reflect the defendant's motivation for behaving in a certain manner, but rather imposes on the defendant the motivation of third parties not connected to the charged crime."[65] In this way, the use of profiles "violates the 'central teaching' of Fourth Amendment

57. United States v. Garcia-Camacho, 53 F.3d 244, 247 (9th Cir. 1995) (quoting Gonzalez-Rivera v. INS, 22 F.3d 1441, 1447 (9th Cir. 1994)) (alterations in original).

58. Typical criticisms can be found in Becton, *supra* note 16, and Kadish, *supra* note 2.

59 *See supra*, note 18.

60. *See, e.g.*, State v. McMillian, 590 N.E.2d 23, 31 (Ohio Ct. App. 1990).

61. FED. R. EVID. 404(a).

62. United States v. Martinez, 182 F.3d 1107, 1111 (9th Cir. 1999).

63. Kadish, *supra* note 2, at 783; *see also* Cloud, *supra* note 27, at 852; State v. Lee, 959 P.2d 799, 802 (Ariz. 1998) ("[U]se of drug courier profile evidence to indicate guilt . . . creates too high a risk that a defendant will be convicted not for what he did but for what others are doing.") (quoting State v. Cifuentes, 830 P.2d 469, 469 (Ariz. Ct. App. 1991)).

64. FED. R. EVID. 404(b).

65. Kadish, *supra* note 2, at 784.

jurisprudence that each case must be examined on its own individualized facts."[66] This is an objection, which I shall more thoroughly discuss *infra*. The assumption that the actions of third parties cannot shed insight upon the actions of an individual is not borne out by empirical analysis.

Drug courier profiles are also criticized on the ground that they are used in racially discriminatory ways. One commentator argues that, "Although the DEA has refused to commit the entire profile to writing, the profile clearly contains a racial component."[67] In a case before the en banc Sixth Circuit, an officer admitted during an evidentiary hearing that seventy-five percent of persons detained through the use of the profile were Black.[68] In a strongly worded dissent, Judge Damon Keith wrote, "The disproportionate number of African-Americans who are stopped indicates that a racial imbalance against African-Americans does exist and is implicitly sanctioned by the law enforcement agency. . . . We cannot allow law enforcement officers to cloak what may fairly be characterized as a racist practice in a generic drug profile that openly targets African-Americans."[69]

If the factual assumptions of this argument are true, it cannot be dismissed lightly. This is true even if it is the case that the profile targets minorities because they in fact are more likely to be drug couriers than non-minorities. Jody Armour refers to decision-makers, who attempt to justify using race in criminal procedure issues on the basis of disproportionate rates of crimes committed by minorities, as "intelligent Bayesians," after the noted mathematician Thomas Bayes.[70] Bayes is recognized for his contribution to the theory of conditional probability. Conditional probability asks, given that event X has occurred, what is the probability that event Y will occur? In Armour's discussion, the presumption is that, if it is known a racial group commits crimes at a higher rate than general society does, then the probability that a person of that group has committed a crime is higher than for an average person. According to Armour, the intelligent Bayesian's motto is, "As much as I regret it, I must act differently toward Blacks because it is logical to do so."[71] However, Armour challenges this reasoning by pointing out that, despite the presumption that Blacks are more likely to commit violent crimes than Whites on a per capita basis, the percentage of the Black

66. Cloud, *supra* note 27, at 920.

67. Sheri Lynn Johnson, *Race and the Decision to Detain a Suspect*, 93 YALE L.J. 214, 234 (1983); *see also* Sandra Guerra, *Criminal Law: Domestic Drug Interdiction Operations: Finding the Balance*, 82 J. CRIM. L. & CRIMINOLOGY 1109, 1147 (1992) (noting that the typical profile appears to target Blacks and Hispanics).

68. *See* United States v. Taylor, 956 F.2d 572, 581 n.1 (6th Cir. 1992) (en banc) (Keith, J., dissenting).

69. *Id.* at 581-82.

70. Jody D. Armour, *Race Ipsa Loquitur: Of Reasonable Racists, Intelligent Bayesians, and Involuntary Negrophobes*, 46 STAN. L. REV. 781, 790 (1994).

71. *Id.*

male population arrested for violent crimes in 1991 was less than 1.7 percent of the population.[72] Thus, whatever probative value the statistical disparity has, it is dwarfed by the unfair prejudice that it inflicts on the other 98.3 percent, and cannot bear the weight of race-based discrimination.[73]

Randall Kennedy analogizes the racial component of the drug courier profile to a sort of "racial tax for the purpose of more efficient law enforcement."[74] Although Kennedy suggests that the use of race in investigations may not be completely illegitimate, he nonetheless argues that such use should not be condoned, because society has the resources to "increase taxes across the board" by hiring more police or stopping everyone randomly, and using race only in extremely rare circumstances.[75]

Moreover, the addition of a racial component creates the possibility of a self-fulfilling prophecy. If the profile assumes that racial minorities are more likely than Whites to be couriers, then the percentage of minorities in the group of people stopped will be higher than the percentage of minorities in the overall population. Even if minorities are in reality no more predisposed than Whites to serve as drug couriers, they will nevertheless appear to be more predisposed, because they will make up a larger proportion of the pool of those persons stopped.

Of course, the same argument can be applied to any system of selection; the use of race-neutral factors may reinforce the belief in the probative value of those factors. A brief response is that: (1) the race-neutral factors are not proscribed by the Constitution as suspect categories the way that race is; and (2) those factors are not immutable the way that race is, so that persons wishing to decrease their probability of being stopped can do so by altering their behavior. These are, of course, not entirely satisfactory answers. Thus, the race-conscious nature of the profile would appear to be validated, when in fact true validation would be lacking.[76]

As this issue has been addressed elsewhere in more detail, I am going to put aside this particular challenge to drug courier profiles. Unless race is the only factor in the drug courier profile—which would make it difficult, if not impossible, to defend in any meaningful way—the profile can be administered in a race-neutral manner, at least on a theoretical basis, by simply eliminating the racial component and using the other facially race-neutral factors.[77] The race-based challenge presents no obstacle to that type of use. Thus, while cognizant of the

72. *Id.* at 791.
73. *Id.* at 796.
74. RANDALL KENNEDY, RACE, CRIME & THE LAW 159 (1997).
75. *Id.* at 161.
76. *See id.* at 22.
77. Whether such factors nevertheless target minorities can be addressed by conventional disparate impact analysis. *See generally* Hazen Paper Co. v. Biggs, 507 U.S. 604, 609 (1993).

problems of the profile as used in real life, I shall assume for the rest of the article that the factors are race-neutral.

III. Probative Value

Does the drug courier profile have any probative value? That is, is there any reason to believe that a person, who engages in a collection of activities that, taken individually, are innocent, may nevertheless be more likely to engage in criminal activity? If not, then the profile cannot be justified even for investigatory purposes, much less for use in trials, and *Sokolow* was decided incorrectly.

A. Probabilistic Evidence

In essence, the drug courier profile purports to be probabilistic evidence. The probability that a person who engages in the conduct highlighted by the profile is a drug courier is higher than the probability for the population at large. Otherwise, there would be no reason to use the profile. Because of this increased likelihood of criminal activity, the police may be justified in detaining that person for further questioning.

When used as substantive evidence at trial, the profile also implies that if the defendant fits the profile, he or she is probably a drug courier. However, "probably" in this context is extremely vague—what is the jury to make of it? The evidence becomes more easily considered if, for instance, the testifying witness can state that a person who fits the profile is z percent likely to be a drug courier. For example, in an otherwise close case, knowing whether z is closer to, say, ninety-five or fifty-five, may be dispositive to the jury's verdict.

Of course, all relevant evidence is probabilistic, in the sense that it tends "to make the existence of any fact that is of consequence to the determination of the action more probable or less probable than it would be without the evidence."[78] Suppose, for example, that prior to the admission of evidence "E," the probability of guilt is assessed at x percent. "E" is relevant if, after it is taken into account, the assessment of probability of guilt is no longer x percent, but is either greater or lesser.

The above example presumes a Bayesian approach to evidence in the sense that the probability of an event (known as the "posterior

78. FED. R. EVID. 401; *see also* Daniel Shaviro, *Statistical-Probability Evidence and the Appearance of Justice*, 103 HARV. L. REV. 530, 538 (1989) ("[A]ll uses of evidence are probabilistic. . . .").

probability") can be revised in light of a related event's occurrence. Bayes' Theorem provides the mathematical formula for doing so. In general, the probability that event X will occur given that event Y has already occurred is:

$$P(X|Y) = \frac{P(X)\,P(Y|X)}{P(X)\,P(Y|X) + P(X')\,P(Y|X')}$$

where P(X|Y) is the probability that event X will occur given that event Y has occurred, P(Y|X) is the probability that event Y will occur given that event X has occurred, P(X') is the probability that event X does not occur, and P(Y|X') is the probability that event Y occurs given that event X has not occurred.[79] Michael Finkelstein and William Fairley suggest that the probative value of each additional piece of evidence can be processed in relation to all the other evidence so far; in other words, given the new piece of evidence, what is the updated probability of guilt?[80]

If event X is thought of as the event that the defendant is guilty, P(X) is the prior estimate of that probability, and P(Y) is the probability that a person who fits the courier profile is indeed a drug courier, then P(X|Y) yields the probability that the defendant is guilty given that he or she fits the drug courier profile.[81]

It is important to observe that Bayes' Theorem is useful primarily as a tool for the trial judge to use in evaluating the admissibility of a given piece of evidence. It is considerably less suited for use by the jury, as it "requires that the decision maker evaluate each bit of evidence as it is introduced." This requirement runs afoul of the usual edict that the jury is "not to form any generalizations until all the evidence is in."[82]

B. Paradoxes of Statistical Evidence

To answer whether P(Y), the probability that a person who fits the drug courier profile is indeed a drug courier, has any meaning, I propose to analogize the profile to a simple hypothetical articulated by Laurence Tribe in his celebrated article, *Trial by Mathematics:*

> A man is found shot to death in the apartment occupied by his mistress. The question is whether she shot him.

79. PAUL G. HOEL ET AL., INTRODUCTION TO PROBABILITY THEORY 16-17 (1971).

80. *See id.* at 498-99.

81. Michael O. Finkelstein & William B. Fairley, *A Bayesian Approach to Identification Evidence*, 83 HARV. L. REV. 489, 500 (1970).

82. RONALD J. ALLEN & RICHARD B. KUHNS, AN ANALYTIC APPROACH TO EVIDENCE 128 (1989).

> Evidence is available to the effect that, in ninety-five percent of all known cases in which a man was killed in his mistress' apartment, the mistress was the killer. How, if at all, may such evidence be used?[83]

A little imagination translates the characteristics of being a mistress, whose lover is murdered in one's apartment, into the equivalent of the drug courier characteristics (i.e., flying to, or from, a source city; paying for a ticket in cash; and so on). The ninety-five percent "fact" corresponds to an estimate of the probability that a person fitting the profile is in fact, factually guilty; with the drug courier profile, this estimate is unknown but presumed to be higher than the probability for a person picked at random. Is there a basis for admitting this evidence?

One immediate problem is that this "evidence" regarding men murdered in their mistresses' apartments, if admitted in the trials of 100 mistresses, might lead to 100 convictions, even though we would only expect ninety-five of them to be guilty. This possibility is especially acute if jurors tend to think of burdens of proof in mathematical terms—that is, preponderance of the evidence as just more than a fifty percent probability of being true[84] and beyond a reasonable doubt as ninety-five percent or greater certainty.[85] Under such a regime, the "evidence" that "ninety-five percent of the time the mistress did it" may, by itself, push juries past that ninety-five percent certainty level in each case, leading to 100 convictions.[86] And yet, there seems to be something drastically wrong with an evidentiary system that would allow defendants to be convicted based solely on what 100 other people have done,[87] particularly since we would expect five percent to be wrongly convicted.

This in fact, is a variation of a set of famous problems articulated by law professors engaged in a debate about the merits of probabilistic evidence. In what is referred to as the "Blue Bus case," the plaintiff is hit by a blue bus. In her lawsuit against the defendant, a bus company, the only evidence she offers is that the defendant operates four-fifths of all blue buses in the city.[88] Tribe points out that, if this evidence is

---

83. Laurence H. Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329, 1341 (1971).

84. *See, e.g.*, Neil B. Cohen, *Conceptualizing Proof and Calculating Probabilities: A Response to Professor Kaye*, 73 CORNELL L. REV. 78, 93 (1987).

85. *See, e.g.*, Shaviro, *supra* note 78, at 532; Tribe, *supra* note 83, at 1372 (constructing an argument against this commonly used figure).

86. For those readers who believe that reasonable doubt requires far more than ninety-five percent certainty, all we have to do is to adjust the numbers in the hypothetical. Because reasonable doubt by definition is less than absolute certainty, assign the number x to the required level of certainty (i.e., $95 < x < 100$). Then assign y to the frequency with which the mistress is guilty such that $x < y < 100$.

87. *See* Charles Nesson, *Reasonable Doubt and Permissive Inferences: The Value of Complexity*, 92 HARV. L. REV. 1187, 1192-94 (1979).

88. *See* Jonathan J. Koehler & Daniel Shaviro, *Verdical Verdicts: Increasing Verdict Accuracy Through the Use of Overtly Probabilistic Evidence and Methods*, 75 CORNELL L. REV. 247, 257

enough to sustain a verdict for the plaintiff, then the owner of the blue buses, "however careful, would have to pay for five-fifths of all unexplained blue bus accidents, a result as inefficient as it is unfair."[89] One solution would be to reduce the recovery by twenty percent in every case. Such a solution would address the problem from the standpoint of the owner of the blue buses, but would result in every plaintiff recovering only eighty percent of his or her damages.

In the "Gatecrasher Paradox," a promoter sells 499 tickets to a rodeo, but 1000 people attend. Therefore, 501 of the attendees did not pay.[90] Unfortunately, there is no way of knowing whether a given person in that group of 1000 paid for admission, though there is a 50.1 percent chance that he or she did not. Surely, the paradox asserts, the promoter cannot recover against all 1000 people.[91]

Legal scholars have wrestled with the question of why the defendants in these cases should not be entitled to verdicts in their favor. Tribe takes a policy-oriented approach, arguing that the plaintiff has an obligation to present some specific evidence, beyond background statistics, that links the defendant to this case, or to provide a satisfactory explanation as to why such evidence is lacking.[92] Any other rule, he argues, "would eliminate any incentive for plaintiffs to do more than establish the background statistics."[93]

Although Tribe makes good points, particularly regarding incentive structures, his criticism is incomplete. Even if the defendant bus company operated ninety-nine percent of the blue buses in town, Tribe would deny recovery unless the plaintiff adduced specific evidence.[94] At the same time, however, Tribe allows for the possibility that background base rate evidence, "when properly combined with other, more conventional, evidence in the same case," may "supply a useful link in the process of proof."[95] Suppose the plaintiff saw that the bus that hit her had an "A" and a "1" in the license plate, and some number of the defendant's blue buses in fact have an "A" and a "1" in their license plates.[96] Is that observation, combined with the knowledge that the defendant operates eighty percent of the blue buses in town, truly

(1990); Tribe, *supra* note 83, at 1341-42. This hypothetical is based loosely on Smith v. Rapid Transit, Inc., 58 N.E.2d 754 (Mass. 1945).

89. Tribe, *supra* note 83, at 1350 (emphasis in original).

90. L. JONATHAN COHEN, THE PROBABLE AND THE PROVABLE 75 (1977).

91. The hypothetical appears to assume that no tickets or receipts were issued, and none of the 499 paying customers has a credit card receipt or a canceled check to prove payment.

92. Tribe, *supra* note 83, at 1349 & n.65; *see also* Lea Brilmayer, *Second-Order Evidence and Bayesian Logic*, 66 B.U. L. REV. 673, 674-75 (1986); Charles Nesson, *The Evidence or the Event? On Judicial Proof and the Acceptability of Verdicts*, 98 HARV. L. REV. 1357, 1379-80 (1985).

93. Tribe, *supra* note 83, at 1349.

94. *Id.*

95. *Id.* at 1350 (emphasis in original).

96. The perceptive reader will see that this fact about the license plates is really just a more specific form of base rate evidence. To a large extent, all evidence is probabilistic in nature. *See* Koehler & Shaviro, *supra* note 88, at 263.

a higher quantum of quality of presentation than the simple base rate fact that the defendant operates, say, ninety-nine percent of the blue buses in town? After all, the troubling nature of the Blue Bus case and the Gatecrasher Paradox stems largely from the fact that application of the base rate evidence will lead to a distressingly high number of incorrect verdicts. But we can hardly expect to do better than ninety-nine percent in a civil case.[97]

Neil Cohen, on the other hand, proposes that one construct a ninety-five percent confidence interval around the probability estimate, and to deny the plaintiff recovery unless the entire confidence interval lies above the burden of proof standard (i.e., fifty percent for civil cases).[98] In essence, Cohen says that it is not enough for the probability estimate to exceed the threshold burden of proof to be confident that the verdict is correct, for that estimate "is more accurately described as an 'estimate' based on a small portion, or a 'sample,' of information rather than as a true value derived from an analysis of all possible information."[99] For example, if thirty out of fifty balls drawn from an urn are white, one would estimate the probability of drawing a white ball from the urn at 0.60,[100] just as it would be if 60,000 out of 100,000 balls drawn are white. However, the probability that the next ball drawn "will be white is ninety-five percent certain to be 0.6 plus or minus 0.14 (that is, between 0.46 and 0.74) [in the first example], whereas based on the information provided by the larger sample, the probability is ninety-five percent certain to be 0.6 plus or minus 0.003 (that is, between 0.597 and 0.603)."[101] In other words, the probability estimate is the same in both examples, but the probability that the probability is accurate increases with the size of the sample.

One theoretical problem with Cohen's approach is that he constructs the confidence interval by assuming that the probability estimate is in fact the actual probability.[102] However, we have no way of knowing what the actual probability is without knowing the actual distribution of balls in the urn. In other words, what we are interested in knowing is whether the ninety-five percent confidence interval around the true probability estimate exceeds the evidentiary threshold (0.5 in civil cases and some much higher but undefined number for criminal

97. For a more detailed criticism and analysis, *see* Shaviro, *supra* note 78.

98. Neil B. Cohen, *Confidence in Probability: Burdens of Persuasion in a World of Imperfect Knowledge*, 60 N.Y.U. L. REV. 385, 402-04 (1985). A confidence interval is a range within which one can expect a stated percentage of trials to fall. In other words, the ninety-five percent confidence interval is the range for which ninety-five percent of the data points will fall within. *See* WARREN WEAVER, LADY LUCK: THE THEORY OF PROBABILITY 319-21 (1963).

99. *Id.* at 398.

100. Cohen does not state explicitly that any ball drawn is replaced before the next draw, but that is implicit in his example.

101. *Id.* at 399.

102. *See* D.H. Kaye, *Apples and Oranges: Confidence Coefficients and the Burden of Persuasion*, 73 CORNELL L. REV. 54, 60-61 (1987).

cases). But to answer that question would require knowing what the true probability is, which, if we knew, would moot the entire exercise. What Cohen proposes is an imprecise approximation—he constructs the ninety-five percent confidence interval around the probability estimate and implicitly assumes that this confidence interval is sufficiently close to the actual probability that there is an insignificant loss of accuracy.[103]

The objections illustrated above apply to the use of such evidence in trials. But what about using the ninety-five percent "evidence" as grounds for detaining the mistress for further questioning? Based on the reasoning in *Sokolow*, it would appear the fact that the mistress fits this "profile" may be sufficient for such purposes. Does the "evidence" regarding mistresses and their dead lovers make it seem more likely that the mistress is, in this particular case, the killer? Certainly one would think that a prudent police detective would, given this fact, treat the mistress as a likely suspect, if not the most likely suspect. After all, there are only a finite number of suspects,[104] and in similar past circumstances, the mistress has been identified as the murderer nineteen times more often than everyone else combined.

The mathematically inclined reader might respond, "No, wait. That ninety-five of the last 100 men shot to death in their mistresses' apartments were killed by the mistresses has nothing to do with this case, just like the fact that a coin has come up heads ninety-five times out of 100 has no impact on whether it will come up heads on the 101st toss." True, a fair coin could come up heads ninety-five times out of 100 and still have only a fifty-fifty chance of coming up heads on the next toss.[105] But in the real world, if a coin came up heads ninety-five times out of 100 one would probably suspect that it is not a fair coin. Similarly, the "evidence" regarding mistresses suggests (but does not prove) that the observation is in fact not random, but rather revealing of some facet of human psychology. It may be, for example, that adulterous relationships between married men and their mistresses are emotionally volatile, often driving one or the other to violence in a fit of passion.

C. Predictive Patterns

At root is the question of whether past historical trends have any predictive value about how individuals act. The hypothetical evidence about mistresses, and, more to the point, the drug courier profile, is

103. The validity of such a confidence interval is the subject of an article by David Kaye and a response by Cohen, with Kaye arguing (as I do here) that all one will have is a "prediction interval for p based on an assumed values for [the probability] of .51." Kaye, *supra* note 102, at 62. For Cohen's response, see COHEN, *supra* note 90, at 78.

104. Necessarily, since there are only a finite number of people in the world.

105. *See* WEAVER, *supra* note 98, at 183.

presumably based on long-term historical observations of how people have acted in certain circumstances. With the mistress example, the database consists of all solved cases in which men have been found murdered in their mistresses' apartments; and with the drug courier profile, the database is the shared set of observations by law enforcement agents. To assess whether historical patterns can shed any useful information, it is helpful to consider other circumstances in which such patterns are employed for decision-making purposes.

1. Yield Management

The best evidence that patterns do in fact provide useful predictive value is the modern business technique of "yield management," used by airlines, hotels, semiconductor fabricators, and rental car agencies.[106] Airlines, for example, seek to sell the optimal mixture of business fare and leisure fare seats.[107] Business fares are expensive but can be bought at the last minute, while leisure fares are cheap but require advance purchase.[108] Obviously, the airline would maximize its revenue if it could sell all business fares, but if it prices all of its seats at the business rate and cannot sell them all, then it will receive $0 revenue for the empty seats and would have been better off pricing them at the leisure rate.

To put it in numerical terms, suppose that Acme Airline flies a 100-seat airplane between Los Angeles and San Francisco. Business fares cost $200, leisure fares cost $100, leisure passengers will not pay business fares, and business passengers would be happy to pay leisure rates but are generally ineligible for them because they do not know in advance when they need to travel. Acme earns $200,000 if it sells all 100 seats at the business rate, compared to $100,000 if it sells all 100 seats at the leisure fare. If Acme charges all business rates and sells seventy-five seats on a given flight, it will earn $150,000 for those seats, and $0 for the twenty-five unsold seats. Had it charged $200 for seventy-five seats and $100 for twenty-five seats, it would have filled those twenty-five seats as well, and received a total of $175,000 for the

106. *See* Alex Taylor III, *Back in the Driver's Seat*, FORTUNE, May 25, 1998, at 212 (hotels, airlines, rental cars); *Windows, Aisle, or ASIC?*, PC WEEK, June 19, 1995, at E3 (semiconductor factories).

107. *See* Scott McCartney, *Ticket Shock: Business Fares Increase Even as Leisure Travel Keeps Getting Cheaper*, WALL ST. J., Nov. 3, 1997, at A1.

108. *See, e.g.*, John Schmeltzer, *Chess Champ Flying Deep Blue Skies*, STAR-LEDGER, Dec. 29, 1997, at 24; *United Unveils 'Deep Blue'-Powered Yield Management System*, AVIATION DAILY, Nov. 7, 1997, at 231.

flight.[109] But if Acme guessed wrong and priced half the seats at business rates and half at leisure rates, it would receive only $150,000.

The trick for the airlines, of course, is determining in advance the proper mix of seats. This is where past seating patterns are helpful. By running those past patterns (which are kept in massive computer databases) through mathematical algorithms and yield-management software, "carriers [can] predict with almost pinpoint accuracy how many business customers would want seats on a particular flight."[110] Similarly, "[b]y analyzing historical data about demand patterns, [rental car agencies] can change prices as often as every fifteen minutes in an effort to rent every car every day at the best possible price."[111]

Note that yield management is similar but not identical to the sampling techniques used by pollsters, like A.C. Nielsen, which provides ratings demographics on television programs.[112] With polls, a sample of the population is surveyed, and its responses are projected onto a much larger set. The technique has no predictive value; it merely allows the observer to determine how the entire population would respond to a given inquiry without having to question every single person. Yield management, on the other hand, uses a large database of past responses to predict how a subset will respond.

Admittedly, this technique is an imperfect analogy to the drug courier profile. Yield management attempts to predict how a group of persons will respond to particular circumstances. For example, as a flight date nears, how many business travelers are likely to want or need a seat on the flight? Yield management does not, however, attempt to predict individual behavior, the way that the drug courier profile does. Nevertheless, both techniques are predicated upon the belief that people can be expected to act similarly to how they have acted in previous similar circumstances.

2. *Insurance Premiums*

A second business example of reliance upon historical patterns to predict human behavior can be found in the insurance industry.[113] An

109. In mathematics, this type of problem is known as a "constrained maximization" problem, for which differential calculus is generally appropriate.

110. McCartney, *supra* note 107, at A10.

111. Taylor, *supra* note 106, at 212.

112. *See* David Bauder, *Television Ratings System Under Attack as Never Before*, ASSOC. PRESS, Dec. 30, 1996, *available in* 1996 WL 4456058; *How Nielsen Measures Ratings*, PEORIA J. STAR, Jan. 5, 1997, at C11.

113. Obviously, not all insurance covers events attributable to human conduct; earthquakes and hurricanes, for example, are natural disasters. However, property damage due to vandalism or negligence is caused by human behavior. Thus, in this subsection, I am speaking of the latter,

insurance company accepts a premium payment q in exchange for the obligation to pay loss L if event X occurs.[114] To make a profit, the insurance company must set the premium q so that it is higher than the expected loss. In other words:

$$q > p(X) * L$$[115]

where $p(X)$ is the probability that event X will occur within the period of coverage.

It is imperative that the insurance company assesses $p(X)$ accurately or else the premiums charged may not be sufficient to cover the payments it must make.[116] Typically, an insurance company calculates $p(X)$—the likelihood of event X—by considering "the historical frequency" of X.[117] In other words, insurers use the past to predict the future.

3. *Stock Market Prediction*

As might be expected, investors in the stock market have long attempted to predict price movements based on observable historical patterns. The branch of stock selection known as "technical analysis" is based upon a belief that one can "study the past—both the movements of common stock prices and the volume of trading—for a clue to the direction of future change."[118] Although technical analysis has its detractors,[119] it "is now considered mainstream securities research, used by major brokerage houses and money managers alike."[120]

Why would technical analysis have any predictive value? The main argument against technical analysis is that, "The stock market has no memory."[121] However, this assertion discounts the same factor that provides probative value to the hypothetical fact regarding mistresses: human psychology. Investor behavior may affect stock prices in

---

though I note that even the former is dependent on the use of historical patterns to predict the future. *See* Alan Sachs, *Special Report: Insurance*, BALT. BUS. J., Aug. 29, 1997, at 12.

114. *See* HAL VARIAN, MICROECONOMIC ANALYSIS 161 (2d ed. 1984).

115. *Id.* In a perfectly competitive market, long-run profits are driven to 0, and q will be equal to $p(X) * L$.

116. Actually, an insurance company could profit even if the premiums do not cover the expected losses, as long as the interest it earns on the premiums generate enough income to offset its losses.

117. VARIAN, *supra* note 114, at 298; *see also* JANET BANFORD, SMARTER INSURANCE SOLUTIONS 6-7 (1996); ROBERT COOTER & THOMAS ULEN, LAW AND ECONOMICS 65 (1988).

118. BURTON GORDON MALKIEL, A RANDOM WALK DOWN WALL STREET 117 (6th ed. 1996). For a more detailed description of the technical analysis of individual stocks, see MARTIN J. PRING, TECHNICAL ANALYSIS EXPLAINED 427-56 (3d ed. 1991).

119. *See, e.g.*, MALKIEL, *supra* note 118, at 140-45.

120. James A. Anderson, *The Trend Is Your Friend*, MONEY, Apr. 1, 1998, at A1.

121. MALKIEL, *supra* note 118, at 140.

irrational but predictable ways, "like pushing new issues too high and out-of-favor companies too low."[122] A recent study showed that the use of a basic technical analysis strategy outperformed a traditional buy and hold strategy, which one would not expect if technical analysis were useless.[123]

The stock market can have "memory" in the form of "support" and "resistance" prices. A support level is a price at which the number of buyers of a stock greatly outnumbers the number of sellers, and where sufficient buying pressure will materialize if the stock price trades down to it.[124] Thus, the stock price is unlikely in the short-term to trade below this level unless there is a change in the company's fundamentals (such as an announcement that earnings will fall short of analysts' estimates),[125] or a market-wide exogenous event (such as an increase in interest rates).[126] Similarly, a resistance level is a price at which the number of sellers of a stock greatly outnumber the number of buyers, thereby putting downward pressure on the stock price.[127] Technical analysts explain that support and resistance levels often exist due to psychological feelings in the minds of investors.[128] For example, if a number of investors buy a stock at $30 a share and it drops to $20, by the time the stock recovers to $30, investors may sell it en masse, thinking that they are lucky to break even. The $30 price becomes a resistance level. Similarly, if the stock drops to $20 again, investors may remember that it rebounded from that level and buy en masse, thus establishing $20 as a support level.[129]

Technical analysis thus develops principles of stock price movement based upon observations of how stocks in general have moved in the past. These principles are then used to predict the movement of individual stocks. In the same way, the drug profile is based upon observations of how drug couriers have acted in the past, and it is then used to predict which persons are likely to be couriers.

Similarly, an increasingly popular investment strategy is known as the "Dogs of the Dow."[130] This strategy, which involves selecting a

122. Paul Sturm, *Technical Knockouts*, SMARTMONEY, April 1998, at 65.

123. *See id.* (citing Blake LeBaron et al., *Simple Technical Trading Rules and the Stochastic Properties of Stock Returns*, 47 J. FIN. 1731 (1992)).

124. PRING, *supra* note 118, at 54.

125. The stock price of Cendant Corp. dropped by almost fifty percent the day after it announced that it would have to restate the previous year's earnings due to accounting irregularities. *Your Money*, L.A. TIMES, Apr. 17, 1998, at D4.

126. In March and April of 1997, the Dow Jones Industrial Average dropped nearly ten percent from its peak after the Federal Reserve raised short-term interest rates. *See* Tom Petruno, *Inflation Signs Drive Dow into 148-Point Skid*, L.A. TIMES, Apr. 12, 1997, at A1.

127. PRING, *supra* note 118, at 54.

128. *See* ROBERT D. EDWARDS & JOHN MAGEE, TECHNICAL ANALYSIS OF STOCK TRENDS 257-58 (7th ed. 1997).

129. "Human nature remains more or less constant and tends to reach to similar situations in consistent ways." PRING, *supra* note 118, at 2.

130. *See, e.g.*, *Dow Dogs Gain Momentum*, USA TODAY, Sept. 5, 1997, at 4B; *see also* TOM GARDNER & DAVID GARDNER, THE MOTLEY FOOL INVESTMENT GUIDE 84 (1996).

subset of the Dow Jones Industrial Average stocks with the highest dividend yields,[131] has produced almost a twenty-six percent annual return over the past twenty years.[132] Based on this past performance, this strategy is touted as a simple way to beat the market. Yet, any continued reliance on this strategy is predicated upon a belief that the past pattern of out-performance by high-yield Dow stocks will continue to manifest itself because the high yield of these stocks indicates that investors have overreacted negatively and that the stocks will recover.

This is an apt moment to point out a danger with backtesting (that is, dredging historical data for patterns): one can often find apparent correlations between unrelated phenomenon. According to the so-called "Super Bowl indicator," the stock market is supposed to have a down year whenever a team from the American Football Conference wins the Super Bowl.[133] Some investors believe that the stock market trend is linked to the length of women's hemlines.[134] What these amusing but silly "indicators" illustrate is that one must ask for some rational explanation to link cause and effect. In the case of technical analysis, the psychology of investors provides a reasonable link between the movement of stock prices in the past and their likely short-term future movement. Similarly, the Dogs of the Dow strategy would appear to foster success because it is a contrarian strategy that buys out-of-favor stocks when their yields are high and prices are low.[135]

4. Graphology

A final example of the utility of past historical patterns is graphology, the practice of handwriting analysis. Although graphology is often dismissed as belonging to the realm of numerology and Tarot cards, it "is slowly gaining acceptance as an employment tool to help determine who's right for a job or promotion . . . ."[136] Some American corporations and many large European firms ask job applicants for handwriting samples, which are then analyzed by graphologists.[137] It has

131. The dividend yield of a stock is the annual dividend rate of the stock divided by its share price. Thus, a fifty-dollar stock paying one dollar in dividends has a yield of 1/50, or two percent.

132. GARDNER & GARDNER, *supra* note 130, at 86.

133. *See, e.g.*, Allan Sloan, *Walstreet: Clash of the Indexes*, NEWSWEEK, Jan. 11, 1999, at 44. Note that in 1998, an AFC team, the Denver Broncos, won the Super Bowl, but the Dow Jones Industrial Average and the Standard and Poor's 500 Index rose 16.1 percent and 26.7 percent respectively. *Investor Spotlight*, L.A. TIMES, Jan. 1, 1999, at C7.

134. *See, e.g.*, Rose DeWolf, *Some Economic Prognosticators Keep a Nervous Eye on Women's Hemlines*, O.C. REG., Sept. 29, 1998, at E6.

135. *See* GARDNER & GARDNER, *supra* note 130, at 87.

136. Tyrone Beason, *Experts in Writing Lend Hand to the Boss*, SAN DIEGO UNION-TRIB., May 25, 1998, at D3.

137. CURTIS W. CASEWIT, GRAPHOLOGY HANDBOOK 2 (1980); *see also* Tim Dawson, *Graphologists Rewrite the Rulebook*, SUNDAY TIMES-LONDON, June 15, 1997 (half of French

also been used by trial lawyers to help select favorable jurors and by detectives to "understand the tendencies of suspected criminals."[138] In 1981, the Library of Congress classified graphology as a "behavioral science."[139]

For example, a person whose handwriting leans to the left is supposedly an extrovert.[140] Like yield management and technical analysis, graphology relies on generalizations drawn from past performance to evaluate an individual: "Analysts don't make judgments based on the handwriting of one person. Rather they try to identify recurring themes in many people's handwriting. Findings that show up regularly in handwriting analyses then become accepted assumptions for graphologists."[141] The theory is that handwriting begins in the brain, with signals going through the motor cortex and nerves to the hand; "[t]he result is an individual one, much like one's electrocardiogram."[142]

IV. False Positives, False Negatives

The fact that probabilistic evidence of the type embodied by the drug courier profile can have some probative value does not, by itself, justify its admission as substantive evidence at trial. To understand why, it is necessary to consider what scientific experimentation classifies as "Type I" and "Type II" errors. A Type I error is a "false positive," or "the detection of an effect that is not there." A Type II error is a "false negative," or "the overlooking of a real effect."[143] Thus, in a criminal trial, the conviction of an innocent person is a Type I error, while the acquittal of a (factually) guilty person is a Type II error.[144] Similarly, at the investigatory stage, the investigation of an innocent person is a Type I error, while the failure to investigate a guilty person is a Type II error.

Type I and Type II errors typically cannot be reduced at the same time, and in fact, are inversely linked.[145] As scholars note:

---

companies, ten percent of Italian and German companies use handwriting analysis before making appointments); Kurt Kragthorpe, *Some People View Analysis As a Science*, SALT LAKE TRIB., Dec. 25, 1997, at C1 (more than 6000 U.S. companies).

138. Beason, *supra* note 136.

139. Kragthrope, *supra* note 137.

140. Beason, *supra* note 136.

141. *Id.*; Dawson, *supra* note 137 (discussing "trait norms").

142. CASEWIT, *supra* note 137, at 5; *see also* Dawson, *supra* note 137.

143. KENNETH R. FOSTER & PETER W. HUBER, JUDGING SCIENCE 75 (1997).

144. *See* Kate Stith, *The Risk of Legal Error in Criminal Cases: Some Consequences of the Asymmetry in the Right to Appeal*, 57 U. CHI. L. REV. 1, 3 (1990).

145. *See, e.g.*, PAUL G. HOEL ET AL., INTRODUCTION TO STATISTICAL THEORY 55 (1971); Joseph Sanders, *From Science to Evidence: The Testimony on Causation in the Bendectin Cases*, 46 STAN. L. REV. 1, 15 & n.61 (1993).

> There is seldom a sharp line that separates unambiguously positive from unambiguously negative results. Instead, the designer of a test usually has to set a threshold for detection of a "positive" result. The higher the threshold, the more likely it is that a "positive" result is correct—but, at the same time, the more likely it is that the test will miss some real positive cases. If the threshold for detecting a disease is too high, a test may identify victims of the disease with few false alarms (that is, with a low rate of false positives.) But, by the same token, the test may overlook many real victims of the disease (that is, it may have a high rate of false negatives).[146]

In medicine, neither error is automatically worse than the other one. A false positive may inflict serious emotional distress. However, a false negative may create a false sense of security resulting in the failure to obtain timely treatment.

In criminal trials, the prosecution must prove the defendant's guilt beyond a reasonable doubt with respect to every element of the crime.[147] This requirement reflects society's view that, when criminal punishment is involved, a Type I error is far worse than a Type II error[148]—hence, the maxim, "It is better to let ten guilty people go free than to convict one innocent person."[149] In fact, society's abhorrence of Type I errors in criminal trials is so strong that once a defendant has tested "negative," so to speak, by being acquitted, we cannot retest that person again, even if we believe that a false negative has resulted.[150] Note that the converse is not true; a showing of actual innocence (i.e., a showing that a false positive has occurred) is sufficient to obtain relief in post-conviction proceedings that would otherwise be denied.[151] This is not to say that society is unwilling to tolerate any Type I errors at all, because reasonable doubt is not the same as absolute certainty.[152]

---

146. FOSTER & HUBER, *supra* note 143, at 75.

147. *In re* Winship, 397 U.S. 358, 364 (1970).

148. John Kaplan, *Decision Theory and the Factfinding Process*, 20 STAN. L. REV. 1065, 1073 (1968).

149. *See, e.g.*, Scott E. Sundby, *The Reasonable Doubt Rule and the Meaning of Innocence*, 40 HAST. L.J. 457, 460 (1989); *see generally* Speiser v. Randall, 357 U.S. 513, 525-26 (1958); Coffin v. United States, 156 U.S. 432, 455-56 (1950).

150. This, of course, stems from the protection of the Double Jeopardy Clause of the Fifth Amendment.

151. *See* Schlup v. Delo, 513 U.S. 298, 313-16 (1995); Murray v. Carrier, 477 U.S. 478, 496 (1986) (holding that a showing of actual innocence would be sufficient to excuse satisfaction of the procedural default standard).

152. *See, e.g., Winship*, 397 U.S. at 370; Commonwealth v. Costley, 118 Mass. 1, 24 (1875); Kaplan, *supra* note 148, at 1071 ("But because no case is doubt free, unless we decide to avoid trying anyone, we will, if we try enough people, inevitably convict an innocent man.").

Incidentally, this understanding of the impossibility of eliminating all errors helps to explain *Herrera v. Collins*,[153] a case that might otherwise appear monstrous.[154] In *Herrera*, the Court engaged in an interesting debate on whether the Constitution permitted the execution of an "innocent" person who has been convicted in a procedurally and substantively fair trial—that is, one in which no legal mistakes occurred. Although the Court ultimately sidestepped the question, Chief Justice Rehnquist and Justices Scalia and Thomas appeared prepared to say that the Constitution does not forbid the execution of such a person.[155] Six Justices were convinced that the execution of an innocent person would violate the Constitution.[156]

The opinions in *Herrera* dance around the notion that the determination of guilt or innocence is a probabilistic inquiry. Other than the real killer (who may or may not have been Herrera), no one knows with absolute certainty who committed the murders. Herrera's submission of new evidence of his innocence (principally in the form of a death-bed admission by a relative) does not prove his innocence; it only makes it seem less likely now that he is guilty than based on what evidence was presented at trial. Thus, the real question is not whether the Constitution permits the execution of an innocent person; it is rather, what does one have to show to be deemed sufficiently likely to be innocent so as to escape the death penalty? From this standpoint, it becomes clear that a majority of the Court (Rehnquist, Scalia, Thomas, O'Connor, Kennedy, and White) believed that a state could constitutionally impose the death penalty without having to eliminate all possibilities of a Type I error.

By minimizing the risk of a Type I error in criminal trials, however, the reasonable doubt standard increases the risk of a Type II error. Without improving the ability of the jury to resolve the question of a defendant's guilt, the only way to reduce the risk of a Type II error would be to lower the standard of proof. But doing that in turn would increase the risk of a Type I error.

For example, two juries reached conflicting verdicts in O.J. Simpson's criminal and civil trials for allegedly murdering his ex-wife and her friend.[157] These verdicts may in fact be consistent in the sense

153. 506 U.S. 390 (1993).

154. *See, e.g.*, Susan Bandes, *Simple Murder: A Comment on the Legality of Executing the Innocent*, 44 BUFF. L. REV. 501, 502 (1996).

155. *Herrera*, 506 U.S. at 417 (assuming without deciding that while "a truly persuasive demonstration of 'actual innocence' made after trial would render the execution of a defendant unconstitutional . . . the threshold showing for such an assumed right would necessarily be extraordinarily high").

156. *See id.* at 419 (O'Connor and Kennedy, JJ. concurring); *id.* at 429 (White, J., concurring); *id.* at 430 (Blackmun, Stevens, and Souter, JJ. dissenting).

157. Stephanie Simon, *Simpson Liable in Slayings, Compensatory Damages Put at $8.5 Million*, L.A. TIMES, Feb. 5, 1997, at A1; Jim Newton, *Simpson Not Guilty; Drama Ends 474 Days After Arrest*, L.A. TIMES, Oct. 4, 1995, at A1.

that they reflect the differing burdens of proof—that the evidence presented in both trials proved Simpson's liability beyond a preponderance of the evidence, but not beyond a reasonable doubt.[158] Nevertheless, it is clear that either a Type I or Type II error has occurred—if Simpson did murder the victims, he has escaped punishment for a criminal act; if he did not, then he should not be laboring under a $33.5 million judgment.[159]

At the investigatory stage, however, the trade-offs between Type I and Type II errors are considerably different from those at trial. A false positive during the investigation simply means that the police investigate a factually innocent person. Admittedly, being investigated can impose significant burdens. So long as the police have probable cause, the factually innocent are subject to searches pursuant to warrants,[160] warrantless arrests,[161] and the filing of criminal charges against them.[162] Once probable cause has been established, the defendant can be subjected to the additional burdens and costs of standing trial.

A false postive at the pretrial stage does result in costs. However, the implications of a false negative are significant as well. A factually guilty person escapes punishment, and law enforcement authorities may focus their investigation on a different person, someone who is factually innocent. Moreover, while there is often a sense that the system worked when a person whom the police believe is factually guilty is nevertheless acquitted,[163] it is hard to see a similar sense of vindication of the system when the police fail to investigate suspects.

The fact that Constitution regards Type II errors as more serious than Type I errors at the pretrial stage is reflected in the lower evidentiary standards required for detaining suspects for questioning, obtaining search or arrest warrants, or indicting a suspect than for convicting him or her. The standard for most civil trials, preponderance of the evidence, reflects the point at which Type I errors are deemed as equally bad as Type II errors; hence, the plaintiff need only prove that its case is more likely to be true than not.[164] Because probable cause is a

158. *See, e.g.*, ALAN M. DERSHOWITZ, REASONABLE DOUBTS: THE CRIMINAL JUSTICE SYSTEM AND THE O.J. SIMPSON CASE 39 (1996).

159. *See* Henry Weinstein, *Size of Punitive Damages Is Justified, Legal Experts Say*, L.A. TIMES, Feb. 11, 1997, at A19.

160. *See* Illinois v. Gates, 462 U.S. 213, 238-39 (1985).

161. *See* United States v. Watson, 423 U.S. 411, 416 (1975).

162. *See* Bordenkircher v. Hayes, 434 U.S. 357, 364 (1978).

163. *See, e.g.*, DERSHOWITZ, *supra* note 158, at 16; W. William Hodes, *Lord Brougham, the Dream Team, and Jury Nullification of the Third Kind*, 67 U. COLO. L. REV. 1075, 1077 (1997). *But see* Alexander M. Bickel, *Judge and Jury—Inconsistent Verdicts in the Federal Courts*, 63 HARV. L. REV. 649, 653 (1950) (arguing that the acquittal of a factually guilty person causes a loss in confidence in the system).

164. *See* Neil B. Cohen, *Confidence in Probability: Burdens of Persuasion in a World of Imperfect Knowledge*, 60 N.Y.U. L. REV. 385, 403 (1985); Kaplan, *supra* note 148, at 1072.

less stringent standard (and reasonable suspicion even lesser),[165] it is clear that at the investigatory stage, where these standards control, Type II errors are worse than Type I errors.

Additionally, probable cause may be based on evidence that would be inadmissible at trial, such as hearsay.[166] Indeed, an indictment may be based entirely on hearsay evidence.[167] This leniency is particularly striking, because one of the primary reasons that hearsay is inadmissible is its perceived lack of reliability compared to in-court testimony.[168]

Thus, when a Type I error is to be avoided at all reasonable costs, evidence whose reliability is questionable may be excluded altogether. But when a Type II error is as significant or more so, such evidence may be considered. In fact, one can draw an imprecise analogy to the scope of discovery allowed under the Federal Rules of Civil Procedure. During pretrial discovery proceedings, parties are given great latitude to seek evidence from the opposing side—including evidence that would not be admissible at trial—so long as the evidence is reasonably likely to yield admissible evidence.[169] This rule also reflects an underlying principle that, during pretrial proceedings, the erroneous discovery of evidence (i.e., evidence that is not admissible and that fails to lead to the discovery of admissible evidence) is preferable to the erroneous withholding of evidence, much more so than at trial.

V. Pitfalls

Given that, at the trial stage, Type I errors—false positives—are to be avoided as much as is reasonably possible, the question to be considered is whether the probative value of the drug courier profile is substantially outweighed by its problems. As I noted earlier, the courts that have excluded profile evidence have generally done so under the "unfair prejudice" prong of Rule 403 of the Federal Rules of Evidence.[170] There are, however, other problems with the use of profile evidence at trial that have not been touched upon, and those are the "confusion of the issues" and "considerations of undue delay, waste of time, or needless

165. Probable cause is the standard that governs most pre-trial proceedings such as obtaining a search or arrest warrant. Reasonable suspicion is the standard that governs the investigative stops pursuant to *Terry v. Ohio*. 392 U.S. 1, 27 (1968).

166. *See* Franks v. Delaware, 438 U.S. 154, 165 (1978); Aguilar v. Texas, 378 U.S. 108, 114 (1964), *overruled on other grounds*; Illinois v. Gates, 462 U.S. 213, 238 (1983); FED. R. CRIM. P. 41(c)(1) ("The finding of probable cause may be based upon hearsay evidence in whole or in part.").

167. Costello v. United States, 350 U.S. 359, 363-64 (1956).

168. *See, e.g.*, Michael L. Siegel, *Rationalizing Hearsay: A Proposal for a Best Evidence Hearsay Rule*, 72 B.U. L. REV. 893, 898 & n.23 (1992).

169. FED. R. CIV. P. 26(b)(1).

170. *See supra* notes 43-46, and accompanying text.

presentation of cumulative evidence" prongs of Rule 403, not to mention the threshold question of relevance.

A. Determining the Probative Value of Profiles

The initial issue in determining the admissibility of a piece of evidence is whether it is relevant; that is, whether it alters the likelihood (either positively or negatively) of the existence of a material fact in dispute.[171] Previously, I demonstrated why the drug courier profile, in theory, could have enough predictive value to justify its use as an investigative tool. That justification does not automatically translate to a similar probative value in the trial of a suspected drug courier. In fact, there is one way in which the profile may not actually even be relevant at a trial.

1. The Problem

Suppose that a defendant is tried for robbery; identity is the only issue. An eyewitness testifies that she saw a white car speed away from the crime scene, and she is certain that the car's license plate had an "A," an "O," and a "K," in that order. During closing argument, the prosecutor says:

> Ladies and gentlemen of the jury, Ms. W testified that the license plate of the getaway car had the letters AOK in that order. The defendant's car (pointing at the defendant) just happens to have the license plate 123AOK. Is that a coincidence? Yeah, right. The odds that Ms. W could have guessed that the plate would have AOK in that order are 17,576 to one. Ms. W must have seen the defendant's car!

Is this argument legitimate? The probability that a license plate with three numbers and three letters would contain the sequence "AOK" is indeed one in 17,576 or 0.000057. The value of each letter is independent from the other, meaning that the fact that the first letter is "A" has no bearing on the value of the second letter.[172] An expert witness would probably have to perform this calculation, but the problem is of trivial difficulty. Where events are independent, the probability that

171. FED. R. EVID. 401.
172. *See* WEAVER, *supra* note 98, at 111.

they will all occur is the product of the individual probabilities.[173] The probability that the first letter is "A" is 1/26, as is the probability that the second letter is "O," and so on. Thus, the probability that a license plate has that sequence (or any particular three-letter sequence) is 1/26 * 1/26 * 1/26, or 1/17,576.[174] If the probability has been calculated correctly, it would seem as if the prosecutor is correct—Ms. W must have seen the defendant's car.

There is, of course, a problem. It is true that the probability that Ms. W could guess that the defendant's license plate has the sequence "AOK" is indeed 0.000057. However, this does not prove at all that it is 99.9943 percent certain that Ms. W saw the defendant's car at the crime scene. Suppose Ms. W in fact was mistaken about the license plate of the getaway car. She believes that she saw "AOK," but the license plate of the getaway car is really "ADK." When the police arrive at the scene of the crime, Ms. W tells them the getaway car was a white sedan with the license plate "AOK." The police access the Department of Motor Vehicles database for white cars with "AOK" in the license plate, which is how they find the defendant.

In such an instance, the 0.000057 probability is mathematically correct, but does not support the inference that the prosecutor is attempting to draw. To be sure, there will undoubtedly be some other evidence in the case tying the defendant to the crime, as it would seem extremely unlikely that a prosecutor would bring a case based solely on one eyewitness' identification of a license plate. But one can easily imagine a case where there is circumstantial evidence that, without the license plate identification, falls just short of proof beyond a reasonable doubt, but exceeds that level with the identification. In such a case, admitting the license plate identification is problematic even though the probability estimate is accurate.

2. *As Applied to Drug Courier Profiles*

It is tempting to simplify this problem by noting that, in the case of a suspected drug courier, the defendant not only was identified as a potential courier by the profile, but also was in fact carrying narcotics. Therefore, it might seem that the appropriate question is: Given that the defendant fit the profile and was carrying narcotics, how likely is it that the defendant was a courier? The problem with this approach is the fact that the defendant fit the profile and the fact that he or she was caught

173. *Id.*

174. Note that the probability would be six times larger if the witness did not recall the specific order of the letters, since there are six ways to arrange a three-letter sequence: AOK, AKO, OAK, OKA, KAO, KOA. *See id.* at 112.

carrying narcotics are not independent. In other words, this is analytically similar to the problem involving the use of the probability that a witness could guess three letters on the defendant's license plate at random; since the police traced the defendant's car based on the identification, the identification in fact strongly correlates to the defendant's car.

Of course, most persons who are caught with narcotics are probably in fact couriers. Some of them, however, may actually be innocent, having come into possession of the drugs unwittingly.[175] Yet, the drug courier profile does not distinguish between the guilty and the unwitting—that is the role of the criminal justice system. In other words, the fact that the defendant fit the courier profile is the reason that he or she was stopped in the first place. To use that fact as evidence of substantive guilt, however, is something like using the fact that the defendant has been indicted as evidence of guilt. Naturally, the indictment cannot be considered in this manner.[176]

It is helpful to think of proceedings before the grand jury and the trial as different kinds of tests designed to answer the same question: "Did the defendant commit this crime?" Trials, of course, are considered more reliable and accurate—at least in terms of avoiding Type I false positives. However, they are also expensive for society, which must pay for the costs of the trial and the prosecution, and in some cases, defense counsel; and for the defendant, who loses time from work, must pay for counsel if he or she can afford it, and so on. Rather than subject every suspect to a trial, society has adopted a two-tier process. The first stage is the grand jury proceeding. If the defendant tests "negative" (i.e., is not indicted), the matter is concluded. If the defendant tests "positive," then the matter continues to the second stage—the trial, which is considerably more refined and accurate.

There are non-legal analogs to this analysis. The standard medical test used to determine if a person has been infected with HIV is nearly 100 percent accurate but takes about a week to process.[177] Recently, the Centers for Disease Control and Prevention began advocating the widespread use of a faster, though less accurate test. Because it is less accurate, the newer test is administered three times. If a person tests positive for HIV once or more under this fast test, the standard one-week test is then used to confirm the result.[178] The advantage of the fast test is that many more people can be tested and cleared than if the traditional test were used from the start.

175. *See* Kadish, *supra* note 2, at 749 n.5 (citing examples).

176. *See, e.g.*, United States v. Utz, 886 F.2d 1148, 1151 (9th Cir. 1989).

177. Chelsea J. Carter, *Government Urges Use of Fast HIV Test*, SAN DIEGO UNION-TRIB., Mar. 27, 1998, at A2.

178. *Id.*

Similarly, there are two tests for comparing DNA strands. The older, more established method is known as "restriction fragment length polymorphism analysis" (RFLP), or the "Southern Blot" test.[179] The newer test is known as "polymerase chain reaction" (PCR).[180] Of the two, PCR is the "quick and dirty" version; results come back as quickly as twenty-four hours after testing, but are subject to a greater margin of error and are less informative.[181] RFLP test results, on the other hand, take several weeks to develop.[182] Since DNA tests do not indicate "unequivocally that a crime suspect is the source of an unknown sample of DNA[,]" but rather that the suspect "cannot be excluded as a possible source,"[183] it does not make sense to use an inconclusive result of the less accurate PCR test to interpret the more accurate RFLP test.

Thus, even if the predictive accuracy of the drug courier profile could be established without an inordinate amount of collateral testimony, the inference to be drawn from the profile may not be supportable because the usefulness of the profile has already been exhausted. To introduce it as evidence is to bootstrap the defendant's arrest into evidentiary significance.[184]

B. Calculating Probabilities Properly

A second difficulty in introducing the drug courier profile at trial involves validation of the underlying foundations of the profile. Since the drug courier profile is really nothing more than probabilistic evidence, the defendant should be entitled to test in court, the mathematical basis of the profile. Probability theory is fairly complex and highly susceptible to incorrect intuition. Validating the mathematical basis of the profile would require substantial amounts of testimony for evidence that offers some, but not significant probative value.

---

179. George Bundy Smith & Janet A. Gordon, *The Admission of DNA Evidence in State and Federal Courts*, 65 FORDHAM L. REV. 2465, 2468 (1997).

180. *Id.* at 2470. PCR was invented by Kary Mullis in 1993, a feat for which he won the Nobel Prize in Chemistry. *See generally* KARY MULLIS, DANCING NAKED IN THE MIND FIELD 3-15 (1998).

181. Smith & Gordon, *supra* note 179.

182. *Id.* at 2471.

183. *Id.* at 2472.

184. Note that bootstrapping does not occur where the profile is introduced as rebuttal evidence, such as in United States v. Beltran-Rios, 878 F.2d 1208, 1212-13 (9th Cir. 1989), where the defense attempted to show that the defendant did not exhibit the tendencies of drug couriers. In that instance, as the *Beltran-Rios* court perceived, the profile was not introduced merely to comment on the meaning of the government's evidence, but to rebut an inference the defendant sought to draw.

1. The Problem

The Supreme Court's analysis of polygraph evidence in *United States v. Scheffer*[185] demonstrates how the validation of a test procedure can spawn unnecessary collateral litigation, thereby justifying the exclusion of evidence relating to that procedure. A polygraph examination, known more conventionally as a "lie detector test," relies on the interpretation of the subject's physiological responses to a series of questions to form an opinion as to whether the subject is being truthful.[186] The Court noted that, "Allowing . . . polygraph evidence would inevitably entail assessments of such issues as whether the test and control questions were appropriate, whether a particular polygraph examiner was qualified and had properly interpreted the physiological responses, and whether other factors such as countermeasures employed by the examinee had distorted the exam results."[187] Moreover, this sort of collateral litigation would take place in every case in which polygraph evidence is introduced.[188] Based partially on that analysis, the Court held that a per se ban on polygraph evidence did not violate a defendant's Sixth Amendment right to present a defense.[189]

Scheffer thus highlights at least two areas of concern in the admission of polygraph results: (1) was the test administered properly (e.g., were proper questions asked); and (2) were the results interpreted properly (e.g., was a sign of nervousness correctly linked to deception)? Based on *Scheffer*'s reasoning, there at least two distinct avenues for mathematical validation of drug courier profiles. First, have the probabilities for the profile been calculated properly? As I shall demonstrate shortly, this is far from a foregone conclusion, for even mathematicians have been known to stumble on probability problems. We should be less sanguine that law enforcement agents have had the type of formal mathematical background to produce reasonably accurate results. Second, even if the probabilities have been calculated correctly, are they being used properly?

Recall that Professor Tribe's hypothetical regarding the mistress with the murdered man in her apartment suggested that the estimated probability that this mistress was the murderess was ninety-five percent. How would this probability be calculated? This is a crucial threshold issue for the use of probabilistic evidence.[190]

---

185. 523 U.S. 303 (1998).
186. *Id.* at 313 n.9.
187. *Id.* (Thomas, J.).
188. *Id.* at 314.
189. The defendant in *Scheffer* had passed a polygraph examination and sought to admit the test result in his trial. *Id.* at 305-306.
190. *See* Mark L. Huffman, Comment, *When the Blue Bus Crashes Into the Gate: The Problem with* People v. Collins *in the Probabilistic Evidence Debate*, 46 U. MIAMI L. REV. 975, 981 (1992).

There are, in fact, three different methods of calculating probabilities. Many academic problems in probability theory comprise what one might call "logical proportion" probabilities, in the sense that the probability is expressed as the ratio of the number of desired outcomes to the total number of possible outcomes.[191] For example, a typical such problem might ask for the probability of drawing a heart or a queen from a standard deck of playing cards. There are sixteen cards that satisfy the criteria,[192] out of a total of fifty-two cards, for a probability of 16/52, or 4/13. This type of probability analysis is not applicable to the mistress example or drug courier profiles, however, because we are not dealing with a closed set such as a deck of playing cards.

A second method of estimating probabilities consists of "frequency analysis." In this process, "the probability of a given outcome in some activity is calculated by observing the frequency with which that outcome occurs; the probability value represents the fraction of times that the particular outcome would be expected to occur in a very large or infinite number of repeated trials of the activity."[193] For example, a gambler at a Las Vegas casino might play the same slot machine 1,000 times in a row and notice that the cherry-cherry-cherry combination showed up twelve times. The gambler would then estimate the probability of hitting the cherry-cherry-cherry combination on a single spin of the slot machine at 12/1000.

The final method of estimating probabilities is a "subjective belief" approach, generally equivalent to a person's degree of belief that an event has or will have occurred.[194] Strictly speaking, this method has little, if any, basis in mathematics.[195] Yet, it has significant support in the legal community (at least, among legal scholars who favor the use of probability theory in trials) on the ground that the other two methods—logical proportion and frequency analysis—are not suited for unique factual situations.[196]

For a concrete illustration of the problems inherent in the calculation of probabilities for use as evidence, consider the (in)famous case of *People v. Collins*.[197] *Collins* was a simple robbery case where identity was the key issue. Because the prosecutor "experienced some difficulty in establishing the identities of the perpetrators of the crime,"[198] he resorted to the use of probability theory to "calculate" the

191. *See* Louis J. Braun, *Quantitative Analysis and the Law: Probability Theory as a Tool of Evidence in Criminal Trials*, 4 UTAH L. REV. 41, 44 (1982).

192. There are thirteen cards with hearts and four queens, but one of the queens is the Queen of Hearts, which cannot be counted twice.

193. Cohen, *supra* note 84, at 391.

194. Braun, *supra* note 191, at 48.

195. *Id.* at 49.

196. *Id.*

197. 438 P.2d 33 (Cal. 1968).

198. *Id.* at 36.

likelihood that there was another couple in Los Angeles who shared the characteristics of the couple being tried.

The victim and another witness testified essentially that the robber was a blonde woman with a ponytail, who escaped in a yellow car driven by a Black man with a mustache and beard.[199] The prosecutor assigned the following probabilities to these traits:

1) yellow car	1/10
2) man with mustache	1/4
3) woman with ponytail	1/10
4) blonde woman	1/3
5) Black man with beard	1/10
6) interracial couple	1/1000[200]

A mathematician witness then calculated the probability that a couple would share all six characteristics was the product of the individual probabilities—in other words, 1/10 * 1/4 * 1/10 * 1/3 * 1/10 * 1/1000, or 1/12,000,000.[201] Significantly, the prosecutor provided no statistical evidence to support the accuracy of these probabilities.[202] As a result, the California Supreme Court held that the prosecutor's line of argument was fatally undermined by the lack of a proper foundation for the witness' testimony.[203] The court actually went beyond this holding and discussed two independent bases that justified exclusion of the evidence, neither of which, however, is relevant to this article.

a. Probability Estimation

It is worth asking how (if at all) the prosecutor could have laid a proper foundation for introducing these probabilities. In other words, let p_{yc} designate the probability that a car in that part of the Los Angeles area would be yellow. How can we be sure that $p_{yc} = 0.1$? If the actual value of $p_{yc} = 0.2$ (meaning that yellow cars were twice as common as the prosecutor estimated), then the probability of fitting all six characteristics would drop to 1/6,000,000. That drop might not be enough to make a difference by itself, but it would significantly affect the probative value of the evidence. A few more inaccuracies could drop the probability to something far less astronomical.

199. *Id.* at 34.
200. *Id.* at 37 n.10.
201. *Id.* at 37.
202. *Collins*, 438 P.2d at 36.
203. *Id.* at 39.

It is clear that p_{yc} cannot be calculated using logical proportion analysis, since there is not an easily quantifiable closed set such as there is with a deck of cards. It may be that ten percent of all cars in the United States are yellow, but this fact would be of questionable relevance if the percentage of cars in California—or Southern California—is significantly greater or fewer than ten percent. Even Los Angeles may not be the proper domain if, for example, many Southern Californians do not drive into San Pedro, where the robbery occurred.

Even frequency analysis may not be suitable for calculating the value of p_{yc}, since it is not at all clear what properly constitutes the repeated event to observe. Suppose that the prosecutor produces a witness who testifies that she sat at the nearest intersection to the scene of the robbery and wrote down the number of yellow and non-yellow colored cars she saw drive by every day for a month, and that ten percent of the cars were yellow. Would this be sufficient to establish $p_{yc} = 0.1$? What if the robbery occurred on a Monday, and the percentage of yellow cars on Monday was actually different from that on other days?

For the moment, it may be easier to recharacterize this problem as determining whether a given coin is fair. For a fair coin, the probability that a toss will result in heads, as represented by p_h, is 0.5. Suppose you toss the coin twice and it comes up heads once and tails once (i.e., either HT or TH). Does this prove that $p_h = 0.5$? Of course not. Even if the coin were biased so that the true value of $p_h = 0.75$ (that is, the coin can be expected to turn up heads seventy-five percent of the time), the probability of yielding HT or TH is still 0.375.[204]

Suppose instead that you toss the coin 1000 times, and the coin turns up heads 510 times. Does this prove that $p_h = 0.51$? No. The estimate that $p_h = 0.51$ is more likely to be accurate than if it had been tossed only 100 times, but note that the ninety-five percent confidence interval for a fair coin (i.e., $p_h = 0.50$) tossed 1000 times is $\pm$ 0.031.[205] This means that if you conduct multiple experiments of tossing a coin 1000 times, the number of times that heads comes up will be between 469 and 531 ninety-five percent of the time.

It may be sufficient for the purpose at hand to conclude that the coin is "probably" fair if it turns up heads 510 times out of 1000, with the knowledge that this result is well within the ninety-five percent confidence interval for $p_h = 0.50$. The acceptable error margin is, of course, dependent on the context of the experiment. If you are betting only one dollar, five heads in ten tosses may be enough for you to accept the coin as fair. But if you are betting something more significant, you will probably want to see quite a few more tosses. This observation

204. The probability of HT is simply $(p_h)(p_t)$, or (.75)(.25) = .1875. Similarly, the probability of TH is $(p_t)(p_h)$, or (.25)(.75) = .1875. Because these are mutually exclusive outcomes, their probabilities can be summed to find the probability of one or the other.

205. Kaye, *supra* note 102, at 60-61.

reflects "second-order uncertainty," which Jonathan Koehler and Daniel Shaviro define as "uncertainty about uncertainty."[206] They explain that two juries may both believe there is a sixty percent chance that the plaintiff in each case should recover, but where the first jury has heard extensive evidence and has a strong grasp of the case, the second jury in contrast, has been presented with very little evidence. The second jury will have much less confidence in its assessment of the probability that the plaintiff should win.[207]

Now return to the problem of determining the value of p_{yc} (the probability that a car is yellow). The most that can be said is, if we assume that the true value of $p_{yc} = 0.10$, the observation by the witness over the specified time period may be within the ninety-five percent confidence interval. We cannot go in the other direction, however, and construct a ninety-five percent confidence interval around the observed ratio of yellow cars to all cars and thereby conclude that the true value of p_{yc} is within that confidence interval.[208] This is a weakness of the frequency analysis approach.

These are formidable obstacles to accurate calculations of probabilities in real-life situations through the frequency analysis method. Even if the probabilities have been calculated accurately, conveying that determination would require a substantial amount of collateral testimony relating to the observations and whether they create the proper foundation for the probability in question.

That leaves the subjective estimate for calculating p_{yc}. This is an especially dangerous approach to use in criminal trials, as scholars note, "Any probabilistic intuition by anyone not specifically tutored in probability calculus has a greater than fifty percent chance of being wrong."[209] Indeed, probability and statistics are deceptively difficult because they appear to be understandably simple, when in fact they are not.[210] A layperson confronted with a problem in advanced calculus

206. Koehler & Shaviro, *supra* note 88, at 251.

207. *Id.*

208. *See supra* notes 98- 103 and accompanying text.

209. FOSTER & HUBER, *supra* note 143, at 119 (quoting MASSIMO PIATTELLI-PALMARINI, INEVITABLE ILLUSIONS: HOW MISTAKES OF REASON RULE OUR MINDS (1994)).

210. *See* AMIR D. ACZEL, PROBABILITY 1, at 3 (1998). Deceptive difficulty is one reason the math section on the Scholastic Aptitude Test (SAT) can differentiate among students' mathematical abilities without involving any concepts above Euclidean geometry. For example, one question might ask:

> If Smith drives to work at forty miles per hour and drives home at sixty miles per hour, what is Smith's average speed that day?

To the untrained mind, the answer would appear to fifty miles per hour, that being the arithmetic average of forty and sixty. However, the average speed is a weighted average, and Smith spends more time driving forty miles per hour than at sixty miles per hour, since it takes longer to cover the same distance (from home to work) at the slower speed. Suppose Smith lives 120 miles away from work. It would take Smith two hours of driving at sixty miles per hour, but three hours at forty miles per hour. Smith drives a total of 240 miles in five hours, which works out to forty-eight miles per hour.

may simply realize that he or she lacks the mathematical background to offer a solution. Thus, such a person confronted with the following problem most likely will have no intuitive feeling: "Give a mathematical argument to show that a heated wire in the shape of a circle must always have two diametrically opposite points with the same temperature."[211]

This self-restraint may not occur with probabilistic inquiries. Consider the following probability problems, all of which, like the wire problem above, are phrased in plain English. Unlike the wire problem, however, they are problems that are far more challenging than they appear. (The solutions to these puzzles are found in Appendix 1.)

The Birthday Problem

How many people must there be in a group so that there is a fifty percent probability that two persons share the same birthday? The obvious guess is 183, which is just above half of 365.[212] It turns out, however, that in a group of twenty-three people, there is a 50.7 percent chance that two of them will share the same birthday.[213] This calculation assumes a perfectly even distribution of birthdays across all days and ignores leap year birthdays of February 29. In fact, in a group of fifty people, there is a ninety-seven percent chance that two of them share the same birthday,[214] a surprisingly non-intuitive result.

The Monty Hall Puzzle

The so-called "Monty Hall" puzzle appears so simple, and yet even mathematics professors have been fooled by it. The set-up follows the famous television game show, "Let's Make a Deal," and posits that Monty Hall (the host) presents you with three doors, marked one, two, and three. He tells you that behind one of the doors is $100,000 (or some suitable prize), and behind the other two doors is nothing. You are to choose one of the three doors. Once you have chosen a door, Hall tells you that he will open one of the two doors you did not choose and show you that it is empty. He then gives you the option of either sticking with the door you chose initially, or switching to the door that he did not

211. R. CREIGHTON BUCK, ADVANCED CALCULUS 97 (3d ed. 1978).

212. *See* JOHN ALLEN PAULOS, INNUMERACY 27 (1988); *see also* K.C. COLE, THE UNIVERSE AND THE TEACUP: THE MATHEMATICS OF TRUTH AND BEAUTY 146 (1998).

213. WEAVER, *supra* note 98, at 134.

214. *Id.* at 135.

reveal. Are you more likely to win if you switch doors, or does it make no difference?[215]

Surprisingly, you should always switch doors. It may appear that the probability of picking the door with the prize starts at 0.33 and improves to 0.50 after Hall opens one of the doors you did not pick. That is not the case, however, because Hall does not pick a random door to open; rather, he acts with knowledge of the location of the prize that you did not have access to when you picked a door initially.

When columnist Marilyn vos Savant first discussed this problem and asserted that one should always switch, among those who (incorrectly) sharply criticized her analysis were mathematics professors.[216] Indeed, one of the most renowned mathematicians of the twentieth Century, Paul Erdos, believed incorrectly that it made no difference whether you switched. If probability theory sometimes can befuddle even the mathematically trained, we should be especially hesitant to assume that the non-mathematically trained will understand and apply it properly.

The Card Game

Much like a shell game, which relies upon physical sleight of hand, there is a swindle involving cards that uses mathematical sleight of hand to cheat the player. The con artist shows you three cards; one is red on both sides, one is black on both sides, and the last one is red on one side and black on another side. The con artist drops the cards in a hat and you pick one without looking in the hat. You are to look at one side of the card. Suppose it's black. The con artist tells you that you know this card is not the red-red card and offers to bet you straight up that it is the black-black card. Should you accept?[217]

It would seem as if the bet is fair, since the card could be the red-black card or the black-black card and it would appear as if either is equally likely. However, the card is actually twice as likely to be the black-black card as it is to be the red-black card. [218]

The Birthday Problem, the Monty Hall Paradox, and the Card Game demonstrate that, where probability calculations are concerned, intuitive estimates can be strikingly inaccurate. This suggests in turn that mathematical validation of the drug courier profile cannot rest on the

215. John Tierney, *Behind Monty Hall's Doors: Puzzle, Debate and Answer?*, N.Y. TIMES, July 21, 1991, at 1.

216. *See* PAUL HOFFMAN, THE MAN WHO LOVED ONLY NUMBERS 235-37 (1998).

217. *See* PAULOS, *supra* note 212, at 64.

218. *Id.* at 64-65.

mere subjective beliefs of law enforcement agents. Instead, a more detailed mathematical analysis of the drug courier profile is required.

b. Independence

As if the complexities of estimating the individual probabilities are not enough, there is a second problem with the prosecutor's analysis in *Collins*. The probability that a couple would share all six characteristics can be determined by multiplying the probabilities of the individual characteristics only if the characteristics are mathematically independent, meaning that "the outcome of one event has no influence on the outcome of the other."[219] As the court noted in *Collins*, however, there is no guarantee that the characteristics set forth by the prosecutor are in fact independent.[220] Note, for example, that whether a man has a mustache and whether a Black man has a beard are probably not independent, since a man with a mustache may be more likely than one without a mustache to have a beard.[221]

Establishing that characteristics are independent is likely to be a very difficult task. It is the reverse of what epidemiology attempts to prove. Epidemiology is "the branch of applied statistics that studies the determinates and correlates of human disease."[222] Epidemiological studies attempt to measure the association between an observed medical illness and potential causes—that is, to show that they are not independent, but rather that a given factor causes a disease.[223] For example, the underlying question in the landmark case of *Daubert v. Merrill Dow Pharmaceuticals, Inc.*,[224] was whether the drug Bendectin, which had been prescribed as treatment for pregnancy-related morning sickness, caused birth defects. A number of medical studies indicated some association between the two, but it was possible to attribute far too much significance to the association and jump to the conclusion that Bendectin caused birth defects.[225] To date, despite dozens of studies of Bendectin, there is still no consensus as to whether the use of the drug causes birth defects.[226]

---

219. *Id.* at 20.

220. People v. Collins, 438 P.2d 33, 39 (Cal. 1968).

221. *Id.* at 39 n.15.

222. FOSTER & HUBER, *supra* note 143, at 150 (internal quotations omitted).

223. Essentially, epidemiologists observe a statistical correlation between a factor and a disease and then attempt to derive a biological understanding of the causation path involved. *See* DAVID E. LILIENFELD & PAUL D. STOLLEY, FOUNDATIONS OF EPIDEMIOLOGY 12 (3d ed. 1994).

224. 509 U.S. 579 (1993).

225. *See generally* FOSTER & HUBER, *supra* note 143, at 2-16.

226. *See id.* at 7 ("The large scientific record available today simply confirms what most epidemiologists would have inferred from the smaller scientific record a decade or two ago: If Bendectin poses any risks at all, they are very small.").

The failure to demonstrate causation, however, does not prove that Bendectin does not cause birth defects. Rather, the criticisms of the studies introduced in *Daubert* focused on their indeterminacy. It is analogous to the fact that while an acquittal in a criminal trial means that the prosecution has failed to prove the defendant's guilt beyond a reasonable doubt, it does not mean that the defendant has proven his or her innocence.[227]

Accordingly, the fact that it is often exceedingly difficult to show that a connection *exists* between two factors also suggests that it may be equally difficult to show that *no* connection exists between two factors.

2. *As Applied to Drug Courier Profiles*

People v. Collins illustrates two complex problems that must be solved before probabilistic evidence can be validated. These problems—estimation of probabilities and establishment of independence—plague the drug courier profile as well.

a. *Probability Estimation*

The prior analysis of *Collins* suggests that it would be difficult to produce a methodology for estimating the probability that a person fitting, say, all seven primary characteristics of the drug courier profile is indeed a courier.

To use the frequency analysis theory of probability, one would need to compare the number of persons fitting all seven characteristics who are couriers to the total number of persons fitting all seven characteristics.[228] Unfortunately, at present, the DEA's record keeping fails to come close to allowing this sort of calculation. The DEA does not even keep track of the number of persons it approaches who turn out not to be carrying drugs,[229] much less the total number of persons it observes (including those who it chooses not to stop) who fit the profile. Thus, the drug courier profile cannot be validated through the use of frequency analysis.

Even if the DEA were to keep accurate track of the statistics regarding the number of persons who fit the profile, there would still be

227. *See* United States v. Watts, 519 U.S. 148, 155 (1997) ("[A]n acquittal is not a finding of any fact. An acquittal can only be an acknowledgment that the government failed to prove an essential element of the offense beyond a reasonable doubt.") (quoting United States v. Putra, 78 F.3d 1386, 1394 (9th Cir. 1996) (Wallace, C.J., dissenting)).

228. *See supra* notes 193 and accompanying text.

229. Becton, *supra* note 16, at 418 n.4.

validation problems. The probability $p_{dc} = N/P$ where p_{dc} is the probability that a person fitting the profile is a courier, N is the number of persons fitting the profile who turn out to be carrying drugs, and P is the number of persons fitting the profile, may not equal the true probability that a person fitting the profile is a drug courier. Indeed, this is the same problem of trying to determine if a coin is fair by tossing it some number of times and observing the results.[230] In a criminal case where the evidence is not overwhelming, the error margin relating to p_{dc} may be significant to the jury's verdict; yet, there is no meaningful way to generate such a margin from the data hypothesized here.

To see the deficiencies of drug courier profiles as substantive evidence, consider how courts treat DNA evidence. DNA, or deoxyribonucleic acid, is being increasingly introduced as substantive evidence in criminal cases where identity is at issue.[231] DNA testing is used where human genetic material (such as blood, hair, or skin) has been left at the scene of the crime. Scientists analyze and compare the sample from the crime scene, referred to as the "known sample," against one from the defendant, which is called the "unknown sample."[232] What is key is that a match between samples does not mean that the defendant has been proven to be the source of the crime scene sample; rather, "the scientists are merely assessing the theoretical likelihood that a randomly selected person from the general population or a certain subsection of the population would match the known sample from the crime scene and the unknown sample from the crime suspect."[233]

How is the likelihood that DNA from a person from the general population would match the known sample calculated? Forensic scientists "establish a DNA data bank from a sample of the population and estimate the frequency of a specific DNA pattern within that population sample."[234] The DNA data bank can be thought of as the equivalent of the total sample of persons observed by DEA agents and who fit the drug courier profile. However, the adequacy of DNA data banks can be challenged on a variety of grounds, including (1) is the database sufficiently large to provide a random sample; and (2) were the DNA samples that make up the database collected in a way so as not to be random?[235] These challenges have been generally unsuccessful,[236] but the lack of success also suggests that the DNA databases are in fact adequate for the purpose at hand.

---

230. *See supra* 205.

231. Smith & Gordon, *supra* note 179, at 2468; *see also* People v. Venegas, 954 P.2d 525, 531 (Cal. 1998).

232. Smith & Gordon, *supra* note 179, at 2472.

233. *Id.*

234. *Id.* at 2474.

235. *See* David H. Kaye, *DNA Evidence: Probability, Population Genetics, and the Courts*, 7 HARV. J.L. & TECH. 101, 120-21 (1993).

236. *Id.*

Because the DEA does not keep track of the number of persons who are stopped but who are not carrying narcotics, there is no equivalent way for challenges to the adequacy of the profile's database to be evaluated. Thus, at present, the government's record keeping is insufficient to allow probabilities to be estimated accurately for the profile. Thus, it is difficult to accord the drug courier profile with the same respect as is given to DNA testing.

b. Independence

A separate problem arising from the drug courier profile is that in practical terms, its use appears to assume that the individual characteristics are independent of one another. Recall that the *Elmore* profile developed by Special Agent Markonni consisted of seven primary characteristics.[237] The profile appears to treat all seven characteristics as independent, when in fact they might not be. It may be that there is some correlation between the use of an alias and the purchase of tickets using $20 bills if, for example, the traveler places a high premium on privacy.[238]

This problem would be of little significance if the profile did not rely implicitly on the product rule. That is, there would be no need to establish the independence of the various factors if the DEA stopped only those travelers who fit all seven primary characteristics, because the seven characteristics can be treated mathematically as one complex characteristic. However, very rarely does a suspect fit every single characteristic.[239] Instead, suspects fit different combinations of the characteristics, which means that there are, for predictive purposes, several profiles, each consisting of some but not all of the traits of the "main" profile. For example, one characteristic is the purchase of an airline ticket using small bills.[240] Suppose DEA agents observe a person who fits the other six characteristics, but who pays for the airline ticket with a credit card. Presumably, the agents would consider stopping this person. Yet, the probability that this person is a drug courier has to be less than if the person had fit all seven characteristics. Depending on the probative values of the particular characteristics, this probability may differ from that of a person who fits all the characteristics except for the use of an alias.

Allowing the DEA to treat six out of seven characteristics as a separate profile so as to avoid having to establish the independence of the

237. *See supra* note 18 and accompanying text.
238. *See* Becton, *supra* note 16, at 449.
239. *See* Cloud, *supra* note 27, at 890-91.
240. *See supra* note 97 and accompanying text.

factors creates its own problem. There is only one way in which a person can fit seven of the characteristics—namely by fitting each one. However, there are seven different ways in which a person can fit six out of the seven characteristics, and twenty-one different ways in which a person can fit five out of seven.[241] Treating each of these combinations as a different profile reduces the statistical certainty for each (as each would be based on fewer data points) and would require even greater record keeping efforts on the part of the DEA.

The alternative would be to determine the individual probabilities for each characteristic, demonstrate that the characteristics are independent, and then apply the product rule to determine the probability that a person fitting the number of observed characteristics would be a courier.[242] In practical terms, this process runs afoul of a number of the obstacles that I have discussed earlier.

First, the difficulty of determining the individual probabilities has been discussed in the previous subsection. Second, the product rule requires that the individual probabilities be independent. In theoretical problems, this is easy to determine. In real life, however, establishing independence is, as noted earlier, the reverse problem of establishing causation—a task that has baffled epidemiologists in a number of high-profile instances.[243] Moreover, epidemiology approaches scientific testing in a controlled and rigorous manner, generally involving control groups and test groups, with single or double masking so as to prevent biases from tainting the results.[244] The drug courier profile, on the other hand, does not appear to have been tested in any remotely similar way.

Thus, any attempt to present to the jury an assessment of the probability that a person fitting some number of the characteristics of the profile is indeed a drug courier runs afoul of the twin problems of calculating the individual problems of each characteristic and of demonstrating that the characteristics are indeed independent. Even if the

---

241. In particular, the number of ways in which a person can fit n out of seven characteristics, where $1 < n < 7$, is given by $7!/(n!2!)$ where $7! = 7 \times 6 \times \ldots \times 2 \times 1$. See WEAVER, *supra* note 98, at 94.

242. For example, suppose that a traveler arrived from a source city, carried no luggage, had a rapid turnaround time, and bought the ticket using $20 and $10 bills. Further suppose that the probabilities that a person fitting each characteristic individually would be innocent are as follows:

- arriving from source city = 9/10
- carrying no luggage = 7/10
- rapid turnaround time = 1/2
- buying ticket in cash = 3/10

If these characteristics are independent from one another, the probability that a person fitting all four characteristics would be innocent is simply the product of the probabilities, or 183/2000, or just over nine percent.

243. *See supra* notes 239-41.

244. LILIENFELD & STOLLEY, *supra* note 223, at 165-66. Single masking means that the subjects are not told if they are being treated, so as to avoid subject bias. With double masking, the observer is also unaware of the distribution of subjects between control and test groups.

prosecution could do so, it would involve a tremendous amount of collateral evidence likely to overwhelm the jury.

C. Using Probabilities Properly

Even if a given probability is considered relevant and probative, and has been calculated properly, the probability may still be misunderstood. This is true because the goal of a criminal investigation is to eliminate, if possible, a false negative, while the goal of an investigation is to eliminate a false positive.

1. The Problem

Consider a test for a disease with a ninety-nine percent accuracy rate both with regard to positive and negative results. That is, ninety-nine times out of 100, a positive result means that the person being tested has the disease being tested for, and ninety-nine times out of 100, a negative result means that the person being tested does not have the disease. If you test positive, how likely are you to have the disease?[245]

The answer is most likely not ninety-nine percent. Here's why. Suppose that ten percent of the populace is actually afflicted with the disease. Out of 1000 subjects, 100 can be expected to have the disease, and 900 will not. Of the 100 who have the disease, ninety-nine will test positive and one will test negative. Of the 900 who do not have the disease, 881 will test negative and nineteen will test positive. Thus, there will be a total of 118 positive tests, of which nineteen—or sixteen percent—are false positives!

The incidence of false positives increases as the accuracy of the test decreases; however, it also increases as the incidence of the population afflicted with the disease decreases. Suppose that only one percent of the population has the disease. Then, out of 1000, ten can be expected to have the disease; 9.9 (call it ten) will test positive. Of the 990 who do not have the disease, 980 will test negative and 9.9 (again, call it ten) will test positive. Thus, fifty percent of the positive results are false positives.

Let me return to the example of the man murdered in his mistress' apartment. Thus far, I have suggested that the observation regarding men and mistresses is probative enough to justify its use for probable cause purposes, but is also sufficiently questionable in reliability to

245. This example is drawn from PAULOS, *supra* note 212, at 66.

justify its exclusion from trial. Yet, there is a question of whether the observation is being presented accurately. Tribe's hypothetical states that it is observed that in ninety-five percent of the known cases, the mistress is the murderer. The use of the word "known" may indicate that Tribe means that in ninety-five percent of the solved cases, the mistress is the murderer; however, the statistic tells us nothing about the total number of cases—solved or unsolved—of men murdered in their mistresses' apartments. For example, suppose that in the history of Los Angeles, 1100 men have been found murdered in their mistresses' apartments. Of those cases, 100 were solved, and ninety-five of the defendants convicted were mistresses and the other five were robbers. That leaves 1000 cases for which the actual killers were never discovered. It may be that, consistent with the ninety-five percent statistic of the solved cases, 950 of those murders were committed by the mistresses as well. On the other hand, it may be that whenever the mistress commits the murder, she is always caught, but that when the killer is not the mistress, the crime becomes very difficult to solve. If that is so, the statistic is of no material value at trial.

2. *As Applied to Drug Courier Profiles*

Although no one has presented an empirical analysis of the accuracy of the drug courier profile[246]—which is one of the problems with using it as substantive evidence—I would be surprised if the accuracy rate comes close to ninety percent. Furthermore, recall that at the trial stage we are concerned primarily with Type I errors, not Type II errors. Thus, the profile will not be particularly helpful even if it were ninety-nine percent likely to identify a drug courier as such, but also fifty percent likely to identify an innocent person as a courier. That is to say, given that a person is a courier, the profile may be accurate at identifying the person. But at the trial what we are interested in is, given that the profile has identified a person as a courier, how likely is it that the person really is acting as a courier?

D. Comparison of Profile Evidence to Other Evidence

Of course, the prosecution has to be able to introduce the defendant's actions as circumstantial evidence of guilt, even if those actions are the same ones that fit the drug courier profile. Where the

246. Becton, *supra* note 16, at 418 n.4; Cloud, *supra* note 27, at 843.

defendant has been caught with controlled substances, the case will typically hinge on the *mens rea* element.[247] In the absence of a confession, which could obviate the need for a trial, the existence of *mens rea* will turn on circumstantial evidence, such as the fact that the defendant paid for a ticket in cash, checked no luggage, and so on and if the defendant testifies, his or her lack of credibility. If believed, this evidence supports an inference that the defendant was aware of the criminal nature of his or her conduct.

But if the prosecution is entitled to present this evidence, why shouldn't it be entitled to take the next step and inform the jury that the defendant must have known he or she was acting as a courier, because his or her actions matched those of other couriers? After all, that the defendant paid for a ticket in cash and checked no luggage are among the very factors that eventually led to his arrest.

The answer is that the drug courier profile is neither direct nor circumstantial evidence of the defendant's guilt. Rather, it is offered to bolster the quality of the circumstantial evidence in the case. The difference is that the individual elements of the profile stand by themselves, but the profile assumes the existence of other facts. For example, evidence that Jones paid for his airline ticket in cash, used an alias, and checked no luggage may be sufficient to establish that he was aware that he was smuggling narcotics, if the jury finds that they occurred and demonstrate criminal intent. The drug courier profile, if offered in conjunction with that circumstantial evidence, would presumably strengthen that evidence, but not in the same way that an additional piece of direct or circumstantial evidence would.

An example may be helpful here. Suppose at the time of his arrest for the stabbing murder of Smith, Jones is found with bloodstains consistent with Smith's blood type (and not Jones's own). This would be some circumstantial evidence of Jones's guilt. However, suppose that the amount of blood found on Jones is not enough (by itself) to support an inference of violence; the relevance is only in the identity of the blood type. The strength of that inference may likely be affected by whether Smith had blood type O (shared by about forty percent of the population) or blood type AB (very rare). In other words, the distribution of blood types among the population bolsters the quality of the circumstantial evidence. The difference between blood type distribution and the drug courier profile is that the former is a medically accepted fact, whereas the latter has yet to be validated to the same extent.

Admittedly, in the absence of the drug courier profile, the jury will be left to assign its own weight to the circumstantial evidence. Evidence

247. *See, e.g.*, United States v. Foster, 939 F.2d 445, 451 (7th Cir. 1991) ("The issue here was whether Foster knew that he was carrying narcotics."). In rare cases, a defendant might argue that law enforcement agents planted the evidence. *See* Kadish, *supra* note 2, at 749 n.5 (citing examples).

that a criminal defendant had a "rapid turnaround time for a very lengthy airplane trip"[248] suggests criminal activity, a prosecutor would argue, because it seems unlikely to be required for legitimate actions. Yet, one can imagine non-criminal reasons for creating such a flight schedule—flying to a funeral perhaps. The jury is expected to use its collective common sense and experience to assign its own estimate of the likelihood (i.e., probability) that a rapid turnaround time for a very lengthy airplane trip is due to criminal activity. Is there a justification for precluding a law enforcement agent from testifying to that which the prosecutor may imply at argument?

The prosecutor's characterization of the evidence is just that—a characterization. The jury will be instructed that nothing that the prosecutor states is evidence.[249] A particularly effective advocate may do a better job of making the circumstantial evidence seem compelling than an average prosecutor, but it is still advocacy.

With regard to the circumstantial evidence, the defendant is free to offer an explanation for the suspicious actions. Perhaps the defendant paid for the ticket in cash due to an aversion to credit cards, and perhaps the defendant checked no luggage out of a desire to leave the airport quickly. The jury can assess the credibility of these explanations and vote to convict or acquit on that basis. These are determinations for which the jury is well suited.[250]

The drug courier profile, on the other hand, is a piece of evidence for which the jury is poorly suited for evaluating. The credibility of the law enforcement agent who testifies as to the profile is generally not in dispute, in the sense that there is no question as to whether the agent believes in the profile. Rather, the jury is asked to determine the appropriate weight that the profile should be accorded, a task requiring a working grasp of probability theory that may not be easily imparted in the context of a criminal trial.

Indeed, any justification for the admission of the drug courier profile as a commentary on the circumstantial evidence would support similar commentary on circumstantial evidence in cases involving other crimes. An obvious example leaps to mind: the murderer of an adulterer profile set forth in Professor Tribe's hypothetical.

248. Elmore v. United States, 595 F.2d 1036, 1039 n.3 (5th Cir. 1979).

249. *See, e.g.*, 9th Cir. Crim. Jury Instr. § 3.5 (West 1997); 5th Cir. Crim. Jury Instr. § 1.06 (West 1997).

250. *See* United States v. Scheffer, 118 S. Ct. 1261, 1266 (1998); *see also* KASSIN & WRIGHTMAN, THE AMERICAN JURY ON TRIAL 67-70 (1988).

VI. Conclusion

Despite the protests of civil libertarians and critical race theorists, the drug courier profile appears to be here to stay. To the extent that the profile is not used as a pretext to stop whomever the police feel like stopping, it can be defended as an investigative tool, at least on a theoretical level. Human behavior does follow predictable patterns, and even if the profile is imperfect, at the investigative stage, society can tolerate false positives.

Courts should, however, generally exclude the profile from trials. The Seventh Circuit's contrary approach of allowing the introduction of the profile as circumstantial evidence is problematic for a number of reasons. First, the profile may not be probative of anything more than the fact that the defendant was likely to be found with narcotics if arrested; in a case where intent is at issue, the profile has little, if anything, to offer—particularly since we don't know how many of the persons arrested due to the profile were actually convicted.

Even if the profile is viewed as having some probative value, as the Seventh Circuit has held, the defendant should be entitled to test the underlying mathematical assumptions of the profile. Validation of the profile, if even possible, would require significant amounts of testimony on issues of probability estimation and the independence of variables. In a case where the circumstantial evidence is strong, the relative probative value of the profile is weak and there should be no reason for the prosecutor to attempt to introduce it. Conversely, where the circumstantial evidence of guilt is weaker, the prosecutor may be more inclined to present the profile in an attempt to bolster the case. But where the overall case is less than overwhelming, the potential errors inherent in the profile may lead to an incorrect result. It is in such a situation that a false positive may have occurred.

Appellate courts that exclude the drug courier profile from trials have therefore reached the right result, though not necessarily for the right reason.

Appendix: Solutions to the Probability Problems

The Birthday Problem

The birthday problem revealed the remarkable result that only twenty-three people need to be in a room before there is a fifty percent probability that at least two of them share a birthday.

The easiest way to calculate this probability is to calculate the probability of the complementary event—that no two people in a group of n persons share a birthday—and subtract that result from 1. Assuming that birthdays are randomly distributed, the probability that two people have different birthdays is 364/365. The first person can have any birthday, and the second person has 364 days out of 365 to have a birthday not in common with the first person. The probability that three people have different birthdays is 1 * 364/365 * 363/365, there being 363 days for the third person to have a different birthday than the first two. In general, the probability pb that n persons all have different birthdays is:

$$pb = [364 * 363 * \ldots * (366 - n)] / 365^{n-1}$$

Thus, 1 - pb yields the probability that at least two persons in a group of n persons shares a birthday. When $n = 23$, $pb = 0.493$, and 1 - $pb = 0.507$.[251]

The Monty Hall Puzzle

In the Monty Hall puzzle, the player was presented with three doors, one of which hid a prize, and asked to pick one. The host then opened one of the other two doors and the player was given the option of keeping his or her first choice or switching to the other unopened door. Although it may seem as if it makes no difference, computer simulations have demonstrated that switching is the better strategy.[252]

The proper way to solve the puzzle is to see what happens if you always switch. Suppose you initially pick the door with the prize behind it. The host opens one of the empty doors, and you switch to the other empty door and lose. Now suppose you initially pick an empty door. The host opens the other empty door, and you switch to the door with the prize and win. So if you always switch, you will win if you initially pick

251. *See, e.g.*, ACZEL, *supra* note 210, at 198-99.
252. HOFFMAN, *supra* note 216, at 238-39.

an empty door and you will lose if you initially pick the door with the prize. Since there are three doors and only one prize, you are twice as likely to choose an empty door initially.[253]

The Card Game

In the card game, a card shark has three cards: one is red on both sides, one is black on both sides, and one is red on one side and black on the other. The card shark mixes them up, draws one, and shows you that it is black on one side. He then offers to bet you that it is the black-black card. It turns out that this is not a fair bet, as the other side of the card is twice as likely to be black than red.

The conventional explanation may seem oblique: "The visible side of the card you picked could be the red side of the red-black card, in which case you win, or it could be one side of the [black-black] card, in which case [the con artist] wins, or it could the other side of the [black-black] card, in which case [the con artist] also wins."[254] Think of it this way: you know it is not the red-red card. That leaves four sides: red, black, black, and black. You are looking at one of the black sides. That leaves one red and two black sides left that the other side of the card could be. Therefore, the other side of the card is twice as likely to be black as it is to be red.

253. That is, there is a 2/3 probability of selecting an empty door and a 1/3 probability of selecting the door with the prize.

254. PAULOS, *supra* note 212, at 64-65.