Logical relevance: problems with the reference population and DNA mixtures in People v. Pizarro

D. H. Kaye

Regents’ Professor and Fellow, Center for the Study of Law, Science, and Technology, Arizona State University, Tempe, Arizona 85287-7906, USA

[Received on 16 August 2004; revised on 19 October 2004; accepted on 20 October 2004]

The concept of relevance is fundamental to the law of evidence. A California court of appeals recently invoked this concept to deem inadmissible estimates of the relative frequency of an incriminating DNA type in major racial groups and computations relating to a mixture of DNA from more than one person. This article shows that these statistics are relevant evidence because they help the factfinder assess the significance of matching DNA types under various hypotheses that might be entertained about the origin of the DNA samples.

Keywords: evidence; relevance; DNA evidence; race and ethnicity; interpretation of mixed stains.

For two decades, the legal system has struggled with DNA evidence. Although the courts have discerned the essential features of the structure of the DNA molecule and the biochemical techniques for analysing the characteristics of particular DNA samples, the logical and statistical aspects of the evidence have proved more difficult. Misinterpretations of DNA evidence are prevalent and persistent.1 This article considers two such problems that are the subject of an opinion in People v. Pizarro,2 a case that has perplexed the California courts for the last 14 years and that has generated a final opinion that is disturbing in its treatment of (1) the presentation of random-match probabilities in a reference population and (2) the interpretation of samples containing DNA from more than one person. According to the Pizarro court, when a crime has been committed by someone whose race or ethnicity is not known, presenting data on the relative frequency in various racial or ethnic groups of the type of DNA found at the scene of the crime is ‘objectionable’ because it is extraneous, potentially irrelevant and prejudicial.3 The court reaches the same conclusion concerning statistics computed under alternative hypotheses about the possible contributors to a mixed stain.4 This article demonstrates that the court’s analysis is defective: the statistics condemned in Pizarro are logically relevant.

One strangely persistent fallacy in the interpretation of DNA evidence is that the relevant ethnic or racial population in which to estimate a DNA profile frequency necessarily is that of the defendant. The issue has been cogently analyzed, and it should be clear that relevant population is the entire class of plausible perpetrators.5 After all, the probative value of the observation that the defendant has the DNA profile found at a crime scene depends on the likelihood ratio for that

1 See, e.g., 3 MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY §25-1 (David Faigman et al. eds., 2d ed. 2002).
3 Id.
4 Id. at 78–79.
evidence. This likelihood ratio is given by (a) the conditional probability of the observed match given that the defendant is the source of the crime-scene DNA divided by (b) the conditional probability of the observed match given that someone else is the source. The numerator has nothing to do with the frequency of the DNA type in any ethnic or racial population. Instead, it depends on the laboratory’s ability to detect matches between two matching DNA samples. The denominator, on the other hand, is related to the frequency of the DNA type in other people, for if the defendant is not the source of the DNA from the crime scene, then someone else is. Sometimes that individual will share the defendant’s race or ethnicity. Other times, the true source will be from another group. Consequently, the defendant’s genetic heritage does not ineluctably define the relevant population in which to estimate a DNA frequency or random-match probability.

Yet, lawyers and courts continue to make this conceptual mistake. A Florida case offers a recent, but by no means the most egregious, example. In *Darling v. State*, a Polish woman living in Orlando was sexually assaulted and killed. Semen was extracted from a vaginal smear. A DNA analyst testified to random match probabilities on the order of one in billions or trillions obtained with ‘the modified ceiling principal [sic] formula’ applied to allele frequency estimates from ‘African-American data, . . . Caucasian data, and . . . Southeastern Hispanics from the Miami area.’ In response, the defendant ‘challenge[d] . . . the expert’s failure to use a Bahamian database.’ This criticism might have been apt if the crime had been committed in one of the chain of islands that are the Bahamas. But if the defendant was not the culprit, then there is little reason to presume that the criminal came to Orlando from places such as Bimini, Cat Island, Exuma, Freeport, Fresh Creek, Governor’s Harbour, Green Turtle Cay, Ragged Island, or Rum Cay. Men living in the Orlando area rather than Bahamians constitute a more appropriate reference population. The Florida Supreme Court seemed oblivious to this fact. It apparently assumed that, ideally, the allele frequencies should have come from a Bahamian database.

---

7 If the probability of laboratory and handling error is very small, the numerator will be close to one.
8 See David Balding & Richard A. Nichols, *DNA Profile Match Probability Calculation: How to Allow for Population Stratification, Relatedness, Database Selection and Single Bands*, 64 *Forensic Sci. Int’l* 125 (1994) (arguing that ‘there are a priori reasons to suppose that the suspect, if innocent, is drawn from the same subpopulation as the criminal’).
9 To say that there is a relevant population is not to say that all members of this population are equally likely to be the source of the trace evidence. Kaye et al., supra note 6, §12.4.1., at 472 n.69. From a Bayesian perspective, such variations can be captured by using functions (that vary across the members of a reference population) for the prior probability distribution and the likelihood. This approach is conceptually valuable, but implementing it directly in litigation is challenging. Id. §12.3.2(b), at 495. Presenting distinct frequencies for DNA types in different subgroups is one way of indicating how the likelihood ratio varies within a broadly defined reference population.
10 An extreme example is *People v. Mohit*, 579 N.Y.S.2d 990 (Westchester Co. Ct. 1992). In this case, the court wrote that inbreeding among Shiite Muslims around Shushhtar, Iran, was ‘of particular importance.’ Id. at 997. But the alleged crime—a sexual assault of a woman in New York—did not occur in or near Shushhtar. If the state was wrong in alleging that the defendant, a physician from Iran, raped his patient, then someone else in Westchester County or thereabouts did.
11 808 So.2d 145 (Fla. 2002).
12 Id. at 152.
13 Id. at 151.
14 Id. at 150.
15 If there is a Bahamian community in Orlando and reason to think that the murderer came from this particular group, the defendant’s criticism of the statistics would have more substance.
16 Nevertheless, the court affirmed the conviction on the theory that the use of the modified ceiling calculation made the figures applicable to the Bahamas. Id. at 159–60.
People v. Pizarro is a step up from Darling, for the California court of appeals recognized that the reference population should be defined as the class of plausible suspects. However, the court’s analysis of the relevance of random-match probabilities in various populations is deficient. In 1990, Michael Pizarro was convicted of raping and suffocating his 13-year-old half-sister. Vaginal swabs from the girl’s body revealed semen from a type B secretor. Pizarro, along with some 8% of the population, was a type B secretor. Additional vaginal swabs and blood samples from Pizarro and the victim were sent to the FBI, which found a multilocus VNTR match. At trial, an FBI analyst testified that ‘[t]he likelihood of finding another unrelated Hispanic individual with a similar profile as Mr Pizarro is one in approximately 250,000’.

Pizarro appealed this conviction, contending that the DNA evidence was inadmissible because

18 Id. at 28–29.
19 Because the sister was a type O nonsecretor, the source of the semen had to be a type B secretor.
20 Id. at 97–98. The pertinent testimony went as follows:

[PROSECUTOR:] What is your opinion as to the chances of another Hispanic male having the same DNA profile as Mr Pizarro?
[ADAMS:] The likelihood of finding another unrelated Hispanic individual with a similar profile as Mr Pizarro is one in approximately 250,000.

[PROSECUTOR:] And this would also be the same statistic for the probability of a match of a DNA profile between the [perpetrator’s DNA] obtained from the vaginal swab?
[ADAMS:] That is correct.

[PROSECUTOR:] Same statistic?
[ADAMS:] Yes.

[PROSECUTOR:] And, again, this is only with Hispanic men?
[ADAMS:] Hispanics, not broken down into gender. [¶] ... [¶]

[PROSECUTOR:] Dr Adams, we have been talking about the chance for a match within the Hispanic community. Would the statistics for a match within the Caucasian community be different?
[ADAMS:] Yes, generally there are going to be some differences in the population data from the different populations. So that’s why we keep them separate. That’s why we have a Caucasian and a Black and a Hispanic, American Indian population because there are differences. [¶] So if I were to compare one person in each of those different populations I would come up—I’m sure I would come up with somewhat different results because in one population that pattern may be very rare, and another population that same pattern may be very common.

[PROSECUTOR:] Have you done any of the calculations necessary to determine what the chances are of having matches of this particular DNA profile within the Caucasian community?
[ADAMS:] Yes.

[PROSECUTOR:] And what are those statistics?
[ADAMS:] The statistics in those cases—in that case comparing the same profile to the Caucasians is much greater. It would be one in 10,000,000.

[PROSECUTOR:] But within the Hispanic group alone it is according to your testimony one in 250,000?
[ADAMS:] Yes, ma’am.

[PROSECUTOR:] What about a situation where someone is half Hispanic and half Caucasian?
[ADAMS:] Well, there is nothing we can do other than to compare them to the two populations and we would use only the smaller of the two in our report. [Adams referred to the number with the smaller denominator.]

[PROSECUTOR:] Why do you use only the smaller of the two?
[ADAMS:] We attempt to be as conservative as possible. The smaller number is less detrimental to the defendant.

Id. This testimony is somewhat misleading in that the computed probability pertains to finding a match in a single, randomly selected individual. The chance that somewhere there is another Hispanic individual with the genotype in question is much larger.
the prosecution had failed to demonstrate that the VNTR testing was generally accepted in the scientific community. The court of appeals remanded the case for a hearing on this issue. After the hearing, the trial court again ruled that the evidence was admissible and reentered the judgment. A second appeal followed. In a prolix opinion, the appellate court determined that it was improper for the prosecution to have offered an estimate of the frequency of a genotype in the Hispanic population—and no other group—when there was no proof (other than that pointing to Pizarro) that the perpetrator of the crime was Hispanic. For this reason, the court of appeals again reversed.

This result is easily justified. If the perpetrator could have come from any of several racial groups, looking to only one racial group for a random match probability could be misleading. However, the *Pizarro* court announced that giving a range of frequencies for the major racial or ethnic groups in the United States is unacceptable. Since providing statistics from several racial groups is the standard way of assessing the significance of a match in cases in which the racial and ethnic status of the perpetrator of the crime initially is unknown, the opinion casts doubt on the outcomes of innumerable cases.

Of course, this might not be a bad thing if the standard approach compromised an important principle in the law of evidence. The *Pizarro* court thought that it did—that there is an unbridgeable gap between scientific and legal reasoning in this situation. Yet, the scientific reasoning that the court questioned is nothing less than the kind of hypothesis testing—considering the principal alternatives and examining the probability of certain outcomes under each of these alternative hypotheses—that dominates modern statistical thinking. In this instance, the DNA expert simply testifies to how surprising the match would be if some major alternatives to the hypothesis that the defendant is the source of the biological samples were true.

The *Pizarro* court rejected this established way of thinking about the significance of a match for four reasons. First, the court asserted that the frequency of the genotype in any given racial or ethnic group is logically irrelevant:

> [I]n the absence of sufficient evidence of the perpetrator’s ethnicity, any particular ethnic frequency is irrelevant. The problem is . . . one of preliminary fact . . . . It does not matter how many Hispanics, Caucasians, Blacks, or Native Americans resemble the perpetrator if the perpetrator is actually Asian. If various ethnic frequencies are presented to the jury, each will have been admitted without adequate foundation.

According to the court, the law’s refusal to consider specific alternative hypotheses is a profound

21 Actually, the state presented frequency estimates in two population groups—Hispanic and Caucasian. But these were chosen because the defendant was said to be half Caucasian and half Hispanic. *Id.* at 97. For brevity, the text of the article refers only to the Hispanic database.

22 As the citations to the scientific literature noted in the opinion reveal, this is not an accepted procedure among forensic scientists. In addition, as the court explains, by suggesting that the defendant’s ethnicity is the only one that merits consideration, the jury could be misled into thinking that the crime had to have been the work of an Hispanic like the defendant. To this extent, the presentation could be unfairly prejudicial (even if, as the state argued on appeal, the number itself probably is larger than that in the relevant population). The prejudice argument, however, is not necessarily convincing, for it would have been easy enough for defense counsel to expose the prosecution’s error through cross-examination or expert testimony of its own.

23 Of course, other approaches are possible, and some have been said to be better suited to forensic proof. See generally, e.g., BERNARD ROBERTSON & G.A. VIGNAUX, *INTERPRETING EVIDENCE; EVALUATING FORENSIC SCIENCE IN THE COURTROOM* (1995) (advocating a Bayesian approach); Balding & Nichols, *supra* note 8 (describing a Bayesian approach); KAYE *et al.*, *supra* note 6, §12.4.3 (discussing such possibilities).

24 3 Cal.Rptr.3d at 104.
manifestation of ‘subtle, even unexpected, differences between the scientific and legal approaches to the same problem’; 25

Here, science promotes consideration of all possibilities, whereas law restricts consideration to possibilities it deems relevant. A conundrum such as this, bound to arise in scientific cases, can be detected and resolved only through the attentive and respectful contemplation of the two disciplines and the mindful evaluation of their separate and intersecting principles. Inevitably, some scientific principles, although correct in their scientific context, will not survive translation into legal application of relevancy principles. 26

These pronouncements might be true in some contexts. When it comes to deciding what evidence is logically relevant, however, it is difficult to conceive of any substantive difference between legal and scientific reasoning. 27 In this instance, estimates of the frequency of a trait in various racial groups are not introduced because they are independently relevant. They are used to indicate the probable value of the fact that the defendant possesses the genetic traits that characterize the perpetrator of the crime. The appropriate foundation for using them consists of proof that (a) for the DNA loci involved, the defendant’s DNA matches the DNA found at the crime scene, and (b) the estimates are produced by a scientifically valid procedure.

The further requirement that the racial identity of the perpetrator be known seems to arise from the court’s recognition that if this fact could be known, some of the statistics would be irrelevant. According to the court, ‘the jury is not assisted by knowing how many Hispanics possess the perpetrator’s traits if the perpetrator is actually Asian’. 28 However, we can never know the perpetrator’s ancestry to a certainty. Suppose that the victim had lived and identified the man who assaulted her as ‘Asian’. The cross-racial identification, made under great stress, could be mistaken. If an Asian defendant wished to argue that the actual rapist might have been Hispanic, it would assist the jury to know whether the DNA types are extremely rare among Hispanics. Indeed, in the limit, if no Hispanic had the requisite genotype, the Hispanic data would refute defendant’s alternative hypothesis. Conversely, if every Hispanic had the incriminating genotype and hardly any Asians did, the DNA evidence would be much less probative of the Asian defendant’s guilt than the Asian-only figure would suggest. To withhold the Hispanic data on the ground that it is irrelevant because the other evidence in the case points to an Asian could be a serious mistake.

Moreover, even if we could know that the perpetrator was, let us say, Asian, with sufficient certainty to exclude all other possibilities from rational discourse, the court’s logic might lead to the paradoxical conclusion that the frequency data for Asians also is irrelevant. Asian-Americans, after all, are not a homogeneous group. There are many subgroups—Chinese, Indonesian, Japanese, and Korean, to name a few—and each subgroup can be parsed still more finely. Would the court conclude that when the most that can be said is that the unknown perpetrator is definitely of Asian ancestry, ‘the jury is not assisted by knowing how many Asians possess the perpetrator’s traits [because] the perpetrator [might] actually [be] of [aboriginal Taiwanese] ancestry’? 29 Contrary to

25 Id. at 105 (note omitted).
26 Id. at 105 (note omitted).
28 3 Cal. Rptr.3d at 104.
29 A possible rejoinder to this subgroup problem is that the statistics for all Asians have some relevance to the Taiwanese
the Pizarro court’s assertions, in a ‘general population case’—one in which the investigation cannot be limited to a particular racial group—the statistics for a range of groups surely are relevant.

Let us turn, then, to the court’s additional reasons for rejecting a range of statistics. ‘Second,’ the opinion claims that ‘the improper mention of ethnicity unfairly and unjustifiably encourages the jurors to focus on ethnicity and race—specifically the ethnicity and race of the defendant, the only suspect before them.’30 Yet, there is no obvious reason to think that telling the jury that the incriminating genotypes occur infrequently in every major racial or ethnic group will encourage the jurors to assume that the crime must have been committed by a member of the defendant’s racial or ethnic group. In fact, giving a range of statistics could discourage the jury from jumping to an unjustified conclusion.

The court’s third argument fares no better than the first two. In fact, it is not really a separate argument. The opinion insists that:

Third, the jury hears unjustifiably damaging evidence because the various ethnic frequencies create a range extending from the most conservative and beneficial to the defendant to the most rare and damning to the defendant. In our example, the evidence against the defendant includes not only the most favourable 1-in-1-million (Caucasian) frequency, but also the most damaging 1-in-10-million (Hispanic) frequency. If the perpetrator is actually Caucasian, only the most favorable 1-in-1-million (Caucasian) frequency is relevant, but the jury nevertheless will hear—and likely focus on—the more damaging 1-in-10-million figure. If the perpetrator is actually Black, only the 1 in 2.5 million (Black) frequency is relevant, but the jury will hear—and likely focus on—the more damaging 1-in-10-million figure. The greater the disparity between the perpetrator’s true frequency and the range’s most damaging extreme, the greater the prejudice the defendant will suffer from mention of that extreme. Without adequate evidence of the perpetrator’s ethnicity, there is no justification for presenting the most damaging frequency.31

This is little more than a restatement of the two previous points,32 and the rejoinder is the same. The fact that because more complete—but unattainable—knowledge would eliminate some subgroup. Yet, the same can be said of the statistics for all Hispanics. They shed some light on the frequencies in other groups. We are all human beings, and DNA types that are very rare in one racial group tend to be rare in others; those that are common in one group tend to be common in others. There are exceptions, of course, and rough racial inferences can be drawn from certain unusually distinctive combinations of alleles, but the commonalities are great enough to meet the law’s rather undemanding standard of relevance.

30 3 Cal.Rptr.3d at 104.
31 Id. (footnotes omitted).
32 It is a little more than that because an accompanying footnote advocates a ‘conservative’ procedure that would hide the fact of racial and ethnic variation from the jury:

A better option when the perpetrator’s ethnicity is not established is presentation of the one most conservative frequency, without mention of ethnicity. In our example, the jury would be told that the perpetrator’s genetic profile is found in no more than 1 in 1 million people—that the profile is at least this rare. . . .

Id. at 104 n.85.

The opinion also floats another suggestion:

A second option when the perpetrator’s ethnicity is not established is presentation of a single frequency calculated from a general, nonethnic database. Again, this method makes no assumptions regarding the perpetrator’s ethnicity and promotes no unwarranted ethnic or racial considerations. If, however, this method is not scientifically valid or results in a frequency that is not considered conservative, it is not a viable option.
alternatives does not make it wrong to contemplate those alternatives, and there is no inexorable prejudice in providing a range of applicable figures.\textsuperscript{33} Finally, the opinion urges ‘cautious evaluation . . . because of the ambiguous nature of artificially defined ethnicities and the uncertainties connected to use of an ethnic database’.\textsuperscript{34} The problem, the opinion states, is that ‘most populations are mixed, that the definitions are to some extent arbitrary, and that they are sometimes more linguistic (e.g. Hispanic) than biological. In fact, people often select their own classification.’\textsuperscript{35} However, ‘cautious evaluation’ is not the same as ignoring the clear differences that appear among the roughly defined racial and ethnic groups.\textsuperscript{36} The limitations in the databases can be made clear to the jurors, who are in a better position to evaluate the possibility that the match is coincidental if they have a range of the possible frequencies that are consistent with the facts of the case. As such, the exclusionary rule announced in \textit{Pizarro} is baseless and counter-productive.

The same kind of dubious reasoning about relevance infected the California court’s response to an ambiguity in the interpretation of a mixed sample. At one VNTR locus (denominated D2) only two alleles could be detected. At other loci, more alleles were present. Because one individual can have no more than two alleles per locus—one from each parent—it was clear that the police were dealing with a mixed DNA sample.

Partly because \textit{Pizarro} had the same genotype as the victim at this locus, the FBI presumed that the perpetrator did as well, and it used this assumption in computing the frequency of the perpetrator’s multilocus genotype. However, there are other possibilities to consider. The perpetrator might have been homozygous for one of the alleles. That is, rather than having two distinct alleles in common with the victim, he might have inherited a single matching allele from both of his parents. Or, he might have had a second allele that did not match the victim’s other allele but that was not detected for technical reasons. Thus, the appellate court decided that ‘if the perpetrator’s

\textit{Id.} However, averaging across racial and ethnic groups, which is roughly what this amounts to, cannot be uniformly conservative. See, e.g., David H. Kaye, \textit{DNA Evidence: Probability, Population Genetics, and the Courts}, \textit{7} \textit{Harv. J. L. \\& Tech.} 101 (1993). Therefore, ‘it is not a viable option.’ \textit{3 Cal. Rptr. 3d at 104 n.85.}\textsuperscript{33} \textsuperscript{34} \textsuperscript{35} \textsuperscript{36} If the categories were entirely arbitrary and misclassifications were dominant, major differences in allele frequencies and genotype frequencies would not be seen.
D2 genotype was discerned solely by reliance on defendant’s D2 genotype, the perpetrator’s genotype was discerned by an improper procedure. Furthermore, that genotype served as inadequate foundation for the relevance of the DNA evidence.\textsuperscript{37}

Rather than stop here, however, the court ‘digress[ed] to discuss the suggestion . . . that because the perpetrator’s genotype cannot be discerned, all possible genotypes should have been accounted for.’\textsuperscript{38} There are differences in the details of the procedures for accounting for the ambiguity, but the scientific consensus is that it is logical and reasonable to compute likelihood ratios or frequencies in the light of alternatives as to the number and nature of contributors to the mixed stain.\textsuperscript{39} Again, the California court charted a new course—exclusion of the evidence that the defendant’s DNA was consistent with the DNA found on the vaginal swabs. It did so on the basis of its understanding of the logic of proof, as illustrated with the following analogy about hair colour:

When evidence is lacking on a certain fact such that the fact cannot be established, the situation does not justify consideration of all possible alternatives to that fact. Only the one fact is relevant. If [an] eyewitness is uncertain about the perpetrator’s hair color, but can narrow the color down to black, brown, or blond, should all three possibilities be taken into account? The logic supporting an affirmative answer states: all possible perpetrators have black, brown, or blond hair; the defendant has black hair; therefore, the defendant is a possible perpetrator. Although initially appealing, this logic ignores the fact that the perpetrator has only one hair color and thus only that one hair color is relevant to his profile; more importantly, it ignores the fact that if the perpetrator actually has brown or blond hair, the defendant simply is not the perpetrator. The correct logic requires a choice of these three possible syllogisms: (1) all possible perpetrators have black hair; the defendant has black hair; therefore, the defendant is a possible perpetrator; (2) all possible perpetrators have brown hair; the defendant has black hair; therefore, the defendant is not the perpetrator; (3) all possible perpetrators have blond hair; the defendant has black hair; therefore, the defendant is not the perpetrator. It would defy the principles of evidence to allow the eyewitness to testify that the perpetrator has black, brown, or blond hair when there is no way of establishing which one hair color the perpetrator actually possesses. This testimony is neither relevant nor probative, but it is potentially damning because it draws the defendant into the pool of possible perpetrators when in reality it more likely excludes him—two of the three possibilities exonerate him.\textsuperscript{40}

Applying this logic to the DNA results, the court reasoned that:

Similarly, only the perpetrator’s one D2 genotype was relevant to his genetic profile. If the prosecution could not establish which genotype the perpetrator possessed at that locus, there was no relevant evidence to admit from that locus. But, as in the analogy,

\textsuperscript{37} Id. at 78.

\textsuperscript{38} Id.

\textsuperscript{39} This view is reflected in the 1992 and 1996 committee reports of the National Academy of Sciences as well as subsequent articles on interpreting mixed stains. Oddly, the \textit{Pizarro} court suggested that the admission of likelihood ratios requires a full-blown Bayesian presentation and analysis. \textit{Id.} at 78 n.64. This is not the case, as likelihood ratios are useful outside of the Bayesian framework. The court, it would seem, simply misread the 1996 NRC report.

\textsuperscript{40} Id. at 78.
the most compelling reason for demanding proof of the perpetrator’s genotype and for refusing to admit evidence of all three possible genotypes was that the other two possible genotypes were more than irrelevant—they potentially proved defendant’s innocence. Thus, the evidence that was admitted to incriminate defendant actually had a greater chance of exonerating him. If the perpetrator was not heterozygous (i.e. if he was either homozygous for the top band or homozygous for the bottom band), defendant did not match the perpetrator and he was excluded as a possible perpetrator. Only if the perpetrator was heterozygous did defendant match and become a possible perpetrator.41

These dicta are patently fallacious and potentially mischievous.42 All manner of evidence in criminal cases narrows the class of suspects but still admits of heterogenous subclasses. That membership in some of these subsets can be inconsistent with guilt does not render the evidence irrelevant or prejudicial. To see this, we need only consider the court’s illustration. If an eyewitness’s testimony can eliminate everyone with red, white, orange, and purple hair, leaving only the drab browns, blacks, and blonds, then the witness’s observation advances the inquiry. It makes it more probable that the defendant, who has black hair, is the criminal. As such, it surely is relevant and probative.43

Indeed, the court’s suggestion that evidence that a criminal has either black, brown, or blond hair exonerates a suspect with black hair is absurd on its face. It rests on the premise that because two of three logically enumerable hypotheses imply innocence, a suspect is probably innocent. This assumes that every state of nature is equally probable—an assumption that is clearly implausible.44

DNA mixtures can be complicated to interpret, but this does not make those interpretations inadmissible.

The latest opinion in People v. Pizarro reflects an incomplete understanding of the legal concept

41 Id. at 78–79.
42 It would render inadmissible testimony that is routinely accepted in paternity cases. Consider a locus in which the mother (M), child (C), and alleged father (AF) all have types A1 and A2. Then the mother and the alleged father can produce children with genotypes (A1,A1), (A1,A2), (A2,A1), and (A2,A2), where the first allele is from the mother, and each of these outcomes is equally likely. The second and third types would be indistinguishable on the basis of standard testing—an A1 and A2 would be detected, but the laboratory could not say which parent transmitted which allele. Hence, the probability that this mother and alleged father would have a child with the alleles A1 and A2 is two out of four, or $\frac{1}{2}$. This probability is the numerator of the paternity index. The denominator is the sum of two quantities. The first term is the probability that M will transmit the allele A1 to C (which is 1/2) times the probability of picking A2 at random from the gene pool (which is the allele frequency, $p_2$). The second term is the probability that M will transmit A2 (also 1/2) times $p_1$ (the frequency of A1 in the gene pool). The resulting paternity index is $PI = 1/(p_1 + p_2)$. There is no reason to ignore this locus in computing the cumulative paternity index. The fact that it could not be ascertained which marker was contributed by the father and which by the mother (and that the man might be excluded if this were known) is taken into account in the computation described above.

43 See, e.g., Fed. R. Evid. 401 (“Relevant evidence’ means evidence having any tendency 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.’

44 The alternative hypotheses listed by the Pizarro court—that ‘the perpetrator was … either homozygous for the top band or homozygous for the bottom band,’ id. at 79—are relatively improbable. Most people are heterozygous, and (without further information on the genotype of the perpetrator), it is more likely that the unknown perpetrator in this case was too. According to the court, William Shields, a population geneticist who testified for the defence, ‘recommended that autorads with mixtures such as this be entirely excluded from the statistical calculations, in part because two of the three possible perpetrator profiles would actually exclude defendant as a suspect.’ Id. at 77. Two other defence experts disagreed with this unusual view.
of relevance. In the presentation of DNA evidence, there is no fundamental conflict between law and science. In particular, scientists, lawyers, judges, and jurors often need to consider a range of hypotheses about the origin of biological stains to make sense of DNA data. The law of evidence permits them to do so.

45 Many of the potentially relevant hypotheses are enumerated in Kaye et al., supra note 6, §12.3.1, at 449.