



2009

## 'False But Highly Persuasive:' How Wrong Were the Probability Estimates in *McDaniel v. Brown*?

David H. Kaye

*Penn State Law*

Follow this and additional works at: [http://elibrary.law.psu.edu/fac\\_works](http://elibrary.law.psu.edu/fac_works)

 Part of the [Criminal Law Commons](#), [Criminal Procedure Commons](#), [Evidence Commons](#), and the [Science and Technology Law Commons](#)

---

### Recommended Citation

David H. Kaye, 'False But Highly Persuasive:' *How Wrong Were the Probability Estimates in McDaniel v. Brown?*, 108 *Mich. L. Rev. First Impressions* 1 (2009).

This Article is brought to you for free and open access by the Faculty Works at Penn State Law eLibrary. It has been accepted for inclusion in Journal Articles by an authorized administrator of Penn State Law eLibrary. For more information, please contact [ram6023@psu.edu](mailto:ram6023@psu.edu).

# “FALSE, BUT HIGHLY PERSUASIVE”: HOW WRONG WERE THE PROBABILITY ESTIMATES IN *MCDANIEL V. BROWN*?

David H. Kaye<sup>\*†</sup>

In *McDaniel v. Brown*, the Supreme Court will review the use of DNA evidence in a 1994 trial for sexual assault and attempted murder. The Court granted certiorari to consider two procedural issues—the standard of federal postconviction review of a state jury verdict for sufficiency of the evidence, and the district court’s decision to allow the prisoner to supplement the record of trials, appeals, and state postconviction proceedings with a geneticist’s letter twelve years after the trial. The letter from Laurence Mueller, a professor at the University of California at Irvine, identified two obvious mistakes in the state’s expert testimony.

This essay clarifies the nature and extent of the errors in this evidence in *Brown*. One might think that the expert’s letter, the opinions of the lower courts, and the briefs—including one from “20 Scholars of Forensic Evidence”—would have done this, but there is more to be said.

## I. THE DNA MATCH

Troy Brown was tried and convicted of a brutal rape in Carlin, Nevada, largely based on DNA evidence. Renee Romero, a criminalist for the county, discovered semen on the victim’s bloody panties. Romero reported that DNA from the semen matched Troy’s at the locations, or loci, for six genes. Her report estimated that the versions of the genes occur in “1 in 18,900 in the Caucasian population, 1 in 2,460,000 in the Black population and 1 in 4,800 in the Hispanic population.” Additional testing showed matches at five VNTR loci. (Variable Number Tandem Repeats are DNA sequences that come in many possible lengths. This makes them very discriminating, but the typing process is laborious and no longer in common use in forensic science.) Romero estimated the random match probability (RMP) for the additional loci to be 1 in 3 million. This quantity is the chance that a randomly selected, unrelated individual would share the loci—a URMP, for short. A more modern computation gives a value below 1 in 150 million.

---

\* Distinguished Professor and Weiss Family Scholar, Pennsylvania State University, Dickinson School of Law. A version of this essay was presented at George Washington University’s Conference, Celebrating Forty-five Years of Statistical Activity of Professor Joseph L. Gastwirth, August 1, 2009, Washington, D.C. I am grateful to James Crow, Edward Cheng, Mitchell Holland, and Kit Kinports for comments and to Laurence Mueller and William Thompson for email exchanges.

† Suggested citation: David H. Kaye, Commentary, “False, But Highly Persuasive”: How Wrong Were the Probability Estimates in *McDaniel v. Brown*?, 108 Mich. L. Rev. First Impressions 1 (2009), <http://www.michiganlawreview.org/assets/fi/108/kaye.pdf>.

The full 11-locus profile would occur in fewer than 1 in 15 billion unrelated individuals (*Appendix A*).<sup>1</sup> This full profile is likely to be unique among individuals in the western United States not closely related to the rapist.

Another possibility, however, is that some relative of Troy's was the rapist. Troy lived in a trailer with his brother, Travis. Another brother, Trent, lived in the same city, and two younger brothers lived on their parent's ranch in Loa, Utah. The most direct way to test the hypothesis of kinship is by DNA tests of the close relatives. One also can calculate the chance that an unsuspected relative would share the profile. Close relatives tend to share more DNA features ("alleles") than do unrelated individuals. The probability that a full sibling would have Troy's 11-locus profile is about 1 in 4500 (*Appendix B*). This sibling-random-match probability (SRMP) is orders of magnitude larger than the URMP, but the match remains quite unlikely.

## II. THE DNA MATCH AND TRANSPOSITION

At trial, the state did not present all the numbers given above. On direct examination, it focused on the 5-locus VNTR match and the 1-in-3-million figure. On redirect examination, however, the expert stumbled in describing probabilities.

### A. The Transposition Error

Romero misrepresented the conditional probability of a match to an unrelated individual as the probability that the DNA discovered in the victim's underwear was Troy's. The mistake occurred when the prosecutor asked for "the likelihood that the DNA found in the panties is the same as the DNA found in the defendant's blood." This is a "source probability"—the chance that Troy is the source given that his DNA profile matches. Using standard mathematical notation, this source probability can be written as  $P(\text{Troy} \mid \text{Match})$ . Romero agreed that "that percentage" could be obtained by subtracting 1 in 3 million from 1, and hence "would be 99.99967 percent." (Actually, there should be four nines after the decimal point.) Brown did not object to this adventure in arithmetic. Mathematically, the characterization of  $1 - \text{URMP}$  as a source probability treats 1 in 3 million as  $P(\text{Unrelated} \mid \text{Match})$ , the probability that an unrelated person is the source given the match. That is, Romero flipped around the hypothesis *Unrelated* and the data *Match* (*Appendix C*).

This transposition of the conditional probability can produce results that range from the approximately correct to the grossly inaccurate. Without discussing the extent of the mathematical error, Mueller's letter stated that this transposition was "so common it has been given a special name, the prosecutor's fallacy." The name is less than felicitous, since naive transposition does not always favor prosecutors (*Appendix C*). Indeed, the fallacy

---

<sup>1</sup> The appendices referenced in this Commentary are available at <http://www.michiganlawreview.org/assets/pdfs/additional-files/kaye-appendix.pdf>.

abounds in the statements of judges, defense counsel, and journalists. Statistics textbooks, evidence casebooks and treatises, and judicial opinions all caution against it. Consequently, the letter is hardly necessary for an appellate court to take cognizance of the transposition. The lower courts were therefore justified in considering the error regardless of whether the Mueller letter is officially part of the record.

### B. Bayes' Theorem

The misrepresentation at trial can be clarified by a correct application of Bayes' Theorem. Using the theorem, the Ninth Circuit railed against the transposition. Judge Wardlaw wrote that Romero's transposition was “especially profound given the weakness of the remaining evidence against Troy.” She explained that:

Statistically, the probability of [a source] given a DNA match is based on a complicated formula known as Bayes's Theorem, . . . and the 1 in 3,000,000 probability . . . is but one of the factors in this formula. Significantly, another factor is the strength of the non-DNA evidence.

But Bayes' theorem is not a “complicated formula.” It is derived in nearly every introductory text on probability or statistics. It has been discussed ad nauseum in law reviews. It states that the probability of a hypothesis changes with new information in the following simple way: *posterior odds = likelihood ratio × prior odds*. This equation applies when there are only two hypotheses as to the source—*Troy* or *Unrelated*.

The right-hand side of the formula is easily computed. The likelihood ratio (*LR*) is  $P(\text{Match} \mid \text{Troy})$  divided by  $P(\text{Match} \mid \text{Unrelated})$ . Troy's DNA will match if he is the source (and if there has been no laboratory or handling error); hence,  $P(\text{Match} \mid \text{Troy}) = 1$ . The probability of a match if the source is unrelated to Troy is the URMP. Therefore,  $LR = 1/(1/3,000,000) = 3,000,000$ . The match to Troy is 3,000,000 times more probable given that Troy as opposed to an unrelated person is the source. Meanwhile, the prior odds reflect the nongenetic evidence in the case. Suppose that before the DNA samples are tested, the odds of *Troy* (based on the other evidence in the case) are 1:1—it is as likely that Troy left the stain as that some unrelated person did. We multiply by the likelihood ratio to obtain posterior odds of 3,000,000:1. The corresponding probability is 3,000,000/3,000,001, or 99.999967 percent—as Romero said.

The problem is that the prior odds could be higher or lower than 1:1. The court of appeals wrote that transposition “could lead to serious error, particularly where the other evidence in the case is weak and therefore the prior probability of guilt is low.” But a very large likelihood ratio swamps even a low prior probability. For example, even if the other evidence were so weak that the prior odds were 1:1000, the posterior odds would be  $(1:1000) \times 3,000,000 = 3000:1$ . The corresponding probability of 99.96668 percent is smaller than Romero's 99.99967 percent, but the discrepancy hardly leaps out as a violation of due process.

### C. The Implications of Transposition

One might argue that even a slight numerical error due to transposition is constitutionally offensive because the witness's description of the "chance that the DNA . . . was from Troy" invites a more serious error. It encourages the jury to think that the source probability is 99.9+ percent even though the figure ignores the possibility that one of Troy's four brothers was the rapist, as well the other evidence in the case. Perhaps this is the point about other "factors in the formula." Under this view, the difficulty with the 99+ percentage in *Brown* is that it has too great a psychological impact on jurors.

But one can support (as I do) a rule of evidence excluding poorly explained and conceptually flawed computations of a source probability as unfairly prejudicial without concluding that a trial judge who fails to exclude such testimony and argument—despite the absence of any objection to it—commits constitutional error. The view that the testimony here was constitutionally impermissible because of its prejudice raises a host of questions. Will a jury hearing the 99+ percent figure be unable to reason effectively about the possibility of a brother or a mistake in handling the samples when defense counsel refers to these matters? A closing argument stating that a highly improbable match means that the defendant is the only person in a locality who realistically could be the source is not inherently unfair. Does this argument become constitutionally impermissible when the prosecution uses the transposed URMP to add that the match establishes a 99+ percent source probability? State and federal courts have allowed DNA analysts to testify that, to a reasonable scientific certainty, a defendant is the source of DNA recovered at a crime-scene. Is that testimony also unconstitutional?

### III. MISCALCULATING THE IMPACT OF TROY'S BROTHERS

The other problem with Romero's testimony is that it overlooked the possibility that the DNA on the victim's panties came from a close relative of Troy's. Troy did not want to advance this defense, but his lawyer asked on re-cross:

Q: Does that statistical probability change with brothers?

A: Yes.

Q: How does it change?

A: With a brother, there would be some genetic relationship. They have a 25 percent chance of sharing both alleles—both bands, and 50 percent chance of sharing one band.

Romero obtained these numbers from a 1992 National Research Council (NRC) report. Mueller's letter pointed out that:

This conclusion is only correct if both parents are heterozygotes and share at most one allele in common. For other possible parental pairs the probability of two sibs matching could be 50% or 100%. Thus, [she] has

chosen a special case which suggests that sibs have the lowest chance of matching that is biologically possible.

This criticism is of little moment. The “special case” of each parent having a distinct allele on each chromosome (heterozygosity) is the norm, but more importantly, the only thing the jury could learn from Romero’s incomplete explanation of identical genotypes by descent is that the chance of a match to a brother must have been far more likely than the URMP of 1 in 3,000,000. This take-home message would have been the same had Romero testified more precisely, adding that the chance of two brothers’ matching at a single locus could have been even larger than 25 percent.

Unfortunately, more serious problems arose when the trial judge asked, “Is there any way to help us with that . . . ?” Applying an equation in the 1992 report to the VNTR loci, Romero testified that the sibling RMP (SRMP) “turns out to be one in 6,500.” The prosecutor had Romero do more arithmetic regarding “the possibility of a brother.” Converting 1 in 6500 into 0.02 percent and showing that its complement was 99.98 percent, Romero agreed that “the likelihood of the parents having one child, and then the very next child having the same genetic code would be .02 percent.” When the defense asked whether that changed at all with two brothers, she answered, “No.”

The Scholars’ Brief presents this exchange as another manifestation of the transposition fallacy and an implicit denial of the obvious fact that the more brothers there are, the greater the chance that at least one will match. It is, however, an accurate statement of the probability for “the very next child.” True, the witness is not amplifying on the implications of her limited statement, but under conventional legal doctrine, she is not required to.

The error in the testimony lies in the figure of 1 in 6500 itself. As Mueller wrote:

Even if we assume that 25% is the proper number to use in this calculation the chance of two brothers matching is  $(0.25)^5 = 1$  in 1024 not 1 in 6500. [T]he error made here by Ms. Romero tends to suggest that the chance of two brothers matching is actually much less than it really is.

Again, it should not take Mueller’s letter for a court to find that the 1 in 6500 figure is wrong. The transcript establishes that Romero was using the 1992 report’s equation, which gives the mean SRMP. It is not mathematically possible for that formula to generate a 5-locus SRMP smaller than 1 in 1024. These are not matters subject to reasonable dispute.

What is debatable is the applicable probability. The Mueller letter suggests that because Troy has four brothers, the SRMP is not the proper figure to use. The probability that at least one of the four untested brothers will match is about four times the SRMP. Furthermore, a better estimate of the SRMP can be obtained by taking into account the genotypes in this case. The standard formula requires estimates of the allele frequencies. These were not part of the trial record and probably are not judicially noticeable, but the letter uses them to arrive at the figure of 1 in 66 for the chance of a match to at least one of the four brothers. The Ninth Circuit was taken with

this discrepancy of “almost one hundred times the probability asserted by Romero.” The majority insisted that the trial presentation “ignored logical implications about Troy’s four brothers, each of whom lived in the general vicinity.” Likewise, the National Association of Defense Counsel’s amicus brief presents 1 in 66 as “the more accurate probability of a sibling match,” and the Scholar’s Brief lists it as a plausible choice for “the true probability.”

Maybe the difference between 1 in 6500 (0.05 percent) instead of 1 in 66 (1.5 percent) is highly prejudicial, but the comparison is misguided. Romero’s 1 in 6500 was supposed to be the SRMP. It should be compared to a properly computed SRMP. For the VNTR loci, the SRMP is about 1 in 263. Adding the allele frequencies not introduced at trial, this probability is 1 in 4563 for all the loci tested. (*Appendix B*). In contrast, the 1 in 66 figure touted in the opinion and the briefs is the probability of a match to *at least* one of Troy’s four brothers. From a Bayesian perspective, this cumulative probability is misleading because it presumes that every brother is equally likely to be an assailant—which is absurd. The Ninth Circuit described the Tenth Circuit state of Utah as “neighboring” and the family ranch in Loa where the two younger brothers apparently lived with their parents as lying within “the general vicinity.” Yet, the driving distance from Loa to Carlin is over 440 miles. The relevance of 1 in 66 therefore rests on such speculations as a 13-year-old brother sneaking away from home to go hundreds of miles to another state. If it is necessary for the prosecution to elaborate on the SRMP at all in its case-in-chief, and if a cumulative probability is to be used (although a Bayesian calculation would be more suitable), then accounting only for the two adult brothers in Carlin seems more reasonable. The probability that one or both would match is approximately 2 in 263 (0.76 percent) for the VNTR loci and 2 in 4563 (0.044 percent) for all the loci. Arguably, the relevant comparison is between Romero’s figure of 1 in 6500 and one of these numbers. It is not between 1 in 6500 and 1 in 66.

#### IV. SUFFICIENCY OF THE EVIDENCE, PREJUDICE, AND DUE PROCESS

What should the Supreme Court do about the transposition of the URMP, the miscalculation of the SRMP, and the failure to compute a cumulative SRMP? Deplorable as much of this is, the Court is not deciding whether such testimony should have been excluded as plain error. For a federal court to grant a writ of habeas corpus, there must be more than a violation of a state rule of evidence. The lower courts basically excised the DNA match whose implications were poorly described and declared that due process was violated because “[t]here was insufficient evidence to convict the Defendant unless the DNA evidence established his guilt.” The dissenting judge on the Ninth Circuit panel applied less drastic surgery to conclude that a rational juror could find guilt beyond a reasonable doubt on the basis of the scientifically valid evidence along with the nongenetic evidence. Judge O’Scannlain maintained that, presented with the 1 in 3,000,000 URMP, the jury could rationally reject the unrelated-source hy-

pothesis. Moreover, even with a cumulative SRMP of 1 in 66, the jury could reasonably conclude that, in light of other evidence indicating the lack of involvement of any brother, Troy almost certainly was the source of the DNA.

But the requirement of proof beyond a reasonable doubt is not the due process value threatened by the distortions of the scientific evidence in *Brown*. The Ninth Circuit should have considered whether the DNA statistics were so misleading that it was fundamentally unfair to allow the trial to proceed without some corrective action. Essentially, *Brown* is not a case about the sufficiency of the evidence of guilt. It is a case of prejudice in how sufficient evidence is presented. This issue is closer than the circuit court’s remarks suggest. The discussion here does not resolve it, but it clarifies the extent to which the testimony was false. If the comparison for due process purposes is between acceptable scientific testimony and the erroneous quantitative evidence, then a more careful evaluation of this disparity is necessary.