Maine Law Review

Volume 76 | Number 1 Article 3

January 2024

The New People v. Collins: How Can Probabilistic Evidence be Properly Admitted?

David Crump

THE NEW *PEOPLE V. COLLINS*: HOW CAN PROBABILISTIC EVIDENCE BE PROPERLY ADMITTED?

David Crump

THE NEW *PEOPLE V. COLLINS*: HOW CAN PROBABILISTIC EVIDENCE BE PROPERLY ADMITTED?

David Crump*

ABSTRACT

The California Supreme Court's decision in *People v. Collins* is a staple in Evidence casebooks. An innovative assistant district attorney in the trial court had presented a mathematician who applied probabilities to questions about the perpetrators' characteristics. The state supreme court disapproved the injection of an equation featuring what mathematicians call the "product rule." The opinion contains thank-goodness-we-escaped-that-disaster reasoning and condemnation of this use of mathematics with probabilities. But the court's analysis probably would be different if the case were decided today, as the "new" *People v. Collins*.

Therefore, this Article considers what the Author calls the new *People v. Collins*: that is, the *Collins* analysis as it would be presented now, as the *Collins* of the present day. The Article concludes that the California court's reasoning was wrong as viewed from today, even if the result is defensible. Its opinion relied on a one-sided characterization of the Assistant District Attorney's evidence and argument. The court's conclusions would have been better presented if they had included balancing in the manner of Evidence Rule 403, of the value of probabilistic reasoning against its tendency to mislead as weighed by the court. And the court declined to consider the principle that no one piece of evidence is required to prove the entire case, by its indicating that the mathematics could not by itself prove guilt.

INTRODUCTION

The California Supreme Court's decision in *People v. Collins*[1] is a staple in Evidence casebooks.[2] An innovative assistant district attorney in the trial court had presented a mathematician who applied probabilities to questions about the

* A.B. Harvard University; J.D. University of Texas School of Law. John B. Neibel Professor of Law, University of Houston Law Center.

1. People v. Collins, 438 P.2d 33 (Cal. 1987). *Contra* Rachals v. State, 361 S.E. 2d 671, 675 (Ga. App. 1987) (rejecting and distinguishing *Collins*). *See also infra* Section IV.B. There have been many citations to *Collins*, including in the last several years. *See e.g.*, Commonwealth v. Widel, No. 701 MDA 2020, 2020 WL 6306000, at *3 (Pa. Super. Oct. 28, 2020) (discussing "shoddy" presentation); Learned v. Ill. Worker's Comp. Comm'n, No. 5-18-0368, 2019 IL App (5th) 180368WC-U, at ¶ 31 (Ill. App. Ct. May 9, 2019) (similar); People v. Julian, 34 Cal. App. 5th 878, 886–87, 888–89 (Cal. App. 2019) (similar).

2. *See* PAUL F. ROTHSTEIN ET AL., EVIDENCE: CASES, MATERIALS AND PROBLEMS 356 (4th ed. 2013) (referring to *Collins*); CHRISTOPHER B. MUELLER ET AL., EVIDENCE UNDER THE RULES: TEXT, CASES, AND PROBLEMS 95 (9th ed. 2019) (same); DEBORAH JONES MERRITT & RIC SIMMONS, LEARNING EVIDENCE: FROM THE FEDERAL RULES TO THE COURTROOM 844–45 (5th ed. 2022) (same).

perpetrators' characteristics.[3] The state supreme court disapproved the injection of an equation featuring what mathematicians call the "product rule."[4] The opinion contains thank-goodness-we-escaped-that-disaster reasoning and condemnation of this use of mathematics with probabilities.[5] But the court's analysis probably would be different if the case were decided today, as the "new" *People v. Collins.*

Therefore, this Article considers what the Author calls the new *People v. Collins:* that is, the *Collins* analysis as it would be presented now, as the *Collins* of the present day. The Article concludes that the California court's reasoning was wrong as viewed from today, even if the result is defensible. Its opinion relied on a one-sided characterization of the Assistant District Attorney (ADA)'s evidence and argument.[6] The court's conclusions would have been better presented if they had included balancing in the manner of Evidence Rule 403, of the value of probabilistic reasoning against its tendency to mislead as weighed by the court.[7] And the court declined to consider the principle that no one piece of evidence is required to prove the entire case, by its indicating that the mathematics could not by itself prove guilt.[8]

This Article contains analysis that is new. Strangely, the phenomenon described in Section II.B, which this Author calls "disaggregation," is presented here for the first time, as well as the role of the defense in encouraging it. So is the analysis of ways to present product-rule equivalents in Part V. The actual application of Bayes' Theorem in Section I.D and Part V contains reasoning that has not been seen before. And there are many points of analysis throughout that are new, because the analysis here is of *Collins,* the new *Collins,* as seen from today's perspective.

Part I of the Article deals comprehensively with the basics of probabilities, including the product rule and Bayes' Theorem (which provides a way to update odds with new evidence). The Author's own original method of defining this Theorem, which is simpler than the usual presentation, can be found in Section I.D. Section II.A sets out the ADA's evidence and argument, including both the lay witnesses and the probability evidence. Section II.B describes a principal reason for the mathematical testimony: the tendency toward disaggregation of items of evidence that fit together, which is encouraged by defense lawyers as part of their legitimate function, but which logically requires an answer such as probabilistic evidence or argument. This Section also describes the way in which courts, less

3. *Collins*, 438 P.2d at 33, 36–37. In addition to recent decisions, see *supra* note 1, there has also been recent commentary on the *Collins* case. *See generally* James Klinedinst, *Probably a Good Case: Using Statistics to Help Juries Determine Comparative Negligence in Florida*, T.M. COOLEY J. PRAC. & CLINICAL L., 33, 44–46 (2015) (suggesting a "better way" to present evidence); Lisa Milot, *Illuminating Innumeracy*, 63 CASE W. RSRV. L. REV. 769, 769–70, 775 (2013) (criticizing "math avoidance" and noting that people do not process probabilities well).

4. *Collins*, 438 P.2d at 36–37.

5. The court's rhetoric in some places is excessive. *See, e.g.*, *Collins*, 438 P.2d at 39–41; *infra* text accompanying notes 71–76. Elsewhere the court labeled mathematics as "a veritable sorcerer in our computerized society"—a strange criticism. *Collins*, 438 P.2d at 33.

6. *See infra* Part III.

7. *See* FED. R. EVID. 403; *see also* CAL. EVID. CODE § 352 (West 2023) (similar); *see infra* note 97 and accompanying text.

8. *See Collins*, 438 P.2d at 40.

admirably, sometimes rely on disaggregation too, including the California Supreme Court in this very case. Section II.C describes the mathematical evidence.

Part III presents the California court's reasoning in the old *Collins* opinion. Part IV, then, reconsiders the court's opinion and contains this Author's own reasoning, which disagrees with several aspects of the court's reasoning, as seen from today. Part V suggests ways in which some kinds of product-rule evidence or argument could be presented to a jury without offending the *Collins* court's reasoning. It shows the new *Collins*, that is. A final part contains the Author's conclusions, which include the observations, first, that probabilities are good evidence, as can be seen in the example of DNA results, and second, that in considering probabilistic testimony, the courts should follow the balancing test usually applied to evidence, which the old *Collins* court did not consider at all.

I. A PRIMER ON PROBABILITIES

A. What Are Probabilities?

A probability is a number between zero and one that expresses the likelihood of an event occurring or a condition existing.[9] A zero probability corresponds to certainty that the event will *not* happen, while a probability of one means that the event is *certain to happen*.[10] The probability that Socrates is still alive today is close to zero, while the probability that the sun will rise in the east tomorrow is virtually one. Diagrammatically, probabilities might look like this:

0 (zero probability)> .001-.999 etc. >1 (certainty)

Flips of coins often are used to illustrate probabilities. If we have a so-called "fair" coin, one that is not loaded either way, and we flip it many times, we could expect that approximately half of the flips will turn up heads and half tails.[11] Mathematicians might hypothesize infinite flips, which would return exactly a half-and-half result. The probability of turning up heads (or tails), then, is ½, or 0.5. Probabilities are more often presented as decimals than numerator-and-denominator fractions.[12]

Consider, on the other hand, a coin that is not fair, meaning that it can be called "loaded." Let us imagine that we do many flips and find that, on average, the coin turns up heads eight times out of ten. The probability of heads, then, is 0.8.[13]

What about the probability of tails, if that of heads is 0.8? If P is the probability of an event happening, then the probability of its not happening is (1 – P).[14] Therefore, the probability of tails is (1 – P), with P standing for the probability of heads, or (1 – 0.8). The probability of tails, then, is 0.2.

If you were to go to the racetrack and bet, you would play against the proprietor's estimate of likelihoods, using what we refer to as "odds" of a particular

9. *See* DAVID CRUMP, HOW TO REASON 389–90 (2014).
10. *See id.*
11. *See id.*
12. *See id.*
13. *See id.* at 391–92.
14. *See id.*

horse winning. Odds are probabilities presented differently.[15] One-to-one odds is the same as a one-half probability, or 0.5. Odds can be converted to probabilities by installing the first number as the numerator of a fraction and using, as the denominator, the sum of the two figures.[16] Thus one-to-one (1:1) odds equals a probability composed of the first number, 1, as the numerator, and 1 + 1 = 2 as the denominator, resulting in a probability of ½ or 0.5.

B. The Product Rule: Combining Independent Probabilities

What happens if we plan to flip the coin twice and try to figure out the probability of seeing two heads? Now we need what is called the "product rule."[17]

The probability of heads on the first flip is 0.5. The probability of heads on the second flip is also 0.5. We multiply the two probabilities. The product is 0.25 (or the equivalent: we multiply ½ x ½ = ¼). This is the probability of two heads turning up. This is how the product rule works.[18]

This result is intuitive. If we flip two coins, only one possible result is two heads, or heads- heads. But there are three other possibilities.[19] One is that we throw the first, and it is heads, but the second is tails. The three possibilities, then, are heads-tails, tails-heads, and tails-tails. There is only one possible way to get two heads. One out of four possibilities is one-fourth, or ¼, or 0.25.

But there is an important qualification to add. The product rule works to mathematical exactness only if the possibilities are what we call "independent."[20] One flip of a fair coin is independent of another flip, and so the product rule works for this two-heads example. But when does the product rule not work so well?

C. Independence: When Does the Product Rule Work?

If the product rule doesn't work when the events are not independent, what does independence mean? Simply put, independence does not exist when the two events are correlated. In other words, the product rule does not work if two events are so related to each other that the probability of one makes the probability of the other more (or less) likely.[21]

Consider this example[22]: A researcher wants to know the probability of two events—the probability of a person (1) who has stage four brain cancer (2) living for five years. She has found the probability of a random person in the population having stage four brain cancer, and she also knows the probability of a random person in the population living for five years. She tries to multiply the two together to find the probability of a person with stage four brain cancer living for five years.

It won't work. This conclusion is intuitive. But mathematically, why won't it work?

15. *See id.* at 391.
16. *See id.*
17. *See id.* at 394.
18. *See id.*
19. *Cf. id.* at 397 (offering a similar example).
20. *See id.* at 394.
21. *See id.* at 394–95.
22. *Cf. id.* (providing a similar but simpler example).

The reason it won't work is that the two events, stage four brain cancer and living five years, are *not independent.* The presence of stage four brain cancer influences the other possibility, living five years. Obviously, stage four brain cancer will shorten many people's lives.[23]

But now we have a catastrophe. Nothing in nature is truly and perfectly independent from another characteristic. The product rule does not work exactly even with our coin flip experiments when we try them in real life. Inevitably, one side will be slightly heavier than another, meaning that there will be a tiny imbalance. If we use this coin in a flip experiment and ask whether there will be two heads in two tries, the product rule will be off by a very small amount.

Yet all kinds of matters are decided by the flip of a coin, ranging from sporting events to who does the laundry among roommates. Are the participants falling for a traditional solution that is flawed because the coin is imbalanced? No, because the coin is close to being a fair coin—a fair piece of metal, if you will—and the deviation from perfection is too small to worry about.

The same thing happens with other matters. Take DNA analysis, for example. We have no proof that one "allele," which is used in the analysis, is exactly and perfectly independent from another.[24] And the product rule is essential to reaching results. We do know that the subjects of analysis are close enough to independence, and so the analysis works.

What is the point, then? The events or characteristics that are the subject of a probability assessment must be independent. But when we say independent, we mean, close enough to independence to make slight correlations irrelevant, because again, nothing in nature is truly, exactly independent from something else.

D. Bayes' Theorem

A mathematical relation called Bayes' Theorem is sometimes used to consider evidence such as that in *Collins.* It is a way of evaluating the effect of new evidence on an existing piece of evidence, particularly a piece of subjectively considered evidence. Let us imagine that a suspect has been arrested in connection with a robbery, quickly and in the same vicinity. An eyewitness identifies the suspect as the perpetrator, but the circumstances lead a reasonable listener (such as a juror) to decide that the witness is entitled to lesser than 100% credibility—perhaps 80%, or perhaps merely 50%. But there is additional evidence. The suspect had on his person a two- dollar bill shortly after the robbery. And the cash taken in the robbery included a two-dollar bill.[25]

The hypothetical juror would probably view this circumstance as corroboration of the eyewitness's identification. But the juror would be very likely to undervalue the corroboration. This situation provides an example for considering Bayes' Theorem.

---

23. *Cf. id.* (providing a similar but simpler example).

24. *See* PAUL C. GIANELLI ET AL., SCIENTIFIC EVIDENCE 42–43 (6th ed. 2020) (discussing allele independence).

25. *See* HOW TO REASON, *supra* note 9, at 397–99 (analyzing a similar example). Technically, the possessive of Bayes should be Bayes's, but the literature generally treats it as Bayes', and this Article therefore follows this convention.

As a first step, we must figure out the likelihood that a random person in this area would have a two-dollar bill on his person. This figure is unlikely to be available to us as an exact number.

We will have to estimate it based on experience. Let us say that the odds of a person having this bill are a thousand to one, or 1,000:1.[26] As an important note, we have now ventured a figure that the California court in *Collins* would have condemned,[27] but without it, we would not be able to use this reasoning method at all.

Bayes' Theorem uses what is called a "likelihood ratio" to compute a relevant probability by combining the eyewitness's 80% credibility with the new two-dollar bill evidence.[28] The likelihood ratio is intuitive; it is the probability of seeing the new evidence if the proposition is true versus seeing the same evidence if the proposition is not true. Let us call the probability of seeing the evidence if the proposition is true P(eviftrue). The probability of seeing it if the proposition is not true, we call P(evifnot). Thus, the likelihood ratio, expressed in notation, is:

Likelihood ratio = P(eviftrue) / P(evifnot).[29]

At this point, this Author usually switches to using odds to signify the likelihood of both the initial odds (Odds(old)) and the new odds (Odds(new)).[30] This approach mixes odds and probabilities, but the two types are really the same concept, and this method simplifies the mathematics. And Bayes' Theorem then emerges to be expressed this way:

Odds(new) = [Likelihood ratio] x [Odds(old)].[31]

So that,

Odds(new) = [P(eviftrue) / P(evifnot)] x Odds(old).[32]

And this is Bayes' Theorem. It is possible, of course, to express the theorem with probabilities only, but the equation becomes far more complex,[33] and this is the reason for substituting odds, which make a simple equation. Then, applying the theorem to the case in which the eyewitness evidence is discounted to 80%[34]:

Odds(new) = (1 / .001) x Odds(old) Odds(new) = 1,000 x 8:2
Odds(new) = 8,000:2 = 4,000:1,

so that the result of applying Bayes' Theorem is odds of 4,000 to 1. This figure corresponds to a probability of 4,000 / 4,001, or in other words, a probability of 0.99975,[35] close to 1.000 (certainty). Without using the mathematics provided by

26. *See infra* Section I.D (discussing and justifying this assumption).
27. *See infra* Part III (discussing the court's analysis).
28. *See* HOW TO REASON, *supra* note 9, at 399–402.
29. *See id.*
30. *See id.*
31. *See id.*
32. *See id.*
33. *See id.* at 401 n.3 (setting out the more complex equation).
34. *Cf. id.* at 402 (applying Bayes' Theorem to a different but similar example).
35. *See* HOW TO REASON, *supra* note 9, at 391 (explaining conversion of odds to probabilities); *see also supra* text accompanying note 16.

Bayes' Theorem, the hypothetical juror probably would underestimate the effect of the corroboration.

II. THE PEOPLE'S EVIDENCE AND ARGUMENT IN *COLLINS*

The description of probabilities above is precisely the basis of the allegedly offending evidence and argument that the ADA used in *Collins* and that the court condemned. But first, the non-mathematical evidence.

A. The Eyewitnesses in the Collins *Case*

The case arose because eyewitnesses described a blonde woman with a ponytail who robbed an elderly woman using a cane by knocking her down and stealing her purse.[36] The woman was quickly picked up by a Black man in a yellow automobile who had a beard and mustache.[37] One eyewitness later identified Janet Collins as the perpetrator, and the man with her turned out to be Malcolm Collins.[38]

But the identifying eyewitness's testimony was impeached. He had earlier picked Malcolm out of a lineup.[39] At the time, he made a statement to the effect that he was not sure of the identification of Malcolm.[40] The man had had a beard at the time of the attack but was shaved at the time of the lineup.[41] In other words, the witness was typical of many at lineups, seeing the subject with an altered appearance inside the police station for the first time after a violent event and in an unfamiliar, artificial setting. At trial, the eyewitness testified that Janet was the perpetrator.[42] And he identified Malcolm with certainty.[43] The defendants testified to an alibi.[44]

The victim of the assault testified that the woman who had robbed her had a blonde ponytail and other characteristics that fit the testimony of the other witness. She was unable to identify either Janet or Malcolm, but testified to her characteristics, as did the other eyewitness.[45] When officers appeared to arrest the defendants, Malcolm ran out the back of their home, hid behind a tree, and eventually was found in the closet of a neighboring home.[46] Defendants were Mirandized and questioned. "[T]he whole tone of the conversation evidenced a strong consciousness of guilt on the part of both defendants who appeared to be seeking the most advantageous way out."[47]

36. People v. Collins, 438 P.2d 33, 34 (Cal. 1968).
37. *Id.*
38. *Id.*
39. *Id.*
40. *Id.*
41. *Id.*
42. *Id.*
43. *Id.*
44. *Id.*
45. *Id.*
46. *Id.* at 35.
47. *Id.* at 36.

At this point, the ADA had an eyewitness identification together with corroborating details from another eyewitness. But would the evidence be sufficient in this condition to provide the jurors with proof that would convince them beyond a reasonable doubt? The impeaching evidence raised that stark question.

But the ADA actually had more evidence. A witness had described a blonde woman with a ponytail and a yellow car with a Black man who had a mustache and beard.[48] One can infer that interracial couples were uncommon at that time. These characteristics fit the descriptions of Janet and Malcolm Collins. How often would an observer encounter all of these characteristics together: a blonde woman, a ponytail, a yellow car, a Black man, with a beard, a mustache, and an interracial couple?

The ADA thought that the probability of all of this evidence coming together was extraordinarily low. And he thought that the jury needed mathematics in the form of probability analysis, to see just how rare the occurrence was.

B. The Disaggregation Tendency: Why Probability Analysis Was Needed

The trouble with the evidentiary weight of this unusual combination was the tendency of human jurors toward what this Author calls "disaggregation." People tend to view corroborating facts in isolation instead of assessing them in combination, or as an aggregation. In fact, as one commentator has written, this phenomenon is supported by psychological research. People in general are not good at reasoning about the manner in which corroborating facts combine:

> "Research demonstrates that people do not process probabilistic information well, that in the face of particularistic information, they cannot integrate the statistical and anecdotal evidence and consequently tend to ignore the statistical information. Intuitive, heuristic, human decision makers must dispense with certain information, and that tends strongly to be the quantitative information. While commentators' arguments have been that the data are inordinately persuasive, the evidence says the reverse is true."[49]
>
> This reasoning resonates with the actual experience of jurors in Collins, who later reported that they had disregarded [the mathematician's] testimony in reaching the verdict, focusing instead on the evidence provided by eyewitnesses to the crime.[50]

In other words, instead of *overvaluing* evidence, like the powerful effect of multiple corroborating factors in *Collins*, jurors are likely to do the opposite: to *undervalue* it. They may tend to reason that "a yellow car doesn't prove guilt beyond a reasonable doubt" and, of course, neither do any of the other characteristics of the Collins couple, taken singly. In fact, the *Collins* court itself lapsed into this kind of disaggregation and fallacious reasoning.[51]

---

48. *Id.* at 34.

49. Milot, *supra* note 3, at 775 (quoting Michael J. Saks & Robert F. Kidd, *Human Information Processing and Adjudication: Trial by Heuristics*, 15 L. & SOC'Y REV. 123, 149 (1980-1981)) (criticizing "math avoidance" and pointing out that people do not process probabilities well).

50. Milot, *supra* note 3, at 775.

51. *See Collins*, 438 P.2d at 40.

And, of course, skillful criminal defense attorneys encourage this tendency toward disaggregation. The Author, who tried many criminal cases before juries in his earlier career, saw this tactic used most prominently in intoxicated driving cases. The defense would be that the defendant wasn't really intoxicated or under the influence. Each piece of evidence could be answered in isolation. A paraphrase might be, "Well, so they say my client talked like he had a wad of cotton in his mouth. That doesn't prove anything, because a lot of people talk like that late at night when they're tired."

And the argument would proceed to isolate each of the other pieces of evidence in a similar manner. "They say he stumbled, but that doesn't prove anything. He tripped over uneven ground. They say he smelled like a brewery, but he tells you that he'd had two beers much earlier, so that doesn't prove anything either." Similar disaggregating reasons would be leveled at the evidence of the defendant's bloodshot eyes, incontinence, weaving while driving, and incoherence. This description of the tactic is not a criticism of defense lawyers. We expect the defender to defend his client by means that connect with jurors.

And even courts have used the same kind of dubious analysis. In fact, the *Collins* court made this mistake. It reasoned that the factors, taken in isolation, could not prove the entire case. For instance, the court said that the crime could have been committed by "a light-skinned Negress."[52] And the Black man could have been "wearing a false beard as a disguise."[53] And there were other instances of this kind of sophistry.[54] This method, of taking the combined characteristics singly and answering each one as if it were the only proof, is exactly the kind of disaggregation that uninformed jurors can be encouraged to use.

But the argument ultimately is fallacious precisely because it disaggregates. And one might well conclude that the same tactic could have been effective at trial in *People v. Collins*. Even if the defender did not encourage disaggregation in actuality, the natural appeal of this kind of reasoning probably would have influenced the jurors. The ADA needed to counter this tendency with evidence or argument encouraging evaluation of the combination of circumstances showing the Collinses to be the perpetrators, rather than taking each piece of evidence separately.

In other words, the ADA needed to encourage the jurors to use the product rule or equivalent reasoning.

C. The Mathematical Evidence and Argument in Collins

The *Collins* court described the allegedly offending testimony as occurring when the ADA called "an instructor in mathematics at a state college."

> The witness testified . . . to the "product rule," which states that the probability of the joint occurrence of a number of [m]utually independent events is equal to the product of the individual probabilities that each of the events will occur. Without presenting any statistical evidence whatsoever in support of the probabilities for

52. *Id*.
53. *Id*.
54. *See id.* at 40–41; *see also infra* text accompanying notes 75–76 (providing additional examples).

> the factors selected, the prosecutor then proceeded to have the witness assume probability factors of the various characteristics which he deemed to be shared by the guilty couple and all other couples sharing such distinctive characteristics.[55]

The prosecutor asked the mathematician to assume these probabilities: 1/10 for a partly yellow automobile; 1/4 for a man with a beard; 1/10 for a woman with a ponytail; 1/3 for a woman with blonde hair; 1/10 for a Black man with a beard; and 1/1000 for an interracial couple together in a car.[56] These assumptions, of course, were adapted to the time in question, the late 1960s, when beards and interracial couples were less common than they are today. The ADA candidly told the jurors that these probabilities were "estimates" and told the jurors to substitute their own estimates if they wished.[57]

The ADA then had the witness apply the product rule to these estimated probabilities.[58] He expressed the result in odds, concluding that there was but one chance in twelve million that any couple possessed the distinctive characteristics of these defendants. He also described the estimates as "conservative," meaning that the odds were even longer, "so that, in reality, 'the chances of anyone else besides these defendants being there, . . . having every similarity, . . . is somewhat like one in a billion.'"[59] This statement, if offered during argument, would have been made as combined with all other evidence—eyewitness identification, defendant's flight, and post-arrest Mirandized statements expressing strong consciousness of guilt—as well as the characteristics and probability evidence.

The trial court understood the relevance of the evidence. In response to the defendant's motion to strike, the trial judge said that the testimony had been received only for the "purpose of illustrating the mathematical probabilities of various matters, the possibilities for them occurring or re-occurring."[60] The court of appeals never indicated that it recognized this purpose. On appeal, one justice disagreed with the court's opinion, saying "I dissent. I would affirm the judgment in its entirety."[61]

This was the evidence and argument that prompted the California court to reverse Malcolm Collins's conviction. Janet Collins did not appeal, so Janet's conviction stood.

III. THE COURT'S OPINION IN *COLLINS*

The court identified three defects in the mathematical evidence. First, it said, the probability estimates were never proved to be accurate.[62] (Of course they were not, and the jury could not have possibly thought that they were intended to be, as

55. *Collins*, 438 P.2d at 36–37.
56. *Id.* at 37 & n.10.
57. *Id.* at 38.
58. *Id.* at 36.
59. *Id.* at 37.
60. *Id.*
61. *Id.* at 43.
62. *Id.* at 38.

the trial judge had correctly recognized,[63] that was not the point.) Second, the factors were not shown to be independent.[64] (Again, that was not the point.[65]) And third, the result was not explained as a probability that a random couple would have exhibited the same characteristics, and the jury was not given a basis to understand that the result was not a probability of guilt.[66]

As to the first alleged defect, that there was no evidentiary proof of the estimated probabilities, the court reasoned as follows:

> [W]e find the record devoid of any evidence relating to any of the six individual probability factors used by the prosecutor and ascribed by him to the six characteristics [of the defendants]. To put it another way, the prosecution produced no evidence whatsoever showing . . . that only one out of every ten cars which might have been at the scene of the robbery was partly yellow, that only one out of every four men who might have been there wore a mustache, that only one out of every ten girls who might have been there wore a ponytail, or that any of the other individual probability factors listed were even roughly accurate.[67]

Again, this was not the point, and it is impossible to imagine that the jurors were under a false belief that the "estimates" given by the ADA were supposed to be mathematically accurate or that they were intended to be understood as such.

In this regard, the ADA told the jurors that his estimates were just his own estimates. In fact, he "invited the jurors to substitute their 'estimates' should they wish to do so."[68] This statement effectively told the jurors that it was the method of reasoning, not the exact result, that was at issue. But to the ear of the court, it was a "curious circumstance" that the ADA had said this after offering his own estimates.[69] And instead of seeing the reasoning for what it was, the court showed how badly it had missed the point. "We can hardly conceive of a more fatal gap in the prosecution's scheme of proof," said the court.[70] This remark assumed that the jurors thought these were mathematically exact figures: an incredible assumption.

The second point was that the factors were not independent, as is required for use of the product rule. The court condemned the evidence as a "glaring defect"[71]:

> No proof was presented that the characteristics selected were mutually independent, even though the witness himself acknowledged that such a condition was essential to the proper application of the 'product rule' or 'multiplication rule.' . . . To the extent that the traits or characteristics were not mutually independent (e.g. Negroes with beards and men with mustaches obviously represent overlapping categories), the 'product rule' would inevitably yield a

63. *Id.* at 37; *see also supra* text accompanying note 60. The trial judge recognized, as the state supreme court did not, that the evidence was helpful in showing how the probabilities should be evaluated by the jury.

64. *Collins*, 438 P.2d at 39.

65. *Id.* at 37; *see also supra* text accompanying note 60.

66. *Collins*, 438 P.2d at 40–41

67. *Id.* at 38.

68. *Id*.

69. *Id.*

70. *Id.* at 39.

71. *Id.*

> wholly erroneous and exaggerated result even if all of the individual components had been determined with precision.[72]

The court's statement of the independence requirement is accurate, as far as it goes. But again, it misses the point, because the evidence was never intended to present an exact result, nor could it possibly have been understood by the jury as presenting an accurate result.

The court's third criticism of the evidence was that it had not been presented as a probability that a random couple would share the described characteristics, as opposed to a probability of guilt. "At best," said the court, the evidence "might yield an estimate as to how infrequently bearded Negroes drive yellow cars in the company of blonde females with ponytails."[73] And thus, surprisingly, the court expressed the intention of the purpose exactly. But then the court proceeded to label this "entire enterprise" as "gravely misguided"[74] and added:

> The prosecution's approach, however, could furnish the jury with absolutely no guidance on the crucial issue: [o]f the admittedly few such couples, which one, if any, was guilty of committing this robbery? Probability theory necessarily remains silent on that question, since no mathematical equation can prove beyond a reasonable doubt (1) that the guilty couple in fact possessed the characteristics described by the People's witnesses, or even (2) that only [o]ne couple possessing those distinctive characteristics could be found in the entire Los Angeles area . . . . [T]he most a mathematical computation could ever yield would be a measure of the probability that a random couple would possess the distinctive features in question. In the present case, for example, the prosecution attempted to compute the probability that a random couple would include a bearded Negro, a blonde girl with a ponytail, and a partly yellow car; the prosecution urged that this probability was no more than one in [twelve] million. Even accepting this conclusion as arithmetically accurate, however, one still could not conclude that the Collinses were probably *the* guilty couple. On the contrary, as we explain in the Appendix, the prosecution's figures actually imply a likelihood of over 40 percent that the Collinses could be 'duplicated' by at least [o]ne other couple who might equally have committed the San Pedro robbery. Urging that the Collinses be convicted on the basis of evidence which logically establishes no more than this seems as indefensible as arguing for the conviction of X on the ground that a witness saw either X or X's twin commit the crime.[75]

This reasoning was another dubious effort by the court to require a particular piece of evidence to prove the case all by itself. The infrequency of the coincidence of this combination of events, said the court, did not allow one to "conclude that the Collinses were probably *the* guilty couple."[76]

But that was not the point. This one single piece of evidence did not have to prove the case all by itself. This reasoning was another deviation from the court's responsibility. The court itself had thus fallen for the disaggregation fallacy. It was enough that the infrequency of the defendant's characteristics contributed,

72. *Id.*
73. *Id.* at 40.
74. *Id.* at 39–40.
75. *Id.* at 40–41.
76. *Id.* at 40 (emphasis added).

even if slightly, to the proof of the case. And, of course, it was intended to be combined with the eyewitness testimony, defendant's flight, and defendants' statements showing consciousness of guilt, in which case one could indeed conclude that the Collinses were *the* guilty couple.

IV. DEFECTS IN THE COURT'S ANALYSIS

A. The Court's Omissions

The court failed to recognize the disaggregation problem. At no point did the justices consider the principal value in the evidence and argument: that of countering the tendency of the jury to consider each piece of circumstantial evidence separately.[77] The mathematical evidence was designed not to show a concrete or exact probability, but to encourage the jurors to consider all of the evidence together, with each piece creating ever greater evidence of the improbability of the circumstances occurring together. And the cumulation was not in terms of addition but by the greater increase of multiplication.[78]

And the court compounded this omission by committing the error of disaggregation itself. Its own misleading argument was to the effect that mathematical evidence could not allow one to "conclude that the Collinses were probably *the* guilty couple."[79] The ADA did not suggest such a conclusion, and the court should not have demanded it.

Furthermore, the court failed to look at the evidence as including eyewitness identifications, which the mathematical method corroborated. If one were to invoke Bayes' Theorem, one could assign even very low odds to the eyewitness identification and still produce compelling evidence that the Collinses were *the* couple who committed the crime.[80]

B. The Court's Three Objections

The court treated the ADA not as an innovative and enterprising attorney, generating a new argument as we learned to do in law school, but rather as though he were a disobedient student. It criticized him for failing to offer proof that the estimates were accurate, which would have been virtually impossible and was not the point; it criticized him for failing to offer proof that the estimates were independent, which again would have been virtually impossible; and it criticized him because the result was not a probability of guilt, when no one had claimed it was.

1. Evidence of the Probability Estimates

The court insisted that one type of error in the mathematical evidence was that the ADA did not offer evidence showing that "one out of ten cars . . . was partly

---

77. *See supra* Section II.B.

78. The trial judge recognized this purpose of the evidence even though the justices did not. *See* Collins, 438 P.2d at 37; *see also supra* text accompanying note 60.

79. *Collins*, 438 P.2d at 40 (emphasis added).

80. *See infra* Section IV.D.

yellow," or that any of the other probabilities he suggested were accurate. The court criticized not only the evidence, but the ADA personally. "The bare, inescapable fact is that the prosecut[or] made no attempt to offer any such evidence."[81]

Presumably, then, if the ADA had made it clear that the probabilities were not intended to be precise, but rather, were illustrative of the manner in which the circumstances cumulated, the court would have withheld this criticism. But the trouble with this conjecture is that the ADA *did* make that point clear. The court labeled his effort to avoid overstatement a "curious circumstance."[82] The court viewed the ADA's invitation to the jurors to substitute their own estimates as an admission that the evidence was a terrible error—indeed, as literally the worst error imaginable—stating that it could "hardly conceive a more fatal gap" in the "scheme of proof."[83]

Actually, contrary to the court's criticism, it is a good thing that there was no such evidence. The trial, even as it was, consumed seven days.[84] If approximations of all of the circumstances had been offered through witnesses in the exacting manner of jury trials, including frequent objections, recesses to consider them, cross-examination, and detailed inquiries into the manner of estimation of the probabilities, the trial likely would have doubled in length.

In other words, the court's reasoning was one of many encouragements of lengthier trials.

And this kind of impetus has resulted in fewer and fewer trials. Today, jury trials are vanishingly rare.[85]

Furthermore, the ADA was encouraging the jury to perform its proper job. The jury's function includes the use of its members' own experiences.[86] That is why the ADA told the jury that it was free to substitute its own estimates for his own. If a juror, in the jury room, had told others, "Yellow cars are rare, perhaps one in ten," and had combined this probability with others by multiplying, the juror would have been performing the traditional, and indeed essential, responsibility of jurors.

2. Independence

"But, as we have indicated," said the court, "there was another glaring defect in the . . . scheme of proof."[87] There was no evidence that the factors were "mutually independent, even though the witness himself acknowledged that such condition was essential to the proper application of the 'product rule' or

81. *Collins*, 438 P.2d at 38.

82. *Id.*

83. *Id.* at 39.

84. *Id.* at 36.

85. Statistically, jury trials are few. *See* DAVID CRUMP ET AL., CASES AND MATERIALS ON CIVIL PROCEDURE 831 (7th ed. 2019) (citing Samuel R. Gross et al., *Don't Try: Civil Verdicts in a System Geared to Settlement,* 44 UCLA L. REV. 1, 2–6 (1996)) (showing figures for California superior court civil jury trials between 1990–1991).

86. *See, e.g.,* Mikolajczyk v. Ford Motor Co., 231 Ill. 2d 516, 554 (2008) (stating that jurors may rely on their own experiences to make determinations).

87. *Collins*, 438 P.2d at 39.

'multiplication rule.'"[88] Thus, concluded the court, to the extent that independence was lacking, "the 'product rule' would inevitably yield a wholly erroneous and exaggerated result."[89] The court gave the example of "Negroes with beards and men with mustaches," which "obviously represent overlapping categories."[90]

Since this fact was "obvious," presumably the jurors were able to perceive it. But mathematical exactness was not the point. The jury could not have failed to understand that the calculation was an example of the method, rather than the calculation of an exact probability, just as the trial court recognized.

And again, nothing is ever perfectly independent in nature. Even a coin-flip experiment, which is ubiquitous in the explanation of probabilities, suffers from this defect. Although theoretically a coin flip repeated infinitely shows half heads and half tails, tiny differences in weights of the two sides would deflect real coins flipped a finite number of times, so that the iconic probability experiment could not itself be based on independent events.[91]

A counterexample can be found in DNA evidence, which is treated by a "library" of ethnic possibilities, with each quality assigned a probability that is multiplied according to the product rule against other qualities.[92] One can surmise that no proof is typically offered to the jury of the various probabilities. And one can imagine that no proof is offered of the independence of the probabilities. The jury is given a probability of one in multiple millions or billions of the circumstances occurring at random.[93]

The court is correct that mustaches and beards are not independent. But the multiplicand that represented this lack of independence in the *Collins* evidence was one-fourth, which is vanishingly small in comparison to other multiplied factors. If there were to be found an error here, it would be harmless. And it is even less significant when one realizes that the point was not the obtaining of an exact probability, but rather the demonstration of how rare circumstances

combine. This is why the ADA told the jurors that they could substitute their own probabilities for the ones he had offered.

3. Not a Probability of Guilt

The court also faulted the evidence because it did not reflect a probability of guilt, but rather a probability of a random couple with all the characteristics of the couple at the crime scene. The court even took the estimate of one in twelve million as indicative of a 40% probability that these qualities were shared by at least one other couple in the metropolitan area.[94] The court therefore charged that the ADA's "entire enterprise" was "gravely misguided."[95]

---

88. *Id.*

89. *Id.*

90. *Id.*

91. *See supra* Section I.C. (explaining independence and its absence in real-world problems).

92. *See* Powers v. State, 343 S.W.3d 36, 45 n.14 (Tenn. 2011) (explaining DNA evidence and its submission).

93. *See id.* at 41.

94. *Collins*, 438 P.2d at 40. The court added an appendix that, in a convoluted way, showed mathematically that if the pool from which the couple was drawn were to increase to twelve million,

Although the court's mathematics are correct as a mathematical proposition, the court's reasoning was misleading. First, the court could not point to anything said by the witness or by the ADA to the effect that twelve million represented a probability of guilt. Instead, the court should have combined its probability analysis with other evidence of guilt, which included identifications, flight evidence, and post-arrest statements indicating guilt. Those additional items of evidence, even if discounted, would have operated together with Bayes' Theorem to produce astronomical odds against another couple's, at random, chancing to be in the position of the Collinses.[96]

Under these circumstances, the court's analysis is equivalent to a concern that the jury might have overvalued the evidence. This concern is addressed in the federal courts by Rule of Evidence 403 by its recognition of a misleading tendency as a weight against admissibility.[97] But the trial court had implicitly held that the evidence did not violate the balancing test of Rule 403 or its California counterpart (California's Rule 352),[98] and the higher court did not conduct any such balancing process. It could not have, because it failed to recognize the probative value of the evidence as the trial court had.[99] An unprincipled weighing of evidence, to assert without balancing that jurors had overvalued it, would make jury trials impossible, because there is no category of evidence that cannot be viewed by a jury as having greater weight than it theoretically should.

At another point, the *Collins* court stated that the ADA's approach "could furnish the jury with absolutely no guidance on the crucial issue: Of the admittedly few such couples, which one, if any, was guilty of committing this robbery?"[100] The error here was the court's assumption, again, that the mathematical evidence was supposed to carry the ball alone across the goal line and score a touchdown. And again, the court disaggregated the evidence. When combined with eyewitness evidence and post-arrest statements, the mathematical evidence could have increased the proof that the Collinses were guilty of this robbery by showing how the corroborating circumstances should have been evaluated.

---

there would be a 41% probability of another couple with similar characteristics. *See id.* app. at 43. The court even unnecessarily added analogies using the transcendental number e. *See id.*

95. *Id.* at 39–40.

96. *See infra* Section V.D (containing a chart illustrating the use of Bayes' Theorem in a similar manner, with only one corroborating item).

97. Rule 403 mandates that evidence be excluded only if its misleading tendency "substantially outweigh[s]" its probative value. In other words, to warrant exclusion, the misleading aspects of the evidence must be not only greater, but "substantially" greater, than its value. *See* FED. R. EVID. 403.

98. *See supra* note 97 and accompanying text. *See also* CAL. EVID. CODE § 352 (West 2023) (allowing exclusion of evidence if probative value is "substantially outweighed" by its tendency to mislead). *See also infra* Section D of the Conclusion (showing how a proper analysis could have been followed).

99. *See Collins*, 438 P.2d at 40.

100. *Id.*

C. The Court's Other Criticisms of the Evidence and Argument

1. The Criticism That People Distorted the Burden of Proof

The *Collins* court also criticized the ADA for his argument to the jury about the burden of proof. This part of the court's opinion also was misguided. To see why, one must consider the nature of adversarial jury arguments from the defense side as well as the prosecution side. The *Collins* court did not do so, and it failed to tell us about the defense argument that the ADA was either answering or anticipating.

Defense lawyers have formulas that they offer juries to interpret or explain the meaning of proof beyond a reasonable doubt.[101] ADAs respond to these efforts with other formulas.[102] The typical criminal trial does not involve an instruction by the judge on the meaning of the beyond-a-reasonable-doubt standard,[103] and it should not, because any definition is sure to be misleading or to nudge the burden one way or the other.[104] In fact, the *Collins* court's own misguided suggestion, which was to couple the proof standard with a standard of "moral certainty,"[105] could be perceived by jurors as watering down the standard—or, alternatively, as raising it.

Although courts cannot readily explain the meaning of the proof standard, greater latitude is afforded attorneys on each side in explaining it.[106] The hope is that, after hearing from both sides, jurors are better able to understand both the importance of the standard and its meaning.[107]

Defense formulas include such nostrums as the civil-criminal distinction and the touchdown analogy, which sound like the following:

> Across the street, in the civil courthouse, they fight about money, and the "greater weight" of the evidence is enough. But here, this defendant's freedom is at stake, and the burden of proof is on the state to convince you, beyond a reasonable doubt, that he is guilty . . . .
>
> In other words, in a civil case, the plaintiff just has to carry the football across the 50- yard line. This is a criminal case, and so the prosecutor has to carry the ball all the way across the goal line and score a touchdown.[108]

And there are other varieties of arguments, including the "aircraft mechanic" explanation. Jurors are asked to imagine themselves as aircraft mechanics. The aircraft has a defect, but it's probably going to be all right on its journey. Still,

101. *See* DAVID CRUMP ET AL., CRIMINAL LAW: CASES, STATUTES, AND LAWYERING STRATEGIES 236–40 (4th ed. 2020) (collecting various explanations).
102. *See id.*
103. *See* Victor v. Nebraska, 511 U.S. 1, 5 (1994) (noting that no definition is required, as long as the beyond-a-reasonable-doubt standard is given).
104. *Cf.* Cage v. Louisiana, 498 U.S. 39, 41 (1990) (per curiam) (holding that instructions such as "grave uncertainty" or "actual substantial doubt" are unconstitutional).
105. *See Collins*, 438 P.2d at 41.
106. *See* CRUMP ET AL., *supra* note 101, at 241.
107. *See id.*
108. *See id.* at 236.

there's a chance that the defect will make the airplane crash. The jury is then asked, if the state's case were that aircraft, would you risk clearing it to fly?[109]

The prosecution's formulas are more prosaic and not nearly as punchy. For example, the jurors may be told that they do not leave their common sense behind when they go into the jury room.[110] Or, they might be told that the burden applies only to the elements of the offense and does not mean proof beyond a shadow of a doubt; it is a proof beyond a reasonable doubt and not beyond all doubt.[111] Alternatively, the explanation may be that the standard refers to a doubt based on reason.[112]

The *Collins* court criticized the ADA for an argument featuring the following statements about the burden of proof:

> [O]n some rare occasion an innocent person might be convicted. [But w]ithout taking that risk . . . life would be intolerable . . . because . . . there would be immunity for the Collinses, for people who chose . . . to go down and push old ladies down and take their money and be immune because how could we ever be sure that they were the ones who did it?[113]

The court's reaction was, "In essence this argument of the prosecutor was calculated to persuade the jury to convict defendants whether or not they were convinced of their guilt to a moral certainty [sic][114] and beyond a reasonable doubt."[115]

The court did not mention whether the ADA's argument was a response to a similar but contrary argument by the defense, as it appears to be. The defense lawyer was competent, as the court observed,[116] and it would have been unforgivable for them to omit emphasizing and explaining the burden of proof. Furthermore, the argument was factually accurate—there always is such a risk, and without that risk, no one would ever be convicted—and it seems less likely to have nudged the jury one way or the other than other standards given above. The court's criticism of this aspect of the argument seems inappropriate.

2. The Criticism That the Underlying Facts Were Not Proved "Conclusively"

One of the more illogical charges by the court is that the ADA did not adequately prove the underlying facts used by the mathematician: the characteristics displayed by the Collinses.[117] The court's reasoning seems to be that before any expert testimony can be built upon it, the underlying facts or data need to be proven absolutely, and that any conceivable criticism of eyewitness testimony prevents use by an expert. The court said "we observe that the prosecutor's theory of probability

109. *See id.* at 238.
110. *See id.*
111. *See id.* at 236.
112. *See id.*
113. People v. Collins, 438 P.2d 33, 41 (Cal. 1968).
114. The "sic" reflects the misleading nature of the moral certainty formula. *Id.* at 241; *see also supra* text accompanying note 105.
115. *Collins*, 438 P.2d at 41.
116. *Id.* at 42.
117. *Id.* at 38.

rested on the assumption that the witnesses called by the People had *conclusively* established that the guilty couple possessed the *precise* characteristics relied upon [in the mathematician's testimony]."[118]

Here, the court was quite wrong. Admitting the testimony of an expert witness does *not* rest upon "conclusive" proof of the "precise" underlying facts and data. This demand by the court would make expert testimony impossible. Experts by necessity must rely on facts established by other means, including evidence in the case, and a requirement of conclusiveness and precision in eyewitnesses could never be met. Instead, the law allows experts to testify on the basis of facts in evidence; and indeed, the Rules of Evidence allow experts to base opinions on facts "made known to them" even if the source is not admissible in evidence:

> An expert may base an opinion on facts or data in the case that the expert has been made aware of or personally observed. If experts in the particular field would reasonably rely on those kinds of facts or data in forming an opinion on the subject, they need not be admissible for the opinion to be admitted.[119]

The court's criticism was contrary to the law. This conclusion is not a major issue in considering the court's main reasoning, but it goes far to demonstrate the court's aversion to mathematical analysis of proof.

V. ARE THERE OTHER WAYS OF PRESENTING THIS KIND OF INFERENCE, SO THAT IT WITHSTANDS CRITICISMS?

Notwithstanding these answers to the court's criticisms, the court, rather than this Author, obviously remains the deciding entity. And the court's rejection of the probabilistic evidence before it is clear. But the point that the jury can and should consider the circumstances cumulatively, rather than disaggregating them, is a valuable point. How can this point be made in a way that might survive the court's criticisms?

A. Court Allowances of Probabilistic Evidence

One possibility is that a court might recognize and allow evidence that is probabilistic. This Article has already provided the example of DNA evidence.[120] Another example is proof of paternity as it existed before DNA analysis was available for the purpose, and also, today, with DNA.

The method was, first, to assume a probability of 0.5 based upon the mother's designation of the father: to assign equal weights to the possibility that the identification of the father was accurate and to the possibility that it was not.[121] The second step was to compare various details of the father's blood chemistry to those of the child and to use known quantities to compute the likelihood of each coincidence.[122] The third was to use Bayes' Theorem to compute a new

118. *Id.* at 40 (emphasis added).
119. FED. R. EVID. 703; *see also* CAL. EVID. CODE § 801 (West 2023) (similar).
120. *See supra* text accompanying note 24.
121. *See* HOW TO REASON, *supra* note 9, at 401 (demonstrating by an example).
122. *See id.*

probability.[123] This last quantity is called a "probability of paternity" and is recognized as such by the courts.[124]

B. Courts That Have Refused to Follow Collins

Yet another possibility is offered by the contrasting decision in *Rachals v. State*.[125] A court may simply refuse to accept the arguments and conclusions made in *People v. Collins*. The evidence in *Rachals* showed that the defendant, a nurse, had killed a patient by administering potassium chloride.[126] An epidemiologist named Dr. Adelle Franks, from the Centers for Disease Control, gave additional testimony. The defendant charged that Dr. Franks's testimony improperly injected probabilities, including, allegedly, a probability of guilt.[127] The evidence did not say that, explicitly:

> The [state] had requested assistance in evaluating the increase in the number of cardiac arrests occurring in the Phoebe Putney Hospital. Dr. Franks went to the hospital . . . and examined the hospital's records for the entire [previous] year. In three of those months, the hospital had experienced no cardiac arrests. In two months, four cardiac arrests had occurred in each month. Accordingly, Dr. Franks stated that the hospital should have between zero and four cardiac arrests in a normal month. However, in the month of November [], eleven cardiac arrests had occurred on the [three] o'clock to [eleven] o'clock shift. The probability of this occurring "by chance alone is less than one in a trillion." In the month of November, five cardiac arrests had occurred in one day and one patient had a total of eight cardiac arrests in that one month. Dr. Franks listed all cardiac arrest patients for the period investigated and the primary nurse on duty with that patient. Rachals was the primary nurse for [eleven] cardiac arrest patients in the month of November. No other nurse was the primary care nurse for more than one cardiac arrest patient. Dr. Franks charted all [twenty-four] nurses for that month and the number of cardiac arrests that occurred on their shift[s], and those that occurred when they were not on shift to calculate a "rate ratio." The "rate ratio" for most nurses was around one, while the "rate ratio" for Rachals "was 26.6, which means that in 26.6 times, it was more likely that a cardiac arrest would occur while she was on duty than when she was not on duty . . . . [T]he rate ratio show infinitely large and unmeasurable [sic] because all of the cardiac arrests that occurred on the [three] o'clock to [eleven] o'clock shift occurred while she was on duty." Rachals's counsel contends that this testimony impermissibly invaded the province of the jury, in that it concluded that if a crime occurred while Rachals was on duty, the odds were 26.6 to 1 that she did it.[128]

The court indicated a suspicion of probabilistic evidence, and it cited *People v. Collins* in this connection.[129]

---

123. *See id.*

124. *See, e.g.*, Griffith v. State, 976 S.W.2d 241, 245 (Tex. App. 1998) (recognizing admissibility of probability of paternity).

125. Rachals v. State, 361 S.E.2d 671 (Ga. App. 1987).

126. *Id.* at 674–76.

127. *Id.* at 674.

128. *Id.* at 674–75.

129. *Id.* at 675.

But the court also cited its own state supreme court, which had encountered probabilistic evidence and approved it.[130] The court had held that there was no error in that case, "'as experts are permitted to give their opinions, based upon their knowledge, including mathematical computations.'"[131] The *Rachals* court also distinguished the case before it from *Collins*:

> In the instant case, however, the number of variables was limited in the controlled situation, and the statistics were not derived from a random sampling. Rather, hospital records established the average number of cardiac arrests for the entire year prior to November 1985, when the dramatic increase in the number of cardiac arrests was noticed. The potential for analytical error that was apparent in *Collins* is absent in the instant case. When in doubt, evidence of this type should be admitted, unless based on sheer speculation . . . although it may be totally rejected by the jury.[132]

It seems doubtful, however, that the evidence was free of the criticisms cast on this kind of evidence by the *Collins* court, particularly that court's concern that the evidence might be taken as a probability of guilt. In other words, *Rachals* seems to provide authority for the use of probabilities in evidence.

C. Using Illustrative Argument

Nothing in *Collins* states that an attorney is in error in giving a final argument to the jury that tells how properly to cumulate corroborating evidence. For example, imagine that the attorney tells the jury:

> When you see a particular circumstance in a crime, and you see it coincide with other circumstances in the same crime, the evidence of both characteristics corroborates the eyewitnesses. And their combination provides much more corroboration than we would get just looking at each characteristic separately.
> This, ladies and gentlemen, is how we should consider the characteristics of the couple in this case. Consider the likelihood of a robbery being carried out by a couple in a yellow car. And the couple being a Black man and a blonde woman. The circumstances don't just add together. It's not an exact mathematical equation, but the circumstances, then, cumulate more powerfully than just their sum—than what we would get from merely adding them.
>
> How rare would you think it would be, that we would have a yellow car, a Black man, who has a beard, and is accompanied by a blonde woman, who has a ponytail? And for each of those unusual characteristics to come together and fit the individuals identified by eyewitnesses?

Presumably this argument is different from the evidence and argument in *Collins*. And presumably, it is more likely to be considered proper. It does not assume any unproven probabilities. It does not violate the requirement that probabilities be independent. And it does not lend itself readily to being considered as a probability of guilt.

130. *Id.*
131. *Id.* (quoting Williams v. State, 312 S.E.2d 40, 73 (Ga. 1983)).
132. *Id.* at 675.

On the other hand, this jury argument fails to capture the way in which occurrences actually cumulate. And it lacks forcefulness. It is the pale ghost of an argument of the kind that usually occurs in criminal courthouses, offered by both assistant district attorneys and defense lawyers.

D. Chart Presentation, to Emphasize the Method Rather Than Any Particular Probability

One possibility for the use of probabilities in court is to emphasize their meaning as a heuristic device.[133] The purpose of the evidence in *Collins* was not to tell the jury about a specific probability, but instead, to show how probabilities combine.[134] In other words, it was to show a method of reasoning from which the jurors could learn and apply the treatment of multiple corroborators: a heuristic.[135] This aspect of mathematical evidence could be enhanced by a presentation that shows many possible outcomes, varying with underlying assumptions, so that it is clear that no actual probability is the object. This kind of presentation might enhance the likelihood that the jury perceives the point, which is the method of thinking about rare events that cumulate.

Above, one can see the method used to compute probabilities in paternity cases.[136] The method begins with an assumption that the odds are one to one that the man identified by the mother is, in fact, the father.[137] This assumption of even odds would be written as odds of 1:1, corresponding to a probability of 0.5: an assumption of relatively poor credibility in the mother, indicating only a grudging kind of halfway confidence in her—a conservative assumption. This assumption then is combined with blood chemistry results by the product rule or Bayes' Theorem to produce a probability of paternity. This figure then is given to a jury if the case is contested.[138]

The method, rather than any particular result, could instead be provided to a jury in a chart that shows varying possibilities depending upon what one assumes initially.

Instead of the paternity example, let us consider a case in which an eyewitness has identified the defendant as having committed the robbery of a convenience store clerk. Further, let us consider that the eyewitness's testimony is corroborated by the finding of a $2.00 bill on the suspect's person, and that a $2.00 bill was taken in the robbery.[139]

---

133. A heuristic, in its broad sense, means a device that enables readers to learn something for themselves. The chart presented by the Author, *infra* p. 59, is a heuristic because it enables jurors to figure out the significance of the corroborating circumstances themselves, rather than simply telling them the answer.

134. *See supra* Section II.C (explaining the actual objective of the evidence).

135. *See* HOW TO REASON, *supra* note 9, at 404–05 (explaining the heuristic function of the chart presented, *infra* p. 59).

136. *See supra* Section V.A.

137. *Id.* (describing the probability of paternity process).

138. *See id.*

139. This example is adapted from a similar but different example provided in a book by the Author that includes treatment of probabilities. *See* HOW TO REASON, *supra* note 9, at 403–05.

Intuitively, one might see that the corroboration enhances the credibility of the eyewitness testimony. But the disaggregation phenomenon acts in strange ways to disrupt this intuitive thought. The defense would perform its responsibility, perhaps, by enhancing this disaggregation.[140] "A $2.00 bill? That doesn't mean he's guilty. A lot of people out in the population have $2.00 bills. Last time I checked, the $2.00 bill was still in circulation. Still legal tender." In this manner, the intuition focuses on the corroboration as a single piece of evidence that is required to prove the case by itself, rather than as combining with the eyewitness testimony to produce overwhelming odds.

So: What are the odds against a random person in the vicinity having a $2.00 bill? In response to the query, "Are $2.00 bills rare?," Google cites a source which says that "$2.00 bills account for just 0.001 of the $2 trillion worth of currency in circulation."[141] But we still will have to make assumptions. If $2.00 bills are one thousandth of the total circulation worldwide, it seems reasonable to estimate that the probability of any random person having a $2.00 bill in his possession is less than one thousandth. Not all bills in circulation are in the United States, not all are owned by individuals, and not all are on people's persons, so the probability is surely smaller.

But this estimate is probably not provable by evidence. If the *Collins* court's requirement of justifying evidence is to be applied,[142] this entire effort is unworkable. In fact, the generally accepted method of providing paternity likewise is inadmissible, because it starts with an unprovable assumption.[143] To prove the requisite facts about $2.00 bills, the proponent would need to commission a very expensive and lengthy sociological study, probably including a survey of many people subject to statistical constraints, done just for the single purpose of demonstrating the effect of the corroboration in this one case.

If this unworkable requirement is not imposed, the proponent can construct a chart displaying different possibilities of the eyewitness's credibility and their combination with the corroborating evidence. The paternity calculation assumes only 1:1 odds or a 0.5 probability of accuracy in the mother's identification of the father.[144] But what if the odds are different: say, 1:9 odds (1:9, meaning only a one-tenth probability of accuracy), or for that matter, 9:1 (9:1, meaning a nine-tenths probability)? Or somewhere in between, other than 1:1?

The proponents could then ask the mathematical witness to apply Bayes' Theorem to these odds in combination with the odds of the corroborating evidence occurring. The mathematician would need to explain Bayes' Theorem to the jury—a time-consuming endeavor, no doubt. The mathematician would apply this method to all one-tenth possibilities on the scale of odds of eyewitness accuracy.

---

140. *See supra* Section II.B. (explaining the disaggregation problem and the role of defense counsel).

141. Jess Catcher, *How Much Are Two Dollar Bills Worth? Discover If Your Rare Bill Is Worth Big Bucks*, YAHOO!LIFE (Oct. 2, 2023), https://www.yahoo.com/lifestyle/much-two-dollar-bills-worth-145915579.html.

142. *See supra* Part III (discussing this requirement).

143. *See supra* text accompanying notes 137–38 (explaining that the method begins with the assumption that the man at issue is the father has a 0.5 probability of being true).

144. *See id.*

The resulting chart might look like this:[145]

Figure 1: Chart for Heuristic Use of Bayes' Theorem	
Initial Odds of Identification Accuracy	*New Odds (rounded), after applying Bayes' Theorem to 1:1,000 Evidence*
0:1 (certainty of nonpaternity)	0:1 (zero probability)
1:9	111:1 (111/112 probability = .99107
2:8, or 1:4	250:1 (250/251 probability = .99602
3:7	429:1 (429/430 probability = .99767)
4:6, or 2:3	667:1 (667/668 probability = .99850)
5:5, or 1:1	1000:1 (1000/1001 probability = .99900)
6:4, or 3:2	1500:1 (1500/1501 probability = .99933)
7:3	2333:1 (2333/2334 probability = .99957)
8:2, or 4:1	4000:1 (4000/4001 probability = .99975)
9:1	9000:1 (9000/9001 probability = .99989)
1:0 (certainty of paternity)	1:0 (1.0 probability)

The chart demonstrates the intuitive result that the corroboration makes the odds of accuracy in the combined result better. It also shows how much better. Even if we have poor confidence in the eyewitness identifications, so that we assign the eyewitness only 1:9 odds of accuracy, the probability of accuracy increases to more than 99%. An initial probability of only one-half, or 0.5, combines with the corroborating evidence to produce 99.9%. And if one assigns 9:1 initial odds to the eyewitness, the combination produces 99.99% odds of accuracy.

But the most interesting feature of this presentation is that it does not focus upon any particular probability. Instead, by reflecting many probabilities, it focuses instead upon the method of reasoning that should combine the new evidence, the corroboration, with the eyewitness identification. Without it, one might well underestimate the impact of the corroborating evidence.

Furthermore, the presentation removes the tendency toward disaggregation that is described above.[146] This tendency is as natural and intuitive as the ability to understand that the corroboration increases the likelihood of accuracy. It can be expected that the defense attorney will encourage the tendency toward disaggregation. The *Collins* court did not consider this issue when it condemned

145. This chart is adapted from another, different one in a book by this Author that covers probabilities. HOW TO REASON, *supra* note 9, at 404.

146. *See supra* Section II.B.

the probabilistic evidence before it, and in fact, it engaged in inappropriate disaggregation itself.[147]

VI. The Unevenly Weighted Balancing Test: Fed. R. Evid. 403 or Cal. Evid. Code 352

The California court did not balance the arguments for and against admissibility under California Evidence Code Section 352.[148] Federal Rule of Evidence 403, which in relevant respects, is substantially equivalent to the California provision but more readily understood nationwide, reads as follows: "The court may exclude relevant evidence if its probative value is substantially outweighed by a danger of one or more of the following: unfair prejudice, confusing the issues, misleading the jury, undue delay, wasting time, or needlessly presenting cumulative evidence."[149] Under the rules, the evidence in *Collins* is certainly relevant, but it can be excluded by reason of its prejudice, misleading tendencies, delay, wasted time, or needless presentation of cumulative evidence. But the rule is unevenly weighted. Exclusion is authorized only if the counterweights "substantially outweigh[]" the probative value of the evidence.[150]

This consideration may be why the trial court in *Collins* refused to exclude the evidence. The evidentiary rules function so that we do not provide the jury with only bits and pieces of what is important. This is why Rule 403 and its California equivalent set up unevenly weighted balancing tests. They are loaded toward admissibility.

It would have been better if the California Supreme Court had followed the logic of the California trial court in recognizing the function of the evidence, as is indicated in Section II.C. of this Article. There was, indeed, probative value in the questioned evidence. The issue remained whether this value was "substantially outweighed" by the counterweights. The most relevant of these factors, considering the court's reasoning, was the concern about possible "misleading" of the jury.

Under the analysis in this Article, this concern would not have "substantially outweighed" the value of the evidence. But even if the California Supreme Court had decided to exclude the evidence, it would have been better if the court had described the alleged ways in which this misleading violated a rule that favored admissibility. An analysis of this kind would have told lawyers how they could properly use probabilistic evidence, both in criminal cases and in civil. It would have made clear how, for example, it is possible to use DNA evidence, which also requires probabilistic analysis.

147. *See supra* Section IV.B.3.
148. Cal. Evid. Code § 352 (West 2023).
149. Fed. R. Evid. 403.
150. *Id.*

CONCLUSION

A. Did the Collins *Court Miss the Point of the Evidence?*

In *Collins*, the prosecuting ADA used probabilistic evidence, as combined by the product rule, to demonstrate how the corroborating circumstances increased the likelihood of the defendant's guilt.[151] But the court provided three reasons for its rejection of the probabilistic evidence used by the ADA.[152] First, the assumed probabilities offered by the ADA were not supported by particularized evidence showing their accuracy.[153] Second, the characteristics that the mathematics expert treated with the product rule were not shown by evidence to be independent.[154] Third, the probability that resulted from such a calculation was not a probability of guilt, as the jury might have concluded; but instead, it was the probability that two random persons would have combined all of the characteristics that the evidence attributed to the guilty couple.[155]

The court unfortunately did not consider the proper purpose of the evidence, which was not the demonstration of any probability of the defendants' guilt. The evidence was, first, a counteractive to the tendency toward disaggregation. In other words, it opposed the propensity of the human mind to view each corroborating circumstance separately and to inquire whether it, in isolation, proved the entire case.[156] Defense attorneys, whose legitimate office is to challenge the evidence of guilt, usually encourage this tendency. "So what if the real guilty parties had a yellow car? Many people have yellow cars out there. That doesn't prove that these defendants are guilty." In fact, courts, including the *Collins* court, sometimes make the same mistake.[157]

In addition, the evidence showed the method by which probabilities combine. Instinctively, jurors must realize that corroborating evidence increases the effect of eyewitness testimony. But just as surely, one can consider that they vastly underestimate the enhancement brought by corroboration. In fact, the combination of different pieces of corroboration, like the yellow car and blonde ponytail in *Collins,* increases the reliability of eyewitness identifications very substantially.

In short, by failing to recognize the disaggregation tendency, or to consider how corroboration naturally can be undervalued, the *Collins* court missed the point.

B. Responses to the Collins *Court's Three Criticisms*

The court's criticism of the ADA's offer of probabilities was also dubious. The objective of the mathematician's testimony did not depend on precise, measurable statistical beginnings.

Instead, it was a demonstration of how different pieces of evidence combine.[158]

151. *See supra* Part II.
152. *See supra* Part III.
153. *Id.*
154. *Id.*
155. *Id.*
156. *See supra* Section II.B.
157. *See supra* Section III.
158. *See supra* Section IV.B.

The ADA did not present any of his assumed probabilities as having resulted from any sociological study or any similar source. The jury, listening to the mathematician's testimony, could not possibly have believed that they were. The ADA told the jury to substitute their estimates if they wished, and thus he called upon the jury to perform its traditional function of using life experiences to evaluate the evidence.[159] Furthermore, the furnishing of concrete evidence about ponytail frequency or other initial statistics would have required the commissioning of a precisely conducted survey or other methodological study.[160] This notion was easy for the court to suggest as a possibility, but it was not really possible. It would have meant an enormous expense for commissioning an expert study of each factor, and an unreasonable lengthening of trial, when precise figures were not even the point of the mathematical evidence.

The court's second point, that the factors treated by the product rule were not independent, has more weight to it, but not much more. The beard-and-mustache combination lacked independence, for example, but this was a minor part of the chain of inferences and harmless error if error at all.[161] One problem with the independence criticism is that nothing in nature is perfectly independent. The example of DNA evidence, which is readily received by courts if properly presented, shows that occurrences that are not perfectly independent can still be useful enough to provide admissible evidence.[162] The jury was capable of using its own observation from driving on streets to estimate the frequency of ponytails and of yellow cars. But the greater problem with the *Collins* court's independence reasoning, once again, was that the evidence was never presented as an accurate calculation. Instead, it was illustrative of the way in which corroborating occurrences combine.[163]

Finally, the court was theoretically correct that the mathematical evidence did not offer a probability of guilt. But it was not presented as a probability of guilt. Instead, it was presented in combination with the eyewitness testimony.[164] The court itself made the mistake of disaggregation, demanding that the probabilistic evidence prove the case all by itself.[165] The court was itself enthralled by the disaggregation mistake that the product-rule evidence was intended to counteract.

C. How Can Product-Rule Evidence Be Properly Presented?

Evidence theoretically identical to that in *Collins* is received by courts routinely as DNA evidence. It has been presented as evidence of paternity, with the same kind of initial assumption that the *Collins* court condemned. These types of evidence exhibit the same alleged flaws that the court held exclusionary.[166] In

159. *See supra* Section II.B.
160. *See supra* Section V.D.
161. *See supra* Section IV.B.2.
162. *See supra* text accompanying note 24.
163. *See supra* Section III.
164. *Collins*, 438 P.2d at 36.
165. *See supra* Section IV.B.3.
166. *See supra* Section V.A.

addition, some courts have disagreed with *Collins* and admitted evidence similar to that in the California case. *Rachals v. State* is an example.[167]

As a substitute, attorneys might consider giving the jury a non-mathematical statement during final argument about how corroborating circumstances combine.[168] This solution is inferior to the kind of evidence offered in *Collins.* But it may be the best that can be done.

Alternatively, an attorney might consider presentation of multiple possible probabilities depending on various possible assumptions. This kind of jury presentation is illustrated above as a chart.[169] This chart avoids what is perhaps the most salient criticism by the *Collins* court, which was that the mathematical evidence could have become confused with a probability of guilt.[170] The chart likewise avoids the criticism that the events are not independent and that they are not proved, because it does not purport to compute an actual probability. It is, instead, a heuristic device, by which a jury can learn how corroborating circumstances might combine.[171]

D. The Major Flaw in the Court's Opinion

The most significant defect in the *Collins* court's analysis, however, was its failure to recognize any positive value in the evidence it condemned. The trial court was more perspicacious, as is indicated in Part III above. The trial judge held the evidence admissible for the "purpose of illustrating the mathematical probabilities of various matters, the possibilities for them occurring or re-occurring."[172] And this was the point, not the "probability of guilt," that the state supreme court hypothesized. That court failed to recognize the actual purpose of the probability analysis: to show how multiple corroborating factors combine and to oppose the human tendency toward disaggregation of the kind described in Section II.B above.

If the court had recognized these purposes, it could have undertaken an analysis of the evidence in light of Rule 403 of the Federal Rules of Evidence or its California equivalent.[173] These rules allow for the balancing of probative value against the prejudicial or misleading nature of the evidence, with exclusion resulting only if the counter-factors "substantially outweigh[]" the probative value.[174] No one can know whether the jurors concocted an improper probability of guilt, because the evidence did not suggest it; it merely introduced an entirely hypothetical set of assumptions and a probability based on those, independently of the eyewitness testimony. The probative value was, as the trial judge said, to "'illustrat[e]'" the way in which corroborating factors combine.[175] The likelihood of

167. *See supra* Section V.B.
168. *See supra* Section V.C. (providing an example).
169. *See supra* Section V.D. (depicting such a chart).
170. *See supra* Section IV.B.3.
171. *See supra* Section V.D. (explaining use as a heuristic).
172. People v. Collins, 438 P.2d at 33, 37 (Cal. 1968).
173. FED. R. EVID. 403; CAL. EVID. CODE § 352 (West 2023).
174. *See id.*
175. *Collins*, 438 P.2d at 37.

guilt, then, was indeterminate, but the combined evidence made it overwhelming. The evidence was intended to serve this purpose—and to minimize the tendency toward disaggregation.

The California Rule provides, "[t]he court . . . may exclude evidence if its probative value is substantially outweighed by the probability that its admission will . . . create substantial danger of . . . misleading the jury."[176] In other words, even if the trial court had found that the "danger" that the jury might take twelve million as a "probability of guilt," the evidence would still not be excluded unless that hypothetical danger was much greater than the danger that jurors might discount the cumulative weight of the corroborative elements or disaggregate them. There really was little chance of the jurors taking the evidence as proving a probability of guilt, especially because the factor probabilities were presented as "assumptions" and because the ADA invited the jurors to substitute their own estimates.[177] The *Collins* court emphasized this aspect of the ADA's questions by italicizing the word "*assume*"[178] and by labeling the invitation to the juror to substitute their own estimates as a "curious circumstance,"[179] which it was not. Instead, one can conclude that the ADA simply treated the probabilities as illustrative and avoided the very fallacy that the court accused him of perpetrating.

Even if the court had found the evidence excludable, its opinion would have been better if the court had undertaken the proper balancing analysis. It would have recognized the possibility that jurors might underestimate the combined weight of the corroborating factors and shown how they actually cumulate. And it would have recognized the tendency toward disaggregation and the appropriateness of its avoidance.

This proper analysis probably would have given guidance to attorneys. It would at least have hinted at acceptable ways to show how evidence combines and to minimize disaggregation. And even if the evidence were to be excluded in *Collins* under this kind of analysis, the court's explanation of the result would have been better.

176. CAL. EVID. CODE § 352 (West 2023).
177. *Collins*, 438 P.2d at 36–37, 36 n.9.
178. *Id.* at 36.
179. *Id.* at 38.