Psychonomic Bulletin and Review (2018) 25:2380–2388
https://doi.org/10.3758/s13423-018-1465-2

BRIEF REPORT

Cognitive-psychology expertise and the calculation of the probability of a wrongful conviction

Jeffrey N. Rouder[1] · **John T. Wixted**[2] · **Nicholas J. S. Christenfeld**[2]

Published online: 8 May 2018
© Psychonomic Society, Inc. 2018

Abstract

Cognitive psychologists are familiar with how their expertise in understanding human perception, memory, and decision-making is applicable to the justice system. They may be less familiar with how their expertise in statistical decision-making and their comfort working in noisy real-world environments is just as applicable. Here we show how this expertise in ideal-observer models may be leveraged to calculate the probability of guilt of Gary Leiterman, a man convicted of murder on the basis of DNA evidence. We show by common probability theory that Leiterman is likely a victim of a tragic contamination event rather than a murderer. Making any calculation of the probability of guilt necessarily relies on subjective assumptions. The conclusion about Leiterman's innocence is not overly sensitive to the assumptions—the probability of innocence remains high for a wide range of reasonable assumptions. We note that cognitive psychologists may be well suited to make these calculations because as working scientists they may be comfortable with the role a reasonable degree of subjectivity plays in analysis.

Keywords Probability · Law · Bayes theorem · Decision-making · Subjectivity

In the course of studying human perception, cognition, and performance, cognitive psychologists often become experts in decision-making in noisy environments. This expertise is applicable to many real-world domains. One example is in the courtroom, where the court decides among competing claims, be they civil or criminal. These claims often involve competing narratives of the facts at hand. As cognitive psychologists, we know that evidence, whether from eyewitness testimony or forensic tests, is not definitive. Instead, it is error prone in as much as there is some probability that the witness is wrong or the test was mishandled. In this sense, we may think of the environment in which courts operate as noisy. The court needs to adjudicate the competing claims—reach decisions—in this noisy environment. We seek to leverage our expertise as cognitive psychologists and apply it to a specific legal case—what many believe to be the wrongful conviction of Gary Leiterman. We first published this analysis in *Jurimetrics* (Wixted, Christenfeld, & Rouder, 2018), and describe it here with expanded commentary as an example to cognitive psychologists of how their expertise may be applied in consequential real-world settings.

Cognitive psychologists are experts in describing how people *should* make decisions in noisy environments, which is the *normative* or ideal-observer approach. Normative approaches include expected utility theory (Luce, 1959), game theory (Neumann & Morgenstern, 1945), the theory of signal detection (Green & Swets 1966), random walk theory (Edwards, 1965; Stone, 1960), Bayesian decision-making (Edwards, Lindman, & Savage 1963; Savage 1951) and the like. These theories assume that people can represent the probability and utility of events, and given these inputs, are ideal at combining and evaluating these probabilities and utilities to reach the most useful decision.

Additionally, cognitive psychologists are experts in *descriptive* decision-making, which is how people actually make decisions. Documented are all sorts of deviations from ideal behavior including the ignorance of base rates (Kahneman & Tversky 1973), the inability to combine information (Tversky & Kahneman 1974), the reliance

✉ Jeffrey N. Rouder
jrouder@uci.edu

[1] University of California, Irvine, CA, USA

[2] University of California, La Jolla, San Diego, USA

on heuristics (Gigerenzer & Goldstein, 1996), and the malleability of memory (Estes 1997; Loftus 2005).

It is this latter expertise, our knowledge of descriptive decision-making, that has been the most relevant to the legal profession. Examples include work on the reliability of witnesses in recalling events and correctly picking out perpetrators in a police line up (e.g., Loftus 1979; Steblay, Dysart, Fulero, & Lindsay 2003). In this paper, however, we leverage our cognitive psychology knowledge of normative decision-making to understand a very particular case—the conviction of Gary Leiterman for murder. In our view, a proper normative analysis, motivated by ordinary issues in cognitive psychology, reveals that it is highly likely that Leiterman was wrongly convicted.

The Leiterman case

Gary Leiterman was convicted in July, 2005, for the 1969 murder of Jane Mixer. He currently has served over 12 years of a life sentence in a Michigan state prison. The story starts in 1969 when first-year University of Michigan law student Jane Mixer was brutally strangled and shot twice in the head. Although the state police performed a thorough analysis of the crime scene, the case went cold and the materials were placed in storage. By 2001, however, there had been a tectonic shift in the state-of-the-art of forensic science. The most important advance was the ability to amplify even small, trace amounts of genetic material so that unique DNA profiles may be uncovered. In line with this trend, materials from Mixer's cold case were brought out of storage in late 2001 and submitted to DNA analysis. DNA material from two unknown males was found and analyzed.

The outcome of DNA analysis is a *profile,* which is a string of numbers indicating values at a small number of sites. These profiles may be compared to other known profiles in a database. The two unknown profiles from Mixer's crime scene were submitted to Michigan's DNA database, which is part of the FBI's national DNA database system, *Combined DNA Index System* (CODIS). Two matches were found within the Michigan state database.

One was from John Ruelas, who had recently killed his mother in January, 2002. A drop of Ruelas' blood was found on Mixer's hand. Ruelas, however, was 4 years old at the time of Mixer's death in 1969. It seems impossible that a 4-year-old would strangle a grown woman, shoot her twice, and move the body some seven miles to where it was eventually found. And so Ruelas was ruled out as a suspect. The other match was from Gary Leiterman, and a drop of clear body fluid, either sweat or saliva, was found on Mixer's pantyhose. Leiterman was 26 years old at the time of Mixer's murder. He was prosecuted and then convicted largely on the strength of the DNA match.

The jury's verdict of guilty

The salient evidence against Leiterman was a DNA match. DNA, though advanced, is not error free. In fact, there are several sources of errors (Thompson, Taroni, & Aitken 2003). The best considered source is the *random match probability*, or the probability that a DNA match would happen by chance to a randomly selected person from the population. State police lab scientists estimated that the random match probability was 1 in 170 trillion for Caucasians. Although there is some controversy about the accuracy of random-match-probability calculations (Thompson, 2009), we take the prosecution's assertion that Leiterman's DNA was indeed on Mixer's sample, and we do not consider this type of error further.

However, the random match error is not the only possible error. The defense raised the possibility of a contamination event. In this scenario, Leiterman's and Ruelas' DNA were not deposited on the crime-scene material in 1969, but in the Michigan State Police Crime Laboratory in early 2002. The contamination event is only plausible if the crime-scene materials, Ruelas' sample, and Leiterman's sample are processed in the same lab at about the same time. Indeed they were! Leiterman was charged with forging a prescription in December of 2001 for Vicodin, a synthetic opioid to which he had become addicted while suffering from kidney stones. He completed a year-long drug rehabilitation program, and the felony charge was never further prosecuted. Nonetheless, under a then new Michigan state law, Leiterman provided a buccal swab (a swab of saliva taken from the inside of the cheek) from which his DNA profile was constructed. Ruelas' had yet to be convicted in early 2002, but the materials from his crime scene, where he murdered his mother, were processed in the same lab during that time period. DNA processing takes place over an extended period of about 6 months or longer, and materials from the Mixer crime scene, Ruelas crime scene, and Leiterman's swab all were processed within a few months of each other—between November 2001 and March 2002—in the Michigan State Lab.

The case critically hinges on whether one believes that a contamination error is plausible. For if contamination is ruled out, then Leiterman's DNA must have been deposited at the crime scene, and, by assumption, he was the murderer. The prosecutors and lab scientists stated that contamination could not have occurred because: 1. the lab followed all federal standards; and 2. the analyses were performed on different days in different rooms. With this

testimony from scientists about DNA, the jury convicted Leiterman. They did so even though there is no plausible explanation of Ruelas' blood on Mixer's crime scene other than contamination.

Although contamination seems plausible, its plausibility has never been quantified. Several writers and experts have suspected it, and it was argued at the trial and on appeal. Unfortunately, during both the trial and the appeal, contamination rates were not known. As a result, the argument reduced to a conflict over assertions. The defense team asserted a possibility of contamination and used Ruelas' profile as evidence. The prosecution team, in contrast, asserted that contamination with Leiterman was an impossibility, and they put lab scientists on the stand to attest to why the procedures used in this case ruled out contamination. Without any other analysis, the jury was asked to choose between a confident prosecution using procedures, numbers, DNA, and science, and a defense team who said it was possible that there was contamination.

Since Leiterman's conviction, the literature on DNA analysis has grown, and there are now estimates of contamination rates. Contamination can be detected in cold cases by including a sample that is known to have no DNA material, and this sample is called a *negative control*. If the target sample is contaminated, presumably the negative control will be as well. In these negative controls, a DNA profile, clearly from contamination, is recovered in 1 in 1500 analyses (Kloosterman, Sjerps, & Quak 2014). If the target sample is contaminated, but the negative control is not, then the contamination is an *undetected* event. Undetected contamination is known to occur, much as it did with Ruelas' sample (which passed the negative-control test). We used the negative-control contamination rate as our estimate of the *undetected* contamination rate because it is the best-known estimate.

The Kloosterman et al. contamination rate of 1 in 1500 can be restated as the lab being 99.93% accurate. The question is whether this high success rate would be sufficient to incriminate Gary Leiterman. At first glance, the 99.93% success rate seems overwhelming, and one might naively think the chances of guilt are 99.93%, but as cognitive psychologists, we know the importance of base rates for rare events like murder and contamination, and we also know how easily these base rates are neglected in practice. Here, we calculate the probability that Leiterman's DNA was deposited on Mixer's clothing at the crime scene in 1969 vs. that it was deposited mistakenly in the lab in 2002. To foreshadow, with reasonable, transparent assumptions based on the latest forensic science, we find there is a 97% chance that Leiterman's DNA was deposited mistakenly in the lab in 2002 as an instance of contamination. To the extent that our assumptions are reasonable, the conviction and continued incarceration of Gary Leiterman is a gross miscarriage of justice.

Calculations of the probability of guilt

We set up the problem in the usual way by first considering two hypotheses.

H_1 : Leiterman killed Mixer,
H_2 : Leiterman was a victim of contamination.

Technically speaking, these are not the only two hypotheses, and there is the event that Leiterman killed Mixer and was the victim of contamination, but this possibility is so remote that we may safely ignore it. For our purposes, the above hypotheses are exhaustive and mutually exclusive, and their probabilities sum to 1.0.

As scientists, we know that the presence of Ruelas' blood on Mixer's clothes implies with exceedingly high probability that there was at least one contamination event in the analysis of Mixer's crime-scene materials. In turn, this fact implies that lab scientists were incorrect in their assertion that there was no possibility of contamination in the lab. Unfortunately, the jury did not consider the presence of Ruelas' blood as consequential in assessing the plausibility that Leiterman's DNA was deposited on Mixer's sample in the lab. Although this consideration by the jury defies logic, we use it as a starting point, as it cannot be relitigated. We calculate the probabilities of DNA deposition at the crime scene in 1969 vs. that in the lab in 2002 without any consideration of Ruelas. These calculations are, therefore, an upper bound on the probability of guilt. Any consideration of Ruelas would lower the probability because it increases the possibility of contamination over the baseline value from the literature.

Our main goal is to compute the probability of guilt conditional on the DNA test results. This probability is also called the posterior probability of guilt, and it is denoted by $P(H_1|D)$, where D denotes the test results. The key step in the calculation is to use Bayes' rule, itself a form of the Law of Conditional Probability. The form of Bayes' rule we find most helpful here is stated in terms of odds. In this case, the posterior odds of guilt are $\frac{P(H_1|D)}{1-P(H_1|D)}$. Because the two hypotheses are mutually exclusive and exhaustive, we note that the complement of guilt is Hypothesis 2, that is

$$\frac{P(H_1|D)}{1-P(H_1|D)} = \frac{P(H_1|D)}{P(H_2|D)}.$$

Hence, the posterior odds of guilt are the posterior ratio, $P(H_1|D)/P(H_2|D)$. This ratio is computed from Bayes rule as:

$$\frac{P(H_1|D)}{P(H_2|D)} = \frac{P(D|H_1)}{P(D|H_2)} \times \frac{P(H_1)}{P(H_2)}.$$

The posterior odds are the product of two terms. The one on the left is the probability of obtaining the DNA results under competing hypotheses; the one on the right is the prior odds of the hypotheses before observing the data. In the following, we define the events that comprise the data and use the literature to specify or estimate the relevant probabilities needed for calculations.

Base rates and prior probabilities

To calculate a posterior guilt ratio, we need to start with prior odds. We use the following arguments to set our values:

- $P(H_1)$. Hypothesis 1 is that Leiterman killed Mixer. Without any further information, we assume that all males in the greater Detroit area in 1969 between the ages of 15 and 60 are equally plausible as the murderer. There were 4 million people living in the area, and we assumed that 1/4 of them, 1 million, were male and capable of the murder. Given this denominator, the a priori probability that Leiterman is the murderer is $P(H_1) = 10^{-6}$.
- $P(H_2)$. Hypothesis 2 is that Leiterman is the victim of contamination in the lab in 2002. There are two choices of development here, and they are largely semantic. One is that we can take the hypothesis as the conjunction of two events: (i) that there was contamination with Mixer's crime-scene samples; and (ii) that contamination was with Leiterman's sample rather than with any of the other samples processed contemporaneously with Mixer's. The other choice is that we can break up the conjunction into separate events. We made this second choice and considered the probability of contamination as a feature of the DNA test rather than as a base rate. We take as the base rate the probability that Leiterman was the victim of contamination given that it occurred. This makes the base rate analogous with that under Hypothesis 1 where the base rate was the probability Leiterman committed the murder given that the Mixer murder occurred. We will incorporate the contamination rate subsequently as part of the data, discussed below. Hence, with this setup, calculating this base rate requires an estimate of how many samples were processed contemporaneously with Mixer's. In 2002, the Michigan State Police Crime Lab processed about 10,000 samples, with about 5000 processed during the time the Mixer evidence was in the lab (Jen 2003). Hence, we estimate that there were approximately 5000 samples analyzed contemporaneously with Mixer. The base rate is thus $P(H_2) = 1/5000 = 2 \times 10^{-4}$.

DNA test results as data

The DNA test results are not a single number but a collection of events. We take the following four events to comprise the data:

- **A match to Leiterman**: There was a definite match between Leiterman's known profile and the DNA found on Mixer's pantyhose. We denote this match event as E_m.
- **Saliva**: The matching DNA on Mixer's pantyhose was consistent with saliva; it was not consistent with blood or semen. We denote this event as E_s.
- **Exclusivity**: The DNA found on Mixer's pantyhose was exclusively from Leiterman; none from Mixer herself was found. We denote this event as E_e.
- **Contemporaneous analysis**: Mixer and Leiterman's DNA were analyzed in the same lab at the same time. We denote this event as E_c.

We are now in a position to calculate posterior odds conditional on these events. We treat these events as statistically independent:

$$P(D|H_1) = P(E_m|H_1) \times P(E_s|H_1) \times P(E_e|H_1) \times P(E_c|H_1),$$

$$P(D|H_2) = P(E_m|H_2) \times P(E_s|H_2) \times P(E_e|H_2) \times P(E_c|H_2).$$

We used the following values in calculation:

- $P(E_m|H_1)$: What is the probability of a DNA match under H_1? A DNA match is not guaranteed even if Leiterman murdered Mixer. In fact, in about half the crimes, a DNA profile is not recoverable because the assailant did not leave enough material for identification (Roman, Reid, Chalfin, & Knight 2009). Moreover, even if the assailant did leave material, there is degradation of the material over the 33 years between the murder in 1969 and the analysis in 2002. The best estimate that we can find is that the probability of degradation is about one-half (Wiser, 2011), and combining these two yields a value of 1/4. Hence, we set $P(E_m|H_1) = .25$.
- $P(E_m|H_2)$: What is the probability of a DNA match under H_2, the contamination hypothesis? Here we ask what is the probability of a match through contamination. The best estimate we can document is 1 in 1500 (Kloosterman et al., 2014). Further discussion is provided in Wixted et al. (2018). Hence, we set $P(E_m|H_2) = 6.67 \times 10^{-4}$.
- $P(E_s|H_1)$: What is the probability that the DNA sample is from saliva under H_1? In about 1/2 of crime scenes

can the perpetrator's saliva be recovered (Cross et al., 2014). Hence, we set $P(E_s|H_1) = .5$.

- $P(E_s|H_2)$: The probability that the DNA sample is from saliva is 1.0 under H_2 as it is known that Leiterman's sample was from a buccal swab.
- $P(E_e|H_1)$: Leiterman's DNA was found on Mixer's crime scene materials but Mixer's was not. This is not a common occurrence. In simulation studies where the victim's clothing is in contact with a perpetrator, about 85% of the time, the victim's DNA is found in greater concentrations on that clothing than the perpetrator's DNA (Breathnach, Williams, McKenna, & Moore 2016). Hence, we set $P(E_e|H_1) = .15$.
- $P(E_e|H_2)$: Under the contamination scenario, it makes sense that Leiterman's concentration is larger than Mixer's. Leiterman's DNA comes from a purposeful swab designed to maximize the likelihood of extracting a profile. Based on the extant literature, Wixted et al. concluded that a lower bound on $P(E_e|H_2) = .575$, a value we use here.
- $P(E_c|H_1)$: Under H_1, there is no requirement that Leiterman's and Mixer's DNA are analyzed contemporaneously. In fact, it seems alarmingly coincidental under this hypothesis that Leiterman and Mixer, two strangers, met in 1969 and then had their DNA sequenced contemporaneously 33 years later. How probable is it that they were sequenced contemporaneously? This probability is hard to compute because there is what is known as *the reference class problem* (Hájek, 2007). We estimate probabilities by asking the frequency of events from a reference class. For example, in flipping coins, we might estimate a probability by asking how many successes from the reference class that is all flips. Sometimes, however, the reference class is not obvious, and it is not obvious here. We know how many samples were sequenced contemporaneously with Mixer's, and that number is 5000; but what is it out of, that is, what is the appropriate reference class? One reasonable reference class is *all the people who were sequenced in the same lab before Mixer's*; another one is *all the people sequenced in all U.S. labs before Mixer's*; and a third is *all the people sequenced to the present time in the Michigan lab or in all U.S. labs*. Our approach is to use the smallest reasonable reference class as to maximize the probability. Such a maximization is in favor of the prosecution as it increases the calculated value of probability of guilt. To this end, we consider the reference class all people whose DNA was analyzed in the Michigan State Lab before or contemporaneously with Mixer's analysis. The number of people in this smallest reasonable reference class is 42,000 (Wixted et al., 2018), and, consequently, $P(E_c|H_1) = 5000/42000 = .12$
- $P(E_c|H_2)$: Under contamination, the probability of contemporaneous analysis must be 1.0.

Calculations

Combining the above equations yields:

$$\frac{P(H_1|D)}{1-P(H_1|D)} = \frac{P(E_m|H_1)\times P(E_s|H_1)\times P(E_e|H_1)\times P(E_c|H_1)}{P(E_m|H_2)\times P(E_s|H_2)\times P(E_e|H_2)\times P(E_c|H_2)} \times \frac{P(H_1)}{P(H_2)}.$$

Substituting the above values yields odds of guilt:

$$\frac{P(H_1|D)}{1-P(H_1|D)} = .0291.$$

Restated, the odds are about 34 to 1 in favor of innocence.

These odds of guilt may be converted to a probability by the equality $p = o/(o+1)$, where p is the probability and o is the odds. In this case, the probability of guilt is about .0283, or a tad less than 3%.

Revisiting the contamination rate

The above analysis reflects critically the contamination rate. If the rate is identically zero, as the Michigan State lab scientists testified, then, indeed, the probability of guilt is identically 1.0. The key to forming a rational opinion about Leiterman is to use a good-faith, informed, reasonable estimate of the contamination rate. Ours comes from Kloosterman et al. (2014), who studied the rate of *detected* contamination events. This rate is surprisingly high, and it indicates that on occasion contamination occurs. How often it is detected when it occurs remains unknown.

What if the contamination rate is much smaller than the value used here? Fig. 1 shows the probability of guilt as

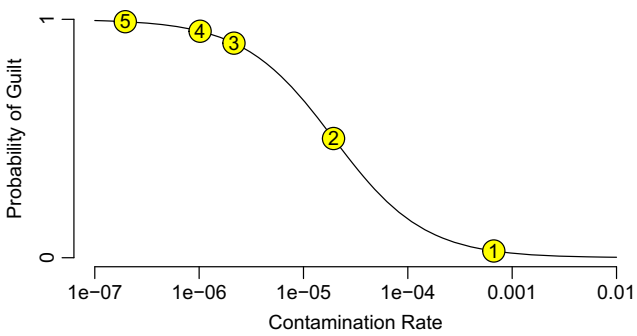

Fig. 1 Probability of guilt as a function of the contamination rate. The Point labeled "1" is the value used, 1 in 1500, from Kloosterman et al. (2014). The other four points correspond to contamination rates needed to support probabilities of 50, 90, 95, and 99%

Table 1 Relationship between contamination rate and posterior probability of guilt

Point	Contamination rate	Pr(Guilt|Data)	Overestimation factor
1	1 in 1500	3%	1
2	1 in 50,000	50%	34
3	1 in 500,000	90%	309
4	1 in 1 million	95%	653
5	1 in 5 million	99%	3400

a function of contamination rate. The contamination rate we used, 1 in 1500, is indicated with Point 1, and this rate corresponds in a 3% chance of guilt. Contamination rates that correspond to guilt chances of 50, 90, 95, and 99% are shown in Points 2 through 5. The values used here are provided in Table 1, and the results are eye opening. For example, if we decided that 95% was a good burden for the prosecution, then, the contamination rate would have to be about 1 in a million to support this burden. Such a rate is 650 times smaller than the one we estimated from the literature. Restated, even if we had overestimated the contamination rate by a factor of 650, the posterior probability of guilt would still not exceed 95%. We list these possible overestimation factors in Table 1 for each point.

Incorporating uncertainty

The values in the preceding analyses represent our best point estimates of probabilities from our reading of available sources. It is prudent to study what may happen if uncertainty is added to these estimates. Fortunately, the Bayesian updating machinery is well suited for this task. We place a prior distribution on each input value and then, following Bayes rule, compute a posterior distribution of guilt. In the above setup, we used ten different values to calculate a point estimate of the probability of guilt. Here we start with distributions instead of points as shown in Fig. 2. Take, for example, the top-left panel, which shows the distribution of the match event, E_1. Previously, we used a value of $P(E_m|H_1) = .25$ for the guilt hypothesis. Now we place a distribution on this probability. Let $\Omega_{m,1}$ be the *log-odds* or logit transform. The subscripts refer to the event, in this case m for a match, and the hypothesis, in this case, 1 for H_1:

$$\Omega_{m,1} = \log\left(\frac{P(E_m|H_1)}{1 - P(E_m|H_1)}\right).$$

The distribution we place is

$$\Omega_{m,1} \sim \text{Normal}\left(a_{m,1}, b_{m,1}^2\right),$$

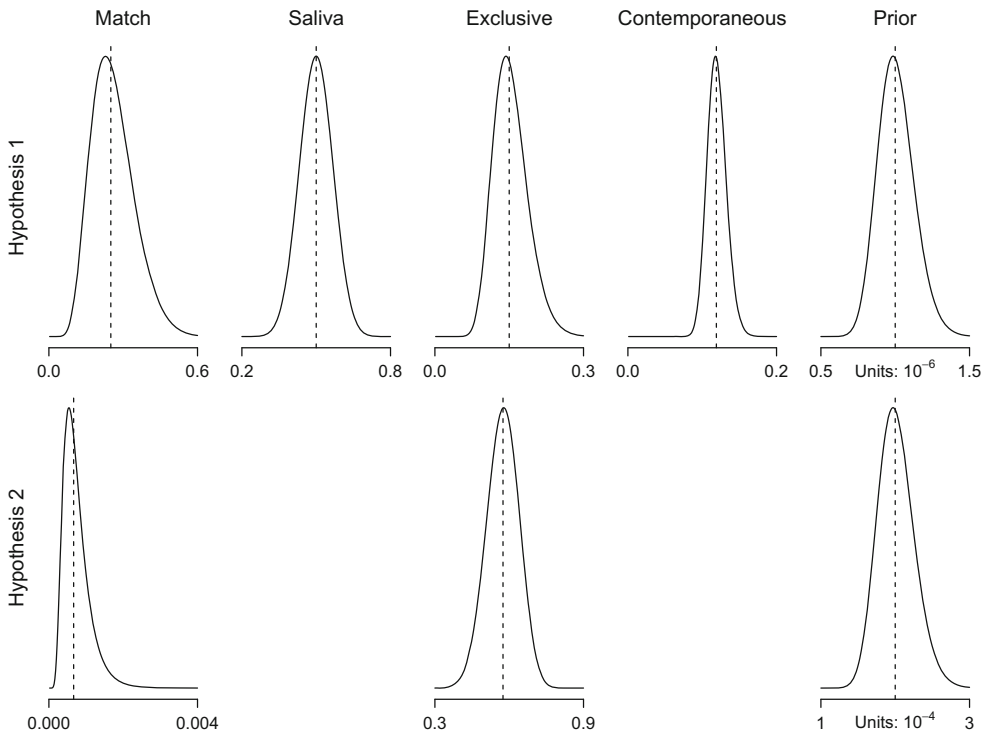

Fig. 2 Distributions on the input values in computing the probability of guilt

where $a_{m,1}$ and $b^2_{m,1}$ are the mean and variance, and must be set judiciously. We used values of $a_{m,1} = \log(1/3)$, which is the log-odds transform of $P(E_m|H_1) = .25$, and $b_{m,1} = 0.46$. The resulting distribution of $P(E_m|H_1)$ is shown in Fig. 2, and the solid line is at the previous point value of .25. Placing normals on log-odds is natural for probabilities, and we do it throughout. The means, a, are set in accordance with the previous point values on probability. The standard deviations, b, are set by considering how much variability there might be in the odds. For example, consider $P(E_m|H_1) = .25$, which corresponds to 1-to-3 odds that there was a DNA match under H_1. We thought the odds could reasonably vary by a factor of 2, say from 1 to 6 to 1 to 1.5. We set b, the standard deviation on the log-odds as $2\log(f)/3$, where f is the factor by which odds might vary. For the match event under Hypothesis 1, we set the factor $f = 2$ (odds from 1 to 6 to 1 to 1.5), and, consequently $b = .46$. We followed the same procedure for all inputs, and the variability in the probabilities in Fig. 2 reflect our best assessment of the range of plausible probabilities.

With these priors on probabilities defined, the last step is to calculate the posterior. The easiest approach is simply to sample from the priors and multiply! We set up iterations where a sample is gathered from each of the distributions in Fig. 2. Then, with these values, we use the above formulas to compute a sample of the posterior probability of guilt. The process is repeated for 100,000 iterations. The resulting histogram is shown in Fig. 3. As can be seen, even with a fair degree of assumed variability, the posterior probability of contamination remains quite high, and the posterior probability of guilt remains quite low.

Discussion

The trial and appeal of Gary Leiterman amounted to adjudicating between two verbal claims. The defense argued that the match could be due to contamination in the lab, the prosecution stated unequivocally that contamination was impossible, and they had several lab scientists back them up in this assertion. Neither the defense nor the prosecution could attempt a good-faith effort at quantifying the possibilities, and without this quantification, the jury sided with the prosecutor and lab scientists. We suspect that if any reasonable quantification had been offered, so long as estimated odds were not off by several orders of magnitude, Leiterman's prosecution could not meet a beyond-a-reasonable-doubt threshold. In our opinion, if the information from the forensic DNA literature that we have now had been available in 2004, Leiterman could have been exonerated by mathematics.

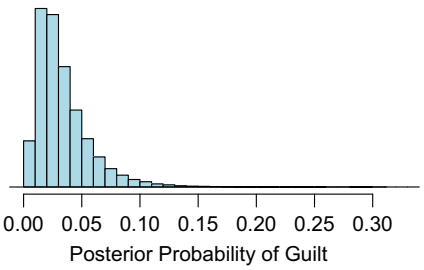

Fig. 3 Posterior probability distribution of guilt when inputs are treated as distributions rather than points

The use of probabilities has had a mixed past in law, and the experience is informative. Perhaps the modern story starts with Savage's (1951) Bayesian approach to decision-making. Finkelstein and Fairley (1970) adopted Savage's approach for the legal system. Here, the jury members would be instructed exactly how to update their prior beliefs about guilt and innocence in a formal fashion. They would be told the relevant Bayes factors in exactly the way Rouder and Morey (2012) recommend that Bayes factors should be reported by investigators to the wider scientific psychology community. And Finkelstein and Fairley (1970) is one of several papers of the era that recommended adopting formal decision-making in the jury process. The endeavor became known pejoratively as *trial-by-mathematics* by Tribe (1971), who provides elegant critiques of the approach. The role of probability remains controversial, timely, and topical in American jurisprudence. Here, we briefly explore some of the critical issues as they relate to our analysis.

We focused on the analysis of the DNA data, and in doing so, we identified four events: the match, the type of bodily fluid the DNA came from, the relative concentrations, and the time and place of analysis. We picked these four because they were related to the DNA test and the probabilities of their occurrences have been the subject of recent research. However, there are other events related to the DNA tests that we may have considered. In favor of Leiterman are two events that we did not include in calculation: One is the presence of Ruelas' match on Mixer's sample, which must be through contamination. Another is that the lab was unable to find DNA on the substrate from Leiterman's buccal-swab reference sample itself when it was subsequently processed. Theodore Kessis, a DNA expert for the defense during the appeal notes this event. He states that it is possible that Leiterman's DNA was missing because it had previously mistakenly been transferred to Mixer's crime-scene evidence. However, there are events that are in favor of the prosecution. In this case, the Michigan State Lab uses two different genetic sequencers for processing crime-scene materials and for the processing of felons for CODIS. If we accept that lab personnel *always* used different sequencers, the chance of contamination is perhaps lowered, a possibility we did not incorporate.

The problem of conditioning on appropriate events goes hand-in-hand with the problem of assigning probabilities to these events. One of the most difficult is the problem of *reference classes* (Hájek 2007). We illustrated it earlier with calculating the probability that Mixer's and Leiterman's samples would be processed contemporaneously. This reference-class problem is ubiquitous throughout the analysis. Take for example the a priori probability that Leiterman murdered Mixer, which we set to 1 in a million. We defined a reference class of males who lived in the Detroit area in 1969 who were in a certain age range, but maybe we should have stratified by other factors, say marital status, as many though not all people who murder unrelated strangers are not married[1] We may have also narrowed or widened our radius in defining the reference class.

Although the tasks of picking events and assigning probabilities to them remain challenging, we should not let the challenge prevent us from providing a quantitative analysis. It is our view that it was the lack of analysis, the inability to bring numbers to bear to the possibility of contamination, that contributed to the fate of Gary Leiterman.

Analysis in our view will remain subjective, but not hopelessly so. The key for us is to embrace the subjectivity with transparency, openness, and humility. We endeavor to state our starting positions, why we make them, and how firm we believe them to be. Then we use Bayesian analysis, which is a direct reflection of the laws of probability, to draw the appropriate inferences. In our experience, most American statisticians are not prepared to engage in the degree of subjectivity that we embrace. The majority of Bayesian statisticians, in fact, are not philosophically committed to a subjective view of analysis. Instead, they use Bayesian methods for their computational power and convenience (Senn, 2011). Cognitive psychologists, in contrast, *know* how much art there is in the state-of-the-art in psychological science. The subjectivity in defining theories and linking them to the relevant data pales in comparison to the problem at hand in computing the odds of Leiterman's guilt. Perhaps this ability to work productively with subjectivity is the cognitive psychologist's greatest advantage. After living our intellectual lives in an exceedingly difficult and messy intellectual world, we are well prepared in temperament and in skills to handle the challenges of quantifying possibilities in the legal world.

[1]It is surprisingly hard to find comprehensive studies of covariates associated with people who kill unrelated strangers. There are many anecdotal accounts, including those that appear in the popular press. We have yet to identify a peer-reviewed publication that provides an explicit inclusion criteria and a systematic survey of even the most basic covariates such as age, marital status, race, IQ, SES, etc.

References

Breathnach, M., Williams, L., McKenna, L., & Moore, E. (2016). Probability of detection of DNA deposited by habitual wearer and/or the second individual who touched the garment. *Forensic Science International: Genetics*, *20*, 53–60.

Cross, T. P., Alderman, M., Wagner, A., Sampson, L., Peters, B., Spencer, M., & Loudenberry, K. (2014). Forensic evidence and criminal justice outcomes in a statewide sample of sexual assault cases. National Criminal Justice Reference Service.

Edwards, W. (1965). Optimal strategies for seeking information: Models for statistics, choice reaction times, and human information processing. *Journal of Mathematical Psychology*, *2*, 312–329.

Edwards, W., Lindman, H., & Savage, L. J. (1963). Bayesian statistical inference for psychological research. *Psychological Review*, *70*, 193–242. Retrieved from https://doi.org/10.1037/h0044139

Estes, W. (1997). Processes of memory loss, recovery, and distortion. *Psychological Review*, *104*, 148–169.

Finkelstein, M. O., & Fairley, W. B. (1970). A Bayesian approach to identification evidence. *Harvard Law Review*, *83*, 489–517. Retrieved from https://www.jstor.org/stable/1339656?seq=1#page_scan_tab_contents

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, *103*, 650–669.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Hájek, A. (2007). The reference class problem is your problem too. *Synthese*, *156*, 563–585.

Jen, K. I. (2003). Michigan's forensic DNA database. Retrieved from www.house.mi.gov/hfa/Archives/PDF/dna.pdf

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, *80*(4), 237.

Kloosterman, A., Sjerps, M., & Quak, A. (2014). Error rates in forensic DNA analysis: Definition, numbers, impact and communication. *Forensic Science International: Genetics*, *12*, 77–85.

Loftus, E. F. (1979). The malleability of human memory: Information introduced after we view an incident can transform memory. *American Scientist*, *67*(3), 312–320.

Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory*, *12*(4), 361–366.

Luce, R. D. (1959). *Individual choice behavior*. New York: Wiley.

Neumann, J., & Morgenstern, O. (1945). *Theory of games and economic behavior*. Princeton: Princeton University Press.

Roman, J. K., Reid, S. E., Chalfin, A. J., & Knight, C. R. (2009). The DNA field experiment: A randomized trial of the cost-effectiveness of using DNA to solve property crimes. *Journal of Experimental Criminology*, *5*(4), 345.

Rouder, J. N., & Morey, R. D. (2012). Default Bayes factors for model selection in regression. *Multivariate Behavioral Research*, *47*, 877–903. Retrieved from https://doi.org/10.1080/00273171.2012.734737

Savage, L. J. (1951). The theory of statistical decision. *Journal of the American Statistical Association*, *46*(253), 55–67.

Senn, S. (2011). You may believe you are a Bayesian but you are probably wrong. *Rationality, Markets and Morals*, *2*, 48–66.

Steblay, N., Dysart, J., Fulero, S., & Lindsay, R. (2003). Eyewitness accuracy rates in police showup and lineup presentations: A meta-analytic comparison. *Law and Human Behavior*, *27*, 523–540.

Stone, M. (1960). Models for choice-reaction time. *Psychometrika*, *25*, 251–260.

Thompson, W. C. (2009). Painting the target around the matching profile: The Texas sharpshooter fallacy in forensic DNA interpretation. *Law, Probability & Risk*, *8*(3), 257–276.

Thompson, W. C., Taroni, F., & Aitken, C. G. (2003). How the probability of a false positive affects the value of DNA evidence. *Journal of Forensic Science*, *48*(1), 1–8.

Tribe, L. H. (1971). Trial by mathematics: Precision and ritual in the legal process. *Harvard Law Review*, *84*, 1329–1393.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131. Retrieved from https://doi.org/10.1126/science.185.4157.1124

Wiser, M. (2011). Time, money running out for state's cold cases. Globe Gazette. Retrieved from http://globegazette.com/news/time-money-running-out-for-state-s-cold-cases/article_d99154ce-9fa9-11e0-a3e3-001cc4c002e0.html

Wixted, J. T., Christenfeld, N. J. S., & Rouder, J. N. (2018). A Bayesian statistical analysis of the DNA contamination scenario. *Jurimetrics*, *58*, 211–242.