(Published in: *Medicine, Science and the Law* (2004) 44 189—192)

# Inference with legal evidence: common sense is necessary, but not sufficient

ROBERT A.J. MATTHEWS

*Visiting Reader, Department of Information Engineering, Aston University*

*Correspondence: R.A.J. Matthews, Department of Information Engineering, Aston University, Birmingham, B4 7ET*

*Tel: 0121 333 4631 E-mail rajm@physics.org*

**ABSTRACT**

Recent cases have highlighted the issue of faulty probabilistic reasoning by expert witnesses in courts of laws. While concern about potential miscarriages of justice is clearly well-placed, the consequences of such faulty reasoning do not seem to be fully appreciated. These are often counterintuitive, as we show with two examples: the Interrogator's Fallacy and the Prosecutor's Fallacy. Both demonstrate the danger of relying solely on 'common sense' when drawing inferences from legal evidence.

## INTRODUCTION

The use of probabilistic inference in legal matters has a long but chequered history. While many of the founders of probability theory were convinced of its value, this sentiment has not been widely shared by the legal profession. As Franklin (2001) puts it in his fascinating study of the early history of probability: "By 1700 law had served its purpose for the mathematical theory of probability. The service was never returned. Legal probability has continued to exist, and it is accepted in legal theory that such notions as proof beyond reasonable doubt involve probability. But all attempts to quantify the concept have been resisted" [p 365].

If this reluctance stems from concern about the ability of otherwise intelligent people to handle probabilistic arguments, it is well-founded. Textbook examples date back at least as far back as the 1960s and the case of *People v. Collins*, in which an expert witness attempted to estimate the probability of guilt simply by multiplying together frequencies of certain traits (Koehler 1997). There have been many examples since, notably *R v. Clark,* in which the accused stood trial for the murder of her two infant sons. At the trial, the possibility that both her sons had died from Sudden Infant Death Syndrome ('cot death') was considered, and an expert witness argued that the probability of a single SIDS death in a family was 1 in 8,543. When asked to estimate the probability of a two such deaths in the same family, the expert witness simply multiplied the two probabilities together – apparently oblivious of the dubious assumption of independence this implies - obtaining impressive-sounding odds of 1 in 73 million against. The calculation was questioned by both another expert witness and by the trial judge in the summing-up. Significantly, however, it was not the egregious abuse of probabilistic reasoning that formed the focus of the judge's comments, but his view that "We do not convict people in these courts on statistics. It would be a terrible day if that were so". This sentiment was shared by the Appeal Court, which overturned Clark's conviction on other grounds.

Whatever the reluctance to accept probabilistic argument in court, the emergence of DNA-based evidence has made it a common feature in many trials. The focus of such evidence is the so-called match probability, that is, the probability of obtaining as good a DNA match as that found with the defendant's DNA, assuming the defendant to be innocent. This convoluted definition is all too easily confused with the probability of the defendant actually being innocent. Given that the match probability is also typically very low ($10^{-9}$ is not uncommon), there is a clear danger of the probability of innocence being taken to be similarly small, and a guilty verdict returned.

This faulty train of reasoning is known as the Prosecutor's Fallacy, a term first coined by Thompson and Schumann (1987). Concern that it has led juries to reach inappropriate guilty verdicts has led to a number of appeals (see, eg Dawid 2002). It has also prompted research into ways of presenting DNA evidence which are less prone to misinterpretation, such as the use of frequency-based formats (Hoffrage *et al.* 2000) and computer-based assistance (Fenton & Neil 2000).

Despite this, there remains considerable scepticism in some quarters of the legal profession about the inherent dangers of this inferential fallacy. In a 1996 ruling on a case involving the Prosecutor's Fallacy, the English Appeal Court declared that the role of a jury is to "evaluate evidence and reach a conclusion not by means of a formula, mathematical or otherwise, but by the joint application of their individual common sense and knowledge of the world to the evidence before them" (quoted in Dawid 2002).

Such confidence in the powers of common sense is hard to square with the results of recent research showing that the way DNA evidence is presented has a dramatic impact on the frequency of inappropriate guilty verdicts (Hoffrage *et al* 2000). It is further undermined by the fact that some common forms of evidence have counter-intuitive properties that have yet to be widely recognised even by those familiar with the perverseness of probabilistic reasoning. It is to these that we now turn.

#### THE INTERROGATOR'S FALLACY

Whatever the qualms about the use of probabilistic reasoning in court, the laws of probability have been shown mathematically to form the bedrock of any coherent and rational assessment of evidence (see, eg Howson & Urbach 1993). Assessing the impact of evidence on belief in a specific hypothesis (eg the guilt of the defendant) is then a matter of applying Bayes's Theorem, according to which evidence $E$ alters the odds on the guilt $G$ of the defendant from its original ("prior") value of $\text{Odds}(G)$ to a new value $\text{Odds}(G \mid E)$, given by

$$Odds(G \mid E) = LR \times Odds(G) \quad (1)$$

where LR is the so-called Likelihood Ratio of probabilities $Pr(E \mid G)/Pr(E \mid {\sim}G)$, with ~ denoting negation (eg ~G = not-guilty = innocent) and "|" is read "given". Evidence $E$ which increases the odds of guilt will lead to $Odds(G \mid E) > Odds(G)$, which by (1) implies that the LR must exceed unity, so that

$$Pr(E \mid G) > Pr(E \mid {\sim}G) \quad (2)$$

In other words, evidence $E$ increases the probability of guilt if and only if $E$ is more plausibly associated with guilt than with innocence. This is entirely in accord with common sense; even so, there are circumstances in which it can lead to counterintuitive consequences. A case in point is confessional evidence, which despite a long and controversial history continues to play a significant role in an estimated 20 per cent of cases (Gudjonsson 1992). It has also been a central feature of some of the most notorious miscarriages of justice of recent times, notably the Guildford Four, Birmingham Six and Judith Ward cases of the mid-1970s.

The prevalence of terrorism-related cases among these miscarriages of justice is striking; it is also symptomatic of a counterintuitive consequence of (2) to which confessional evidence is especially prone. According to (2), confessions only contribute weight of evidence against a defendant if they are more plausibly associated with guilt than innocence. While this condition may seem entirely uncontroversial, it is likely to be breached in terrorism cases, as terrorists are specifically trained and motivated to resist interrogation (see e.g. Coogan 1993). As such, those guilty of terrorism crimes are *less* likely to confess than those unconnected with such crimes, and (2) no longer holds. Indeed, obtaining a confession in terrorism cases plausibly constitutes a reversal of the inequality (2), and thus provides evidence of the defendant's *innocence*.

This failure to consider whether (2) holds true when considering evidence of guilt has been called the Interrogator's Fallacy (Matthews 1995). While it can have counterintuitive consequences, the cause of this inferential fallacy is clear enough, as is the remedy: one must always ensure evidence does not violate the principle that beliefs only gain support from evidence more plausibly associated with the belief being true than false.

Our second example of the dangers of relying on common sense is somewhat more subtle, and centres on the Prosecutor's Fallacy.

## BASIS OF THE PROSECUTOR'S FALLACY

In the argot of probability theory, the Prosecutor's Fallacy arises through so-called transposition of conditioning; that is, confusion between the two conditional probabilities $Pr(A \mid B)$ and $Pr(B \mid A)$. Such confusion is often easily detected: for example, most people can appreciate the huge difference between the probability that one has the flu, given symptoms of fever $Pr(flu \mid fever)$, and the (very high) probability of having a fever, given infection with influenza, $Pr(fever \mid flu)$. In the case of DNA evidence, however, the fallacy consists of confusing the match probability, $Pr(DNA\ match \mid innocence)$ with the probability of the defendant actually being innocent, given the DNA match evidence, *Pr(innocence | DNA match)*. As both these conditional probabilities are somewhat abstract, it is hardly surprising confusion arises.

In the considerable literature on the Prosecutor's Fallacy, there is a widespread assumption that – as its name implies – the fallacy always favours the prosecution. It does not appear to be widely recognised that the fallacy has a counterintuitive feature which allows it to favour the defence, as we now show.

## CONDITIONS FOR THE FALLACY TO BENEFIT THE DEFENCE

By Bayes's Theorem, we have, as before

$$Odds(G \mid E) = LR \times Odds(G) \quad (3)$$

For DNA evidence, the false-negative rate is sufficiently low that the numerator of the LR is almost exactly 1, while the denominator is the match probability $Pr(E \mid \sim G)$. As already noted, the essence of the Prosecutor's Fallacy lies in confusing $Pr(E \mid \sim G)$ with $Pr(\sim G \mid E)$, the probability of innocence given the DNA evidence. The effect of this confusion can, however, be counterintuitive. To see this, convert (3) into probabilities using $Pr(\ldots) = Odds(\ldots)/[1 + Odds(\ldots)]$, set $Pr(E \mid G) \approx 1$ and bring the two conditional probabilities involved in the Prosecutor's Fallacy together as a ratio:

$$Pr(E \mid \sim G) / Pr(\sim G \mid E) = Odds(G)/ Pr(G \mid E) \quad (4)$$

From (4) we see that the consequences of committing the Prosecutor's Fallacy depends on two factors: the prior odds of guilt *Odds(G),* and the probability of guilt in the light of the evidence, *Pr(G | E).* While the latter is unknown, its value lies in the unit interval; hence, rearranging (4) we find

$$Pr(E \mid \sim G) \geq Pr(\sim G \mid E) \times Odds(G) \quad (5)$$

Thus the match probability $Pr(E \mid \sim G)$ will *exceed* the true conditional probability of innocence, $Pr(\sim G \mid E)$ if $Odds(G) > 1$, ie if the prior probability of guilt is deemed to exceed 0.5. In other words, if jury members are already convinced of the defendant's guilt on a simple balance of probabilities and then commit the Prosecutor's Fallacy, they will *over*-estimate the probability of the defendant being innocent.

Despite its name, therefore, the Prosecutor's Fallacy is capable of supporting the case for the defence. This counterintuitive result has obvious relevance to appeals made on the grounds that the original jury committed the Prosecutor's Fallacy. Whether the fact that the fallacy can favour the defence, let alone the conditions under which it does so, would be recognised by anyone relying on common sense alone is open to question.

## CONCLUSION

Current concern about the potential for miscarriages of justice arising from faulty probabilistic reasoning is undoubtedly well-placed. Probability theory is replete with counterintuitive results, and recent court cases have demonstrated that egregious errors can be made even with basic concepts in probability. We have shown, however, that there are more subtle inferential traps awaiting the unwary. The Interrogator's Fallacy can turn an admission of guilt into impressive evidence of innocence, while the Prosecutor's Fallacy - to which DNA evidence is especially prone - can work in favour of the defence. Both fallacies demonstrate the inadequacy of common sense in assessing legal evidence. Expositions of the mathematical details underpinning such fallacies may not be appropriate for courts of law, but a greater awareness of their existence surely is.

## REFERENCES


Coogan T.P. (1993) *The IRA* London HarperCollins

Dawid, A.P., (2002) Bayes's Theorem and weighing evidence by juries *Proc Brit Acad* **113** 71-90

Fenton N.E. and Neil M., (2000) The Jury Observation Fallacy and the use of Bayesian Networks to present Probabilistic Legal Arguments. *Math Today* **36** 180-187.

Franklin J. (2001) *The Science of Conjecture* Baltimore, Johns Hopkins University Press

Gudjonsson G.H. (1992) *The Psychology of Interrogations, Confessions and Testimony* Chichester Wiley

Hoffrage U., Lindsey S., Hertwig R., and Gigerenzer G. (2000) Communicating statistical information. *Science* **290** 2261-2262.

Howson C. & Urbach P. (1993) *Scientific Reasoning: the Bayesian Approach* Chicago, Open Court Publishing

Koehler J.J. (1997). One in millions, billions, trillions: lessons from *People vs. Collins (1968)* for *People vs. Simpson (1995). J Legal Education* **47** 214-223.

Matthews R.A.J. (1995) The Interrogator's Fallacy *Bull Inst Maths Apps* **31** 3-5

Thompson W.C. and Schumann E.L. (1987) Interpretation of statistical evidence in criminal trials: the prosecutor's fallacy and the defense attorney's fallacy. *Law and Human Behaviour* **11** 167-187