Evidence, probability and relative plausibility

Colin Aitken (ID)
School of Mathematics, The University of Edinburgh, UK

Franco Taroni
School of Criminal Justice, The University of Lausanne, Switzerland

Silvia Bozza
Department of Economics, The Ca' Foscari University of Venice, Italy

Abstract
A comparison is made between probability and relative plausibility as approaches for the interpretation of evidence. It is argued that a probabilistic approach is capable of answering the criticisms of the proponents of relative plausibility. It is also shown that a probabilistic approach can answer the problem of overlapping where there is evidence that each side claims supports its theory of what happened.

Introduction

Over a period of 25 years and three editions, a seminal book *Statistics and the Evaluation of Evidence for Forensic Scientists* has described the underlying ideas and mapped the progress of these ideas in the area of the book title, an area which has come to be known as forensic statistics. The first edition of Statistics and the Evaluation of Evidence for Forensic Scientists was published in 1995 (Aitken, 1995). As the manuscript neared completion, there were two suggestions of the publisher that did not proceed to the published version of the book. The first related to the title where it was suggested to include the word 'interpretation' until it was pointed out that the book was about 'evaluation' not 'interpretation' and that the two words were not synonymous. The second suggestion was that the cover could include a picture of a fingerprint. It was pointed out in this case that there was very little on fingerprints in the book. There is an analogy of evidence evaluation with the scales of justice. Thus, it was that the book cover included scales of justice, images of which have appeared on the two subsequent editions.

Corresponding author:
Colin Aitken, School of Mathematics, The University of Edinburgh, King's Buildings, Peter Guthrie Tait Road, Edinburgh EH9 3FD, UK.
E-mail: C.G.G.Aitken@ed.ac.uk

There may have been little on interpretation in the first edition, a reflection on the relationship between statistics and forensic science at the time. In the second edition, in 2004, there were about 80 pages on interpretation and a discussion of transfer evidence, reflecting the work on propositions done in the late 1990s; see Cook et al. (1998a, 1998b). Now, 25 years after the appearance of the first edition, there are about 300 pages on interpretation within nearly 1000 pages of material, an increase which can only partially be explained by an increase in type size and margin sizes from the previous two editions. However, the title remains unchanged to ensure continuity of ideas from the beginning.

The primary concern of the first edition was with technical matters. DNA profiling was in its relative infancy, having been around less than 10 years. Its introduction helped the introduction of probabilistic reasoning to the courts. There were many teething problems, as the court transcripts of the time show. These were being overcome and there were thoughts as to which other areas of forensic science may be amenable to probabilistic treatment. Much probabilistic work had been done on evidence in the form of the refractive index of glass fragments with a seminal paper by Lindley (1977). The data for that paper had been provided by Ian Evett, who has published many papers on statistics, probabilistic inference and forensic science over the years. Since then, the technical aspects of the evaluation of scientific evidence have become increasingly accepted in the courts and the emphasis has become increasingly on interpretation and the best approach for the communication of evidential value to the fact-finder, judge or jury depending on jurisdiction.

The likelihood ratio is generally accepted as the best way of evaluating evidence. That it is the best way was first explained by Good (1989) and it is explicitly recommended by the European Network of Forensic Science Institutes (ENFSI) in their guidelines for evaluative reporting (ENFSI, 2015). One can read:

> Evaluation of forensic science findings in court uses probability as a measure of uncertainty. This is based upon the findings, associated data and expert knowledge, case specific propositions and conditioning information. (ENFSI Guideline, point n. 2.3, at p. 6)

and

> Evaluation will follow the principles outlined in Guidance note 1 (refer to paragraph 4.0). It is based on the assignment of a likelihood ratio. Reporting practice should conform to these logical principles. This framework for evaluative reporting applies to all forensic science disciplines. The likelihood ratio measures the strength of support the findings provide to discriminate between propositions of interest. It is scientifically accepted, providing a logically defensible way to deal with inferential reasoning. (ENFSI Guideline, point n. 2.4, at p. 6)

The likelihood ratio is part of the odds form of Bayes' theorem. Consider two propositions in the context of a criminal trial, H_p: the defendant is guilty, H_d: the defendant is not guilty. The value of evidence E is of interest. The odds form of Bayes' theorem states that

$$\frac{\Pr(H_p \mid E)}{\Pr(H_d \mid E)} = \frac{\Pr(E \mid H_p)}{\Pr(E \mid H_d)} \times \frac{\Pr(H_p)}{\Pr(H_d)}. \tag{1}$$

The term $\Pr(H_p)\Pr(H_d)$ is the odds in favour of guilt before the presentation of the evidence (prior odds), the term $\Pr(H_p \mid E)\Pr(H_d \mid E)$ is the odds in favour of guilt after the presentation of the evidence (posterior odds) and the third fraction $\Pr(E \mid H_p)\Pr(E \mid H_d)$ is the likelihood ratio, the term which converts prior odds to posterior odds. For those familiar with probabilistic reasoning, the interpretation of the likelihood ratio is straightforward. It can be interpreted to someone with words like: 'Consider your odds in favour of guilt prior to hearing the evidence. Listen to the evidence and determine how much more likely the evidence is if the defendant is guilty than if they are innocent. Multiply your prior odds by that amount and the result is your posterior odds in favour of guilt.' This wording is not new. In the context of the

Dreyfus case, Darboux et al. commented that

> since it is absolutely impossible for us [the experts] to know the *a priori* probability, we cannot say: this coincidence proves that the ratio of the forgery's probability to the inverse probability is a real value. We can only say: following the observation of this coincidence, this ratio becomes *X* times greater than before the observation. (1908: 504)

Unfortunately, for those unfamiliar with probabilistic reasoning (and, we suggest, this may be most people in the context of a criminal trial, judge, advocates and barristers, jurors), this interpretation is far from straightforward. Thus it was, as the technical matters became more acceptable, the emphasis of research on the role of probabilistic reasoning and statistics in forensic science moved from the technical matters towards interpretation. Lay people, in a statistical sense, have difficulty understanding the statistician's verbal description of the likelihood ratio. Many errors of interpretation are given in Aitken et al. (2021).

The underlying premise of these ideas is that legal reasoning is probabilistic in nature. In a civil trial, judgment is based on a balance of probabilities. In a criminal trial, judgment is based on a finding beyond reasonable doubt, where that concept may be considered probabilistically. However, an alternative premise has recently been suggested. Instead of probabilistic reasoning as the basis for juridical proof the suggestion is that juridical proof should be based on a concept known as *explanationism*. The concept of explanationism has been developed because of apparent analytic difficulties detected in the probabilistic paradigm. It has been argued that the strengths of the competing paradigm involving explanatory inferences (referred to as *relative plausibility*) have become too persuasive to ignore. It is explained here that probabilistic reasoning can overcome these 'analytic difficulties'. As a consequence, probabilistic reasoning remains the best way in which to evaluate and interpret evidence.

Interpretation – probability and relative plausibility

Allen and Pardo (2019) in the introductory paper to a special issue of International Journal of Evidence and Proof discuss the merits of the approach known as relative plausibility to interpretation as a replacement for probabilistic reasoning and argue that a paradigm shift is being experienced. The remainder of the special issue contains 20 short papers discussing the relative merits of relative plausibility and a final rejoinder from Allen and Pardo. The interested reader can study these papers at their leisure. Here we comment briefly on four problems that Allen and Pardo (2019) argue render probability implausible. These are the problems of:

- the subjective nature of an assignation of coherent numbers to evidence;
- the item-by-item argument associated with a probabilistic account, rather than a holistic argument;
- the inconsistency between the assumptions of probability and those of legal doctrine and jury instructions;
- the inconsistency of probabilistic thresholds for standards of proof with the legal process and goals for these standards.

Allen and Pardo allow for the relationship between probability and uncertainty. The terms 'probabilistic reasoning' and 'probability calculus' are synonymous. The only way to combine uncertainties about events or different items of evidence is to use the rules of probability.

The comments about reasoning under uncertainty are made in the context of the law. However, the ability to reason coherently in the presence of uncertainty is necessary in many disciplines, not just the law. Probabilistic reasoning is the best way to do it. Even critical commentators (e.g. Allen, 2013) note that 'none of the conceptualizations of probability except that as a subjective degree of belief can

function at trial'. Note also that the choice of probability as a framework for reasoning under uncertainty is not made to suggest that it provides a model for the working of the legal process as a whole. Instead, it is argued that probability helps (i) formulate principles for logical modes of reasoning, (ii) formulate relevant questions and (iii) define the boundaries between the competences of expert witnesses and legal decision-makers. There is no suggestion that the responsibility for the reasoning and the decision-making processes is to be delegated to an abstract theory or that statisticians and other reasoners with probability should comment on the administration of the law. However, there are many instances in reports of criminal cases of incorrect probabilistic reasoning. At least eight different types of fallacious probabilistic reasoning in legal cases identified in the literature are gathered together in Aitken et al. (2021). It is to aid in the reduction of these instances of fallacious reasoning that the arguments in support of probabilistic reasoning in the criminal justice system are presented here.

A fifth problem, known as the overlapping problem (Pardo, 2013), is also discussed. This problem concerns the situation where the same evidence is interpreted differently by the prosecution and defence.

The subjective nature of an assignation of coherent numbers to evidence

Allen and Pardo (2019) cite *Anderson* v *Liberty Lobby, Inc.*, 477 U.S. 242, 252 (1986), which explains that summary judgment depends on whether 'reasonable jurors could find by a preponderance of the evidence that the plaintiff is entitled to a verdict' and *Reeves* v *Sanderson Plumbing Prods., Inc.*, 530 U.S. 133, 150 (2000), which explains that the standard for judgement as a matter of law 'mirrors' the summary-judgment standard and *Jackson* v *Virginia*, 443 U.S. 307, 318 (1979), which explains the sufficiency standard for criminal convictions depends on whether 'any rational trier of fact could find guilt beyond a reasonable doubt'. Allen and Pardo argue that the legal standards for measuring sufficiency of the evidence in civil and criminal cases depend on being able to separate reasonable from unreasonable conclusions. They then argue that under the subjective account every decision is reasonable (at least so long as it is otherwise internally consistent). The subjective account thus fails to explain the central feature of the legal doctrine of reasonableness.

This argument conflates internal consistency with reasonableness. It is to be hoped that supporters of probabilistic reasoning are reasonable people who can distinguish reasonable inferences and decisions from unreasonable (though internally consistent) inferences and decisions. It is also to be hoped that the numerical assignations of probability to uncertain outcomes are chosen carefully to represent their measure of belief and are not unconstrained 'in any way by the quality of the evidence or its probative value in proving the facts of consequence at trial' (Allen and Pardo, 2019: 12). Also, decisions by fact-finders are ultimately subjective, whether or not numbers are involved. Those who argue numerically assign numbers to measures of uncertainty whereas those who do not argue numerically assign verbal descriptions to them. Both lines of argument are subjective. However, errors of interpretation are more easily exposed if a numerical argument is used than a verbal argument. Many errors of interpretation are listed in Aitken et al. (2021). It is not that probabilistic reasoners do not make errors, only that their reasoning is more explicit than that of those who argue verbally. An important advantage of the probabilistic approach is that the lines of reasoning of an argument expressed by an expert, a judge or a party to the case are made explicit and it is easy to argue against numerical values chosen. It is easier to analyse critically a probabilistic statement about the relationship between a certain finding and a proposition than a verbal statement that the finding is 'consistent with' the proposition or that a given proposition is a logical consequence of a given observation, for example.

The item-by-item argument associated with a probabilistic account rather than with a holistic argument

The probabilistic account (in both objective and subjective forms) is inconsistent with how fact-finders process and reason with evidence, which occurs holistically and not in the item-by-item fashion envisioned by the

account. Although the account may nevertheless have some normative or educational value in understanding aspects of evidence, it is plainly inconsistent with how actual evidence evaluation works. This inconsistency is a problem for any such empirical account that purports to explain actual standards of proof because it appears to require or imply that fact-finders evaluate evidence in a manner that they obviously do not. (Allen and Pardo, 2019: 13)

Bayes' theorem provides a very good method for reasoning with evidence (e.g. Hahn et al., 2013) and for updating beliefs as evidence is led (e.g. Oaksford and Chater, 2007). Consider the evidence E in (1) as E_1. Take the posterior odds of (1), with E_1 replacing E, as the prior odds for new evidence E_2. Then,

$$\frac{\Pr(H_p \mid E_1, E_2)}{\Pr(H_d \mid E_1, E_2)} = \frac{\Pr(E_2 \mid H_p, E_1)}{\Pr(E_2 \mid H_d, E_1)} \times \frac{\Pr(H_p \mid E_1)}{\Pr(H_d \mid E_1)}. \tag{2}$$

The posterior odds in favour of guilt given the two pieces of evidence is easily obtained through a sequential process of multiplication, with due attention paid to dependence between the two pieces of evidence. The extension to more pieces of evidence follows analogously. As a trial progresses, the beliefs can be updated sequentially. This requires each piece of evidence to remain uncertain throughout the trial. It is not the case that a piece of evidence, once heard, is either supportive of H_p or H_d with an associated assigned probability 1 representing certainty.

Allen and Pardo (2019: 12) in footnote 43 disagree with this approach to assimilation of evidence

What any offer of evidence means depends on all the evidence in the case. What might appear initially to be inculpatory may actually turn out to be exculpatory. Thus, serial updating is a waste of time and also confuses what is a learning process with an updating process. Trials are educational processes where fact-finders learn about the parties' explanations rather than updating processes that involve updating a complete and coherent set of priors. The fact-finder typically has no idea what those 'priors' are until all the evidence is in.

Evidence is led sequentially. It is hard to imagine a fact-finder withholding (at least privately) a belief in the ultimate issue until all the evidence has been heard. The value of each new item of evidence, E_n say, is dependent on all the previous items of evidence, $E_{n-1}, \ldots, E_1$ and the posterior odds after the introduction of each new item of evidence is an update on the previous posterior odds. The updating process caters for initially inculpatory evidence turning out to be exculpatory.

As well as incremental updating, a holistic view of the evidence is provided by so-called Bayesian networks or Bayes' nets (Roberts and Aitken, 2014; Robertson and Vignaux, 1993; Taroni et al., 2014). Each piece of evidence at a trial and the ultimate issue (e.g. guilt or innocence) is represented by a node in the network. The relationships amongst the different pieces of evidence and the ultimate issue are represented by lines joining appropriate nodes. The strengths of the relationships are noted in an adjoining conditional probability table. Software is available to process Bayes' nets. As each piece of evidence is considered, the Bayes' net can be updated incrementally as with the formulaic updating of (2).

It has been commented that Wigmore charts (Roberts and Aitken, 2014; Wigmore, 1937) and Bayes' nets 'will generally reflect the standpoint of the individual constructing the graph, rather than any putative "objective truth"' (Dawid et al., 2011: 147). This is a valid point. A Bayes' net helps an individual to structure the evidence and the relationships amongst different pieces of evidence better to understand the overall case. It is a benefit of the Bayes' net that an individual can be assisted in their structuring of the evidence. Bayes' nets from different individuals considering the same trial may be compared. Similarities and differences in their understanding of the evidence will be easy to determine. Debate amongst the individuals concerned will, hopefully, lead to a common Bayes' net and a consensus view of the relationship between the evidence and the ultimate issue.

It is not suggested in consideration of a case that all possible combinations of items of evidence with associated uncertainties be considered by an evaluator of the evidence. Of course, no human being has the

time to do the computations. That is why Bayes' nets were developed (see Lauritzen and Spiegelhalter, 1988 with a hint 'local computations' in the title of the paper of the method described). As an example of this approach in practice consider an example involving offender profiling (or, more correctly, specific case analysis); see Aitken et al. (1996). There, Bayes' nets were developed in a collaboration amongst senior police officers, statisticians and computer scientists to add structure to the investigation of sexually motivated child homicides. Various nets were developed and subjective conditional probabilities included based on the opinions of the police officers. One of the group was a very assiduous police officer who had compiled a database of all known cases of sexually motivated child homicides (sadly, a large number). From this database, frequentist conditional probabilities were able to be obtained and compared with the subjective conditional probabilities of the senior police officers. The two sets of probabilities were very similar. On reflection, this result is not surprising. The officers concerned were very experienced in these types of cases and had a very good feel for the values of the conditional probabilities with which the group was trying to work.

Bayes' nets provide the ability to deconstruct a complicated scenario into local relationships. It is then possible for the appropriate conditional probabilities to be assigned subjectively. With suitably experienced scientists or police officers these subjective probabilities will be close to reality. The ability of experienced individuals to express their uncertainties as probabilities enables Bayes' nets to lead to a more accurate outcome than their detractors suggest. See Biedermann et al. (2011).

It is also possible to consider 'qualitative Bayesian networks'. Qualitative Bayes' nets have been developed and supported by scientific and legal literature (see, e.g. Robertson and Vignaux, 1993). The logic dates back to Wigmore (1913). The qualitative relationships amongst variables cover qualitative influences which define the sign of the direct influence one variable has on another. This can be easily translated into sentences like 'variable A makes variable B more likely'.

The ability of Bayes' nets to deal with complex scenarios is not just applicable to scientific evidence. Other examples of their use in discussions of criminal cases are given by Edwards (1991), Kadane and Schum (1996) and Thagart (2003).

The inconsistency between the assumptions of probability and those of legal doctrine and jury instructions

Allen and Pardo (2019) note that 'legal doctrine and jury instructions typically dictate that the standard of proof applies to the individual elements, not to cases as a whole. Thus, in a two-element claim, A and B, if the plaintiff proves each to 0.6, then the plaintiff wins according to the law. However, if these two elements are independent of each other, then the probability of plaintiff's claim is 0.36, not greater than 0.5. This inconsistency has come to be known as the conjunction problem.' Many legal scholars cite this problem as a reason why probabilistic reasoning cannot be used in a court of law.

This conjunction problem was described earlier by Cohen (1977). Denote the elements A and B as in Allen and Pardo (2019)'s notation. They are both part of the plaintiff's case. Both need to be 'proven' on the balance of probabilities to satisfy the burden of proof. There is one item of evidence, denoted E, and A and B are deemed independent, given E. The evidence is deemed to support both elements as it can be shown that $\Pr(A \mid E)$ and $\Pr(B \mid E)$ are both greater than 0.5. However, while both individual elements have associated probabilities greater than 0.5 their product can have a probability less than 0.5. For example, assume $\Pr(A \mid E)$ and $\Pr(B \mid E)$ are both equal to 0.6; both 'prove' the plaintiff's case on the balance of probabilities. However, A and B, given E, are independent and $\Pr(A, B \mid E) = \Pr(A \mid E) \times \Pr(B \mid E) = 0.36$; the case would fail on the balance of probabilities.

The example given by Cohen (1977) is that of a car driver who sues their insurance company because it refuses to compensate them after an accident. The two component issues (elements) that are disputed are, first, the circumstances of the crash (A) and, second, the terms of the driver's contract with the

insurance company (B). Suppose each of these issues was determined with a probability of 0.6 in favour of the driver. Using the argument above, the probability of the joint issue would be 0.36.

However, further analysis (see, for example, Nesson, 1985) reveals that the plaintiff's case (story) has a higher probability than any other case (story). Denote the complement (negation) of A as $\bar{A}$ and the complement of B as $\bar{B}$ such that $\Pr(\bar{A} \mid E) = 1 - \Pr(A \mid E)$ and $\Pr(\bar{B} \mid E) = 1 - \Pr(B \mid E)$. Then, for example, $\Pr(A \mid E) = 0.6 \Rightarrow \Pr(\bar{A} \mid E) = 0.4$ and $\Pr(B \mid E) = 0.6 \Rightarrow \Pr(\bar{B} \mid E) = 0.4$. There are four possible cases: $\{A, B\}$, $\{A, \bar{B}\}$, $\{\bar{A}, B\}$ and $\{\bar{A}, \bar{B}\}$ with associated probabilities, conditional on E,

$$\Pr(A, B \mid E) = \Pr(A \mid E) \times \Pr(B \mid E) = 0.6 \times 0.6 = 0.36;$$
$$\Pr(A, \bar{B} \mid E) = \Pr(A \mid E) \times \Pr(\bar{B} \mid E) = 0.6 \times 0.4 = 0.24;$$
$$\Pr(\bar{A}, B \mid E) = \Pr(\bar{A} \mid E) \times \Pr(B \mid E) = 0.4 \times 0.6 = 0.24;$$
$$\Pr(\bar{A}, \bar{B} \mid E) = \Pr(\bar{A} \mid E) \times \Pr(\bar{B} \mid E) = 0.4 \times 0.4 = 0.16.$$

The case with the highest probability, conditional on E, is $\{A, B\}$. It should not be necessary that, in order to for a case to be proven on the balance of probabilities, the probability of the case is greater than the probabilities of all other cases put together. It should only be necessary that the case has a higher probability than that of each of the other cases taken individually.

A similar approach to this conjunction problem is best described by the illustration of what is known as Linda's example (Sides et al., 2002). Here, the evidence E is the following information about Linda: 'she is 31 years old, single, outspoken, and very bright; she majored in philosophy and, as a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations'.

The elements are

- A: Linda is a bank teller.
- B: Linda is active in the feminist movement.
- (A, B): Linda is a bank teller and is active in the feminist movement.

A majority of respondents across a variety of studies ranked the conjunction of A and B as more probable than A (see Hertwig and Chase (1998) for a review of findings; the original report is Tversky and Kahneman (1983)). This ranking is in contradiction to the laws of probability that state that, for two events A and B, $\Pr(AB) = \Pr(B \mid A)\Pr(A) \leq \Pr(A)$, since $\Pr(B \mid A) \leq 1$.

This paradoxical finding can be understood by considering relative values of unconditional and conditional probabilities. Consider a definition that evidence E favours (A, B) more than it favours A if and only if

$$\frac{\Pr(A, B \mid E)}{\Pr(A, B)} > \frac{\Pr(A \mid E)}{\Pr(A)}.$$

It can be shown that this inequality holds if and only if

$$\Pr(E \mid A, B) > \Pr(E \mid A).$$

This inequality asserts that Linda is more likely to be single, outspoken etc. on the assumption that she is a feminist bank teller than on the assumption that she is a bank teller. This is a reasonable assertion.

The numbers in the example from Nesson (1985) are for illustrative purposes only. The important principle in the discussion of Linda's example is that it is the relative values of the probabilities, not their absolute values, that are important. For a discussion of this principle, an example is given in Cheng (2013) with a rebuttal in Allen and Stein (2013), a rebuttal which is, in turn, rebutted below. Consider (Cheng, 2013) independent elements A and B in a case, and evidence E such that $\Pr(A \mid E) =$

0.9 and $\Pr(B \mid E) = 0.4$. Then, it can be argued that $A|E$ favours the plaintiff's case and $B \mid E$ favours the defence case. However, suppose the prior probabilities are $\Pr(A) = 0.5$ and $\Pr(B) = 0.2$. Then

$$\frac{\Pr(A \mid E) \times \Pr(B \mid E)}{\Pr(A) \times \Pr(B)} = \frac{0.36}{0.10} = 3.6 =$$

$$\frac{\Pr(A, B \mid E)}{\Pr(A, B)} > \frac{\Pr(A \mid E)}{\Pr(A)} = \frac{0.9}{0.5} = 1.8,$$

and

$$\frac{0.36}{0.10} = 3.6 = \frac{\Pr(A, B \mid E)}{\Pr(A, B)} > \frac{\Pr(B \mid E)}{\Pr(B)} = \frac{0.4}{0.2} = 2.0.$$

In both examples, evidence E favours (A, B) more than it favours A and more than it favours B. Also, $\Pr(A, B \mid E) = 0.36$ which is less than 0.5. However, this does not mean the plaintiff's case fails. The correct comparison for $\Pr(A, B \mid E)$ is not with 0.5 but with $\Pr(A, B)$. Given the assumed prior probabilities, E can be said to favour the plaintiff's case (A, B). The importance features of these results are (a) not that $\Pr(A \mid E) > 0.5$ and $\Pr(B \mid E) < 0.5$ but that $\Pr(A \mid E) > \Pr(A)$ and $\Pr(B \mid E) > \Pr(B)$ and (b) not that $\Pr(A, B \mid E) < 0.5$ but that it is greater than $\Pr(A, B)$. The absolute values of the probabilities are not important; it could be very difficult to assign numbers to these probabilities, the priors and the posteriors (those conditional on E), in many cases. However, it is reasonable for a trier of fact to make subjective judgements about the relative values of $\Pr(A \mid E)$ and $\Pr(A)$, of $\Pr(B \mid E)$ and $\Pr(B)$ and hence of $\Pr(A, B \mid E)$ and $\Pr(A, B)$. The trier may reason, for example, that B is still less likely to be true than false after E has been presented but it is not as less likely to be true than false than it was before E was presented. These subjective judgements need not be numerical, they can be verbal descriptions of relative values. In the words of Allen and Stein (2013), the approach based on relative probability and subjective judgement 'outscores' other approaches with respect to the factors of coherence, causal specificity and evidential support, factors which Allen and Stein (2013) say also inform plausibility and, by definition, the approach based on relative probability satisfies these factors within the probabilistic paradigm.

There is another conjunction problem, one which is discussed by Dawid (1987) and Cohen (1988). This is concerned with evidence evaluation, not the posterior probabilities of a case or elements of a case. This problem has only one element (better thought of here as a proposition), which is denoted H_p to be consistent with regular discussions about evidence evaluation and the likelihood ratio, the function of the evidence which provides the solution to this conjunction problem. The negation of H_p is denoted H_d. There are two pieces of evidence, denoted E_1 and E_2, which are assumed independent. Assume the conditional probability of H_p, conditional on E_1, is 0.6. Acceptance of E_1 enables H_p to be found on the balance of probabilities ($\Pr(H_p \mid E_1) > 0.5$). Assume, similarly, the conditional probability of H_p, conditional on E_2, is also 0.6. Acceptance of E_2 also enables H_p to be found on the balance of probabilities ($\Pr(H_p \mid E_2) > 0.5$). Thus, both E_1 and E_2 favour H_p in that these posterior probabilities are greater than 0.5, the assumed prior probability for H_p, $\Pr(H_p) = 0.5$, the neutral position in a civil case where the initial probabilities for H_p and H_d are taken to be equal to 0.5. However, taken together, the product $\Pr(H_p \mid E_1) \times \Pr(H_p \mid E_2) = 0.36 < 0.5$. It may then be thought that E_1 and E_2, taken together, give a result that finds for the negation of H_p since the joint probability is less than 0.5 when, singly, they both find for H_p. This appears paradoxical. However, note that the probability of interest is $\Pr(H_p \mid E_1, E_2)$. The error often made is to equate this probability to $\Pr(H_p \mid E_1) \times \Pr(H_p \mid E_2)$. The paradox may be resolved through use of the likelihood ratio.

As E_1 and E_2 have been assumed independent, given H_p or H_d, $\Pr(E_1, E_2 \mid H_p) = \Pr(E_1 \mid H_p) \times \Pr(E_2 \mid H_p)$ and $\Pr(E_1, E_2 \mid H_d) = \Pr(E_1 \mid H_d) \times \Pr(E_2 \mid H_d)$. Then

$$
\begin{aligned}
\frac{\Pr(H_p \mid E_1, E_2)}{\Pr(H_d \mid E_1, E_2)} &= \frac{\Pr(E_1 \mid H_p)}{\Pr(E_1 \mid H_d)} \times \frac{\Pr(E_2 \mid H_p)}{\Pr(E_2 \mid H_d)} \times \frac{\Pr(H_p)}{\Pr(H_d)} \\
&= \frac{\Pr(H_p \mid E_1)}{\Pr(H_d \mid E_1)} \times \frac{\Pr(E_2 \mid H_p)}{\Pr(E_2 \mid H_d)} \times \frac{\Pr(E_1)}{\Pr(E_1)} \\
&= \frac{\Pr(H_p \mid E_1)}{\Pr(H_d \mid E_1)} \times \frac{\Pr(E_2 \mid H_p)}{\Pr(E_2 \mid H_d)} \times \frac{\Pr(H_p)}{\Pr(H_d)} \times \frac{\Pr(H_d)}{\Pr(H_p)} \\
&= \frac{\Pr(H_p \mid E_1)}{\Pr(H_d \mid E_1)} \times \frac{\Pr(H_p \mid E_2)}{\Pr(H_d \mid E_2)} \times \frac{\Pr(H_d)}{\Pr(H_p)} \times \frac{\Pr(E_2)}{\Pr(E_2)} \\
&= \frac{0.6}{0.4} \times \frac{0.6}{0.4} \times \frac{\Pr(H_d)}{\Pr(H_p)} \\
&= 2.25 \times \frac{\Pr(H_d)}{\Pr(H_p)}.
\end{aligned}
$$

Assume again $\Pr(H_p) = \Pr(H_d)$, and so $\Pr(H_d) / \Pr(H_p) = 1$. Then

$$
\Pr(H_p \mid E_1, E_2) = \frac{2.25}{1 + 2.25} = 0.69,
$$

which is greater than 0.6. Thus, under the stated assumptions, the combination of the two items of evidence yields a higher probability for H_p than when considered separately. The problem of conjunction disappears. The flaw in the argument about the problem of conjunction is to assume independence of E_1 and E_2 implies $\Pr(H_p \mid E_1, E_2) = \Pr(H_p \mid E_1) \Pr(H_p \mid E_2)$ which, as the above argument shows, is not the case.

The first problem is faced by the triers of fact, the judges and juries. A decision has to be made, guilty or not guilty in a criminal case or a finding for the plaintiff or defendant in a civil case. In these situations, the burden of proof is, for example, beyond reasonable doubt in a criminal case, or on the balance of probabilities in a civil case. The decision should then be made on the situation which is best supported by the evidence. For those of a probabilistic bent, while the absolute values of the probabilities may be of interest, it is the relative values of the probabilities of the various stories provided by the prosecution and defence or plaintiff and defendant that are important. The second conjunction problem refers to the problem faced by an expert witness. They have to offer an opinion on the value of their evidence. This is best done with the likelihood ratio.

The likelihood ratio provides the best approach to the evaluation of evidence. It is a statistic which is a function of the probabilities of evidence, conditional on propositions, or elements. Consideration of this statistic is a matter for an expert witness. The inverse probabilities, those which are functions of probabilities of propositions or elements, conditional on evidence, are posterior probabilities. Consideration of these is a matter for a trier of fact.

In the absence of appropriate data, verbal summaries of relative values will suffice. The important principle is that it is relative values that matter, values that can be expressed verbally.

The inconsistency of probabilistic thresholds for standards of proof with the legal process and goals for these standards

A probabilistic approach to decision-making eases consistent decision-making. In a trial there are two sides, a plaintiff and a defendant in a civil trial and a prosecution and a defence in a criminal trial. Evidence is led and at the conclusion of the presentation of all the evidence and the summing up,

a decision has to be made by the fact-finder (judge or jury) to find for the plaintiff or the defendant in a civil trial or to find the defendant in a criminal trial guilty or innocent. As before, denote the plaintiff's or prosecution's case H_p, the defendant's case H_d and the totality of the evidence E. The odds version of Bayes' theorem (1) indicates how the posterior odds in favour of H_p may be determined from knowledge of the prior odds and the likelihood ratio.

The fact-finder using this approach is presented at the conclusion of the presentation of all the evidence and the summing up with a value for the posterior odds $\Pr(H_p \,|\, E)/\Pr(H_d \,|\, E)$. For ease of notation, denote the posterior odds Pst and the prior odds Pri. The fact-finder can use their own value judgement system to determine their decision. If it is a civil trial, perhaps a judgement that $Pst > Pri$ will suffice for the fact-finder to find for the plaintiff. If it is a criminal trial, some multiplying factor $K > 1$ should be needed in order to find the defendant guilty only if $Pst > KPri$ where the choice of K is a matter of personal choice, perhaps dependent on the severity of the charge against the defendant. It may be that a loss function needs to be considered, the value of which could affect the value of K (Taroni et al., 2010). These ideas are not new in legal disciplines, originating with Kaplan (1968) and Kaye (1988, 1999). Furthermore, such an approach has the attraction for forensic scientists that it is based on decisions. A description of the methodology and analysis of some forensic applications is given in Aitken et al. (2021).

It is assumed that H_p and H_d are mutually exclusive. It is to be hoped any overlap in the two explanations of events leading to the trial will have been agreed by the two sides and what is left for the trial has no overlap. There is no need for H_p and H_d to be exhaustive. It may be very difficult for them to be so. There may be some explanations which are left unstated at the beginning of the trial. However, it is the posterior odds (a misnomer if H_p and H_d are not exhaustive) that are important. If H_p and H_d are not exhaustive then it is not possible to derive posterior probabilities $\Pr(H_p \,|\, E)$ and $\Pr(H_d \,|\, E)$ from their ratio. However, it is still possible for E to be evaluated as the change in value from the prior odds to the posterior odds.

The overlapping problem

Rule 401(a) of the Federal Rules of Evidence states that '[e]vidence is relevant if it has any tendency to make a fact more or less probable than it would be without the evidence'. Equation (1) shows that this is the case if the likelihood ratio is different from one. Pardo (2013) argues that this condition is not necessary. It is possible for a likelihood ratio to be one but for the evidence to be relevant. In any given trial, there may be evidence that each side claims supports its theory of what happened, an 'overlap' of evidence. The example given to illustrate this point is paraphrased here. The defendant, an inmate at a maximum security prison, was charged with two counts of battery on prison guards. There was an altercation between the defendant and guards after the defendant refused to return a food tray in his cell. Evidence (E) was led that the defendant had not received a package sent to him by his family, details of the package and the mail procedures at the prison. The two competing theories for the assault were that (a) the defendant was frustrated and angry about not receiving the package, withheld the tray and attacked the guard (H_p); and (b) he withheld the tray to obtain the sergeant's attention and in response the guards attacked him (H_d). The argument against the use of the likelihood ratio as a measure of relevance is that the likelihood ratio in this example is one (and hence irrelevant using FRE 401) but the evidence is relevant as it provides information about the events that gave rise to the assault. Pardo (2013) noted that 'there is no reason to believe that this evidence supports one theory over the other. In other words, the likelihood ratio is 1.' He continued to argue that '[i]f this conception were true, the jury would have been forced to choose between who started the altercation without giving them any information about the events that precipitated it. The apparent relevance of this evidence – despite its likelihood ratio of 1 – provides a *reductio ad absurdum* for the theory.' Technically, this argument suggests that, while condition (a) of FRE 401 may be a sufficient condition for relevance, it is not a necessary one; evidence may still be relevant

even if the likelihood ratio is 1. However, there is more that can be said about the example of the altercation between the inmate and the security guards.

For evidence interpretation, the propositions being considered need to be specified clearly in a framework of circumstances. Furthermore, it is important to separate propositions from explanations (Evett et al., 2000; Hicks et al., 2015). The importance of this separation is discussed in Aitken et al. (2021).

> For example, statements like 'the crime stain originated from the PoI (person of interest)' and 'the crime stain originated from some unknown person who happened to have the same genotype as the PoI' represent explanations. The probability of the evidence that the DNA profile of the crime stain genotype matches the profile of the PoI given the first explanation is 1, but the probability of the evidence given the alternative explanation is also 1. Thus, the likelihood ratio is 1 and would not help advance resolution of the issues in the case. (2021: 489)

The quoted statements from Aitken et al. (2021) represent explanations, not propositions. Both statements explain the evidence of the DNA profile of the crime stain. Thus, it is not possible to determine the value of the evidence.

It is possible to define propositions so that the value of the evidence would be determined. For example, the prosecution proposition (H_p) could be that the source of the crime stain is the PoI and the defence proposition (H_d) could be that the source of the crime stain is some person unrelated to the PoI. The evidence (E) is the DNA profile of the crime stain and the DNA profile of the PoI. The value of the evidence is, as usual, the likelihood ratio $\Pr(E \mid H_p)/\Pr(E \mid H_d)$.

Pardo (2013) states that '[t]he two competing theories for the assault were that (a) the defendant was frustrated and angry about not receiving the package, withheld the tray and attacked the guard; and (b) he withheld the tray to obtain the sergeant's attention and in response the guards attacked him.' However, in the context of the discussion above about propositions and explanations, these two 'competing theories' are explanations, not propositions.

The corresponding propositions are better phrased as (a) the defendant attacked the guard (H_p) and (b) the guard attacked the defendant (H_d). The evidence (E) is that the defendant had not received a package sent to him by his family, details of the package and the mail procedures at the prison. It is then the jury's job to determine the likelihood ratio. Formally, in this example it is the ratio of the probability the defendant had not received a package sent to him by his family, details of the package and the mail procedures at the prison given that the defendant attacked the guard to the probability the defendant had not received a package sent to him by his family, details of the package and the mail procedures at the prison, given the guard attacked the defendant. This ratio may be better understood in terms of likelihood.

There is a reason why the ratio of probabilities, $\Pr(E \mid H_p)/\Pr(E \mid H_d)$, is known as the likelihood ratio. The numerator, $\Pr(E \mid H_p)$ depends on E and H_p. In its dependence on the former, E, it is thought of as a probability, in its dependence on the latter, H_p, as a likelihood. The dependence is emphasised by the use of the word 'given'; $\Pr(E \mid H_p)$ is the probability of E given H_p but is the likelihood of H_p given E. A similar argument applies to the denominator, $\Pr(E \mid H_d)$. When considering likelihood, it is E that is considered as fixed, as in the case in the example here where E is that the defendant had not received a package sent to him by his family, details of the package and the mail procedures at the prison. Further details of these ideas are given in Lindley (2014), from where the immediate discussion above was taken.

The jury is interested in the ratio of the likelihood the defendant attacked the guard given the defendant had not received a package sent to him by his family, details of the package and the mail procedures at the prison to the likelihood the guard attacked the defendant, given the defendant had not received a package sent to him by his family, details of the package and the mail procedures at the prison. There is no particular reason why this ratio should be one. There is no *reductio ad absurdum* for the likelihood ratio theory. With careful consideration of, and distinction between, explanations and propositions, the overlapping problem ceases to be a problem.

Another example of evidence that has a likelihood ratio of 1 and is relevant is given by Pardo (2013).

A witness will testify that someone matching the defendant's description was seen fleeing a crime scene. The defendant claims that it was his identical twin and introduces evidence establishing the twin's existence. Suppose there is no reason to believe the testimony distinguishes the defendant from his twin.

In a comparison of the likelihood of the defendant's guilt versus that of his twin, there does not appear to be any reason to think the likelihood ratio is different from 1: the ratio of the probability the person fleeing the crime scene 'matched' the defendant's description given the defendant's guilt to the probability the person fleeing the crime scene 'matched' the defendant's description given the defendant's identical twin's guilt will be 1. However, the evidence is relevant.

There are two pieces of evidence: E_1, the person fleeing the crime scene matched the defendant's description and E_2, the defendant has an identical twin. There are two propositions, H_p: the defendant is guilty; H_d: the defendant is not guilty. The prior odds in favour of the defendant's guilt is $\Pr(H_p)/\Pr(H_d)$. The likelihood ratio for E_1 is

$$\frac{\Pr(E_1 \mid H_p)}{\Pr(E_1 \mid H_d)},$$

which would be expected to be very large. The likelihood ratio for the combined evidence (E_1, E_2) is

$$\frac{\Pr(E_1, E_2 \mid H_p)}{\Pr(E_1, E_2 \mid H_d)} = \frac{\Pr(E_2 \mid E_1, H_p)}{\Pr(E_2 \mid E_1, H_d)} \frac{\Pr(E_1 \mid H_p)}{\Pr(E_1 \mid H_d)}.$$

There is a good argument that

$$\frac{\Pr(E_2 \mid E_1, H_p)}{\Pr(E_2 \mid E_1, H_d)} = \frac{\Pr(E_2 \mid H_p)}{\Pr(E_2 \mid H_d)} = 1.$$

The existence of the identical twin is independent of the evidence of the sight of the person fleeing the crime scene and of the defendant's guilt.

Another example given by Pardo (2013) to show the difficulty of a probabilistic approach to evidence evaluation illustrates how the use of a likelihood ratio highlights the differing factors which need to be considered when assessing guilt. The example concerns two poisons. The prosecution alleges (H_p) that the victim died of poisoning, and the defence (H_d) contends that the victim died from some other cause. There is evidence that at 12:00 p.m. on the day he collapsed and died, the victim's lunch contained a poison (A) that is fatal for 90% of people who ingest it. There is also evidence that at 12:30 p.m., the victim ingested a second poison (B) concealed in a drink that completely counteracts the first poison A but that is fatal for 80% of people who ingest it. Then $\Pr(H_p \mid A) = 0.9$, $\Pr(H_p \mid B) = 0.8$.

Assume H_p and H_d are mutually exclusive and exhaustive; i.e., it is assumed the victim is dead and the point of issue is whether the victim died of poisoning or of some other cause. The evidence is the ingestion of the two poisons A and B. The posterior odds in favour of H_p are

$$\frac{\Pr(H_p \mid A, B)}{\Pr(H_d \mid A, B)}.$$

As B counteracts A, it is reasonable to assume that

$$\frac{\Pr(H_p \mid A, B)}{\Pr(H_d \mid A, B)} = \frac{\Pr(H_p \mid B)}{\Pr(H_d \mid B)} = 0.8/0.2 = 4.$$

It is unnecessary given these probabilities to consider likelihood ratios for the value of the evidence of

ingestion of A and then B. The posterior odds

$$\frac{\Pr(H_p \,|\, A, B)}{\Pr(H_d \,|\, A, B)} \neq \frac{\Pr(H_p \,|\, A)}{\Pr(H_d \,|\, A)} \cdot \frac{\Pr(H_p \,|\, B)}{\Pr(H_d \,|\, B)}.$$

The posterior odds may be written as

$$\frac{\Pr(H_p \,|\, A, B)}{\Pr(H_d \,|\, A, B)} = \frac{\Pr(A, B \,|\, H_p))}{\Pr(A, B \,|\, H_d)} \frac{\Pr(H_p)}{\Pr(H_d)}.$$

Consideration of the prior odds is a matter for the triers of fact, not for any expert witness. The likelihood ratio may be written as

$$\frac{\Pr(A, B \,|\, H_p)}{\Pr(A, B \,|\, H_d)} = \frac{\Pr(B \,|\, A, H_p)}{\Pr(B \,|\, A, H_d)} \frac{\Pr(A \,|\, H_p)}{\Pr(A \,|\, H_d)}.$$

There is no information given in the narrative above that enables determination of these probabilities.

A final example from Pardo (2013) is that of a lottery.

There is a murder victim. The motive appears to be that the victim ran an illegal lottery and refused to pay the winner. It is unknown who actually won the lottery. The prosecution claims it was the defendant, and the defendant claims it was a rival. The defendant purchased one of the one thousand total lottery tickets and the rival purchased ninety-nine tickets. Suppose there is also evidence that the other nine hundred tickets were never sold and account has been taken of them.

The probability the defendant won the lottery is 1/100 or 1/1000, depending on whether or not the other nine hundred tickets were entered in the lottery or not. Similarly, the probability the rival won the lottery is 99/100 or 99/1000, depending on whether or not the other nine hundred tickets were entered in the lottery or not. The ratio of the two probabilities is 1/99, regardless of whether or not the other nine hundred tickets were entered in the lottery or not.

Conclusion

It has been written '[w]e believe it unfortunate that the triers of fact, jurors or judges (in a bench trial), never see evidence marshalled and arguments displayed in a Wigmorean manner. Wigmore's intended audience for his methods consisted of attorneys *preparing* for trial; he said almost nothing about the impact of his methods on the behaviour of triers of fact.' (Kadane and Schum, 1996: 258.) Likewise, it is unfortunate that barristers, advocates and attorneys do not make better use of the insights into the structure of the entirety of the evidence at trial afforded by the use of Bayesian networks. A well-presented Bayesian network would aid prosecutors and defenders in their case summaries at the beginning and end of a trial and help triers of fact in their deliberations throughout the trial.

Probabilities are subjective beliefs, the likelihood ratio is the measure for the value of evidence and the Bayesian approach is the approach to update beliefs as evidence is presented. Bayesian networks are a tool for use in practice to clarify the structure and dependencies between propositions and masses of evidence. Quantification through probability provides an approach to the evaluation of evidence which is clearer and more easily analysed critically than a verbal description of the value. A sequential probabilistic argument provides a coherent approach to evaluation and an improved assessment of the evidence than an argument that tries to consider all the evidence as a whole or one that mistakenly combines posterior probabilities.

Many have echoed Mr Justice Holmes' 1897 opinion that 'For the rational study of the law the black-letter man may be the man of the present, but the man of the future is the man of statistics' (at p. 469) (Holmes, 1897). Twining (1990) added – to the previous Holmes' statement – that

> […] the calculus of chance, Bayes' Theorem, and a general ability to spot fallacies and abuses in statistical argument should be part of the basic training of every lawyer. (1990: 14)

Twining (1990) then reiterated the concept by writing

> The continuing 'probabilities debate' is more problematic. For, while it raises fundamental problems in probability theory and logic, it has important practical implications. The 'unreality' of some of these debates is in part attributable to the over-concentration of naive examples, such as the rodeo and blue and green bus problems, which do not take account of important procedural and other contextual factors. But theories of probability are central to evaluation of evidence and to elucidating basic evidentiary concepts such as 'standard of proof', probative value and 'prejudicial effect'; and statistical evidence is of increasing practical importance in many different types of proceeding. […] The lawyer of today needs to be a master of elementary statistics. The general aspects of this can be accommodated within the logic of proof […]. (1990: 209)

There have been numerous calls to pay attention to these concepts, ranging from Eggleston, who wrote 'it is the purpose of this book to discuss the part which probabilities play in the law, and the extent to which existing legal doctrine is compatible with the true role of probabilities in the conduct of human affairs' (1983: 3) to the former President of the Italian Supreme Court, who recently noticed:[1]

> The challenge shifts to the field of the concrete effectiveness of the procedural model, while a new professional figure of magistrate is emerging in judicial practice. This would be a man of all-round culture, not only juridical but also humanistic and scientific, a responsible and effective evaluator of the fact and interpreter of law, a good reasoner and quality decision maker, skilled in the exercise of the art of judging, expert in logical-inferential techniques and in the verification of statistical-probabilistic schemes, as well as in the techniques of argumentative writing. However, he should be free from constraints and conditions other than the law, the reason and the ethics of limit and doubt. (Canzio, 2021: 91)

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

ORCID iD

Colin Aitken https://orcid.org/0000-0001-8457-4383

1. La sfida si sposta, a ben vedere, sul terreno della concreta efficacia del modello processuale, mentre va stagliandosi all'orizzonte, nella pratica giudiziaria, una nuova figura professionale di magistrato. Un uomo di cultura a tutto tondo, non solo giuridica ma anche umanistica e scientifica, un responsabile e ed efficace valutatore del fatto e interprete del diritto, un buon ragionatore e decisore di qualità, abile nell'esercizio dell'arte del giudicare, esperto nelle tecniche logico-inferenziali e nella verifica degli schemi statistico-probabilistici, come pure nelle tecniche della scrittura argomentativa. E però libero da vincoli e condizionamenti che non siano la legge, la ragione e l'etica del limite e del dubbio.

References

Aitken CGG (1995) *Statistics and the Evaluation of Evidence for Forensic Scientists*. Chichester: John Wiley & Sons.

Aitken CGG, Connolly T, Gammerman A, et al. (1996) Statistical modelling in specific case analysis. *Science and Justice* 36(4): 245–255.

Aitken CGG, Taroni F and Bozza S (2021) *Statistics and the Evaluation of Evidence for Forensic Scientists*, 3rd ed. Chichester: John Wiley & Sons.

Allen RJ (2013) Taming complexity: Rationality, the law of evidence and the nature of the legal system. *Law, Probability and Risk* 12(2): 99–113.

Allen RJ and Pardo MS (2019) Relative plausibility and its critics. *International Journal of Evidence and Proof* 23(1–2): 5–59.

Allen RJ and Stein A (2013) Evidence, probability and the burden of proof. *Arizona Law Review* 55: 557–602.

Biedermann A, Taroni F and Thomson WC (2011) Using graphical probability analysis (Bayes' nets) to evaluate a conditional probability. *Law, Probability and Risk* 10(2): 89–121.

Canzio G (2021) Alle radici dell'errore giudiziario: 'heuristics and biases'. In: Luparia Donati L (ed) *L'Errore Giudiziario*. Milano: Giuffrè Francis Lefebvre, 81–91.

Cheng E (2013) Reconceptualizing the burden of proof. *Yale Law Journal* 122(5): 1254–1279.

Cohen JL (1977) *The Probable and the Provable*. Oxford: Clarendon Press.

Cohen JL (1988) The difficulty about conjunction in forensic proof. *The Statistician* 37(4/5): 415–416.

Cook R, Evett IW, Jackson G, et al. (1998a) A hierarchy of propositions: Deciding which level to address in casework. *Science & Justice* 38(4): 231–239.

Cook R, Evett IW, Jackson G, et al. (1998b) A model for case assessment and interpretation. *Science & Justice* 38(3): 151–156.

Darboux JG, Appell PE and Poincaré H (1908) Examen critique des divers systèmes ou études graphologiques auxquels a donné lieu le bordereau. In: *L'affaire Dreyfus – La révision du procès de Rennes – Enquête de la chambre criminelle de la Cour de Cassation*. Paris: Ligue Française des droits de l'homme et du citoyen, 335–393.

Dawid AP (1987) The difficulty about conjunction. *The Statistician* 36(2/3): 91–97.

Dawid AP, Schum DA and Hepler A (2011) Inference networks: Bayes and Wigmore. In: Dawid AP, Twining W and Vasilaki M (eds) *Evidence, Inference and Enquiry*. Oxford: Oxford University Press for The British Academy, 119–150.

Edwards W (1991) Influence diagrams, Bayesian imperialism, and the *Collins* case: An appeal to reason. *Cardozo Law Review* 13(2/3): 1025–1074.

Eggleston R (1983) *Evidence, Proof and Probability*. London: Weidenfeld and Nicolson.

ENFSI (2015) ENFSI guideline for evaluative reporting in forensic science. Dublin: ENFSI.

Evett IW, Jackson G and Lambert JA (2000) More on the hierarchy of propositions: Exploring the distinction between explanations and propositions. *Science & Justice* 40(1): 3–10.

Good IJ (1989) C319: weight of evidence and a compelling metaprinciple. *Journal of Statistical Computation and Simulation* 31(2): 121–123.

Hahn U, Harris AJL and Oaksford M (2013) Rational argument, rational inference. *Argument and Computation* 4(1): 21–35.

Hertwig R and Chase VM (1998) Many reasons or just one: how response mode affects reasoning in the conjunction problem. *Thinking and Reasoning* 4(1): 319–352.

Hicks TN, Biedermann A, de Koeijer JA, et al. (2015) The importance of distinguishing information from evidence/observations when formulating propositions. *Science & Justice* 55(6): 520–525.

Holmes OW (1897) The path of the law. *Harvard Law Review* 10(8): 457–478.

Kadane JB and Schum DA (1996) *A Probabilistic Analysis of the Sacco and Vanzetti Evidence*. New York: John Wiley & Sons Inc.

Kaplan J (1968) Decision theory and the factfinding process. *Stanford Law Review* 20: 1065–1092.

Kaye DH (1988) What is Bayesianism? In: Tillers P and Green ED (eds) *Probability and Inference in the Law of Evidence, the Uses and Limits of Bayesianism*, Boston Studies in the Philosophy of Science. Dordrecht: Springer, 1–19.

Kaye DH (1999) Clarifying the burden of persuasion: What Bayesian rules do and do not do. *International Journal of Evidence and Proof* 3(1): 1–28.

Lauritzen SL and Spiegelhalter DJ (1988) Local computations with probabilities on graphical structures and their application to expert systems. *Journal of the Royal Statistical Society: Series B* 50(2): 157–194.

Lindley DV (1977) A problem in forensic science. *Biometrika* 64(2): 207–213.

Lindley DV (2014) *Understanding Uncertainty*, revised ed. Hoboken, NJ, USA: John Wiley and Sons Inc.

Nesson C (1985) The evidence or the event? On judicial proof and the acceptability of verdicts. *Harvard Law Review* 98: 1357–1392.

Oaksford M and Chater N (2007) *Bayesian Rationality: The Probabilistic Approach to Human Reasoning*. Oxford: Oxford University Press.

Pardo MS (2013) The nature and purpose of evidence theory. *Vanderbilt Law Review* 66(2): 547–614.

Roberts P and Aitken CGG (2014) *The Logic of Forensic Proof*. London: Royal Statistical Society. Available at: https://rss.org.uk/news-publication/publications/law-guides/, accessed, 16 July 2022.

Robertson B and Vignaux GA (1993) Taking fact analysis seriously. *Michigan Law Review* 91(6): 1442–1464.

Sides A, Osheron D, Bonini N, et al. (2002) On the reality of the conjunction fallacy. *Memory and Cognition* 30(2): 191–198.

Taroni F, Biedermann A, Bozza S, et al. (2014) *Bayesian Networks for Probabilistic Inference and Decision Analysis in Forensic Science*, 2nd ed. Chichester: John Wiley & Sons.

Taroni F, Bozza S, Biedermann A, et al. (2010) *Data Analysis in Forensic Science; a Bayesian Decision Perspective*. Chichester: John Wiley & Sons.

Thagart P (2003) Why wasn't O.J. convicted? Emotional coherence and legal inference. *Cognition and Emotion* 17(3): 361–383.

Tversky A and Kahneman D (1983) Extensional versus intuitive reasoning: the conjunction fallacy in probability judgment. *Psychological Review* 90(4): 293–315.

Twining W (1990) *Rethinking Evidence. Exploratory Essays*. Evanston, IL: Northwestern University Press.

Wigmore JH (1913) The problem of proof. *Illinois Law Review* 8(2): 77–103.

Wigmore JH (1937) *The Science of Judicial Proof: as Given by Logic, Psychology and General Experience and Illustrated in Judicial Trials*, 3rd ed. Boston, MA: Little, Brown.