1997

Of Flutes, Oboes and the As If World of Evidence Law

Richard O. Lempert
University of Michigan Law School, rlempert@umich.edu

Recommended Citation

Lempert, Richard O. "Of Flutes, Oboes and the As If World of Evidence Law." *International Journal of Evidence and Proof* 1 (1997): 316-320. DOI: https://doi.org/10.1177%2F1365712797001special09

Of flutes, oboes and the *as if* world of evidence law

By Richard Lempert
Francis A. Allen Collegiate Professor of Law and Professor of Sociology, University of Michigan

Reading Allen's article, I am reminded of a cold war parable I heard during the 1960s. It concerned a flute and an oboe who joined an orchestra one year and immediately set to quarrelling. The flute was distressed because whenever it was playing at its lyrical best the oboe would enter, drowning it out. The oboe was affronted because its deepest, most sonorous passages were invariably ruined by the high-pitched flute butting in. When the orchestra split up for the summer and these quarrelsome instruments went their separate ways, the flute, as it angrily contemplated the oboe, found itself stretching on tiptoes and trying to speak in its lowest voice. The oboe, on the other hand, despite its annoyance with the flute, could not resist speaking in falsetto and hunching over as it played. When the orchestra reassembled in the fall, it had two new clarinets.[1]

The story comes to mind because Allen is a prominent Bayesioskeptic while I have been labelled a Bayesian enthusiast,[2] yet I find myself agreeing with much of what Allen says in his article which, though styled a 'preliminary inquiry' seems to be a culmination of his thinking about the role of Bayes' theorem in understanding evidence and proof. Yet, I don't think we have both become clarinets; rather, Allen is sounding very much like an oboe, for he seems to be saying much that I have said all along.[3]

Let me highlight areas of agreement. First, Allen agrees that there can be settings in which Bayes' theorem is normative. Second, he does not deny that Bayes' theorem may be useful as an analytic tool. These are essentially the premises on which my article 'Modeling Relevance' (Lempert, 1977; see also Lempert, 1988) rests; namely, Bayes' theorem is a useful analytic tool for understanding the law's idea of logical relevance because it captures the way the law's fact finders *should* (not necessarily 'do') reason about evidence.

1 The story was current during the so-called 'Kruschev Thaw' at the start of Lyndon Johnson's Great Society programme and before the Gulf of Tonkin incident and massive escalation of the Viet Nam war. Its point was that both the Soviet Union and the United States seemed to admire features of the other and despite their competition and angry confrontations might be on convergent paths.

2 I do not know where the terms 'Bayesioskeptic' or 'Bayesian enthusiast' come from, but I have no better shorthand to define what are thought to be conflicting positions about how evidence is, might be or should be processed.

3 This is not to say that there are not some double bassoons among the ranks of Bayesian enthusiasts, with whom both Allen and I would disagree.

Third, Allen allows that even at trials Bayes' theorem may provide fact finders with an appropriate framework for making sense of statistical evidence and maybe even certain non-statistical evidence. Using Bayes' theorem to present statistical evidence is something I (Lempert, 1991), and others (Evett and Werrett, 1990; Evett, 1992b; Evett *et al.*, 1992; Kaye 1993), have advocated when DNA evidence is offered, and perhaps when other statistical evidence is presented as well (Lempert, 1988).

Fourth, Allen agrees that Bayes' theorem may be useful for juridical objectives other than understanding or improving the way fact finders reason, and specifically as an analytic tool.[4] This is the use I make of Bayes' theorem in my article 'Modeling Relevance' (Lempert, 1977) and it is a use I defend in my article 'The New Evidence Scholarship' (Lempert, 1988).

Finally, I agree with Allen that except in a few situations involving statistical evidence, fact finders should not be instructed in the use of Bayes' theorem or told to apply Bayes' theorem to non-statistical legal evidence (Lempert, 1988; 64–5). I also agree with some of Allen's arguments regarding computational complexity, but like him do not believe that this is the only reason why jurors should not ordinarily be encouraged to employ Bayesian modes of evaluating trial evidence.[5] But this does not move me towards flutehood because I have never taken a contrary position. I made clear at the end of 'Modeling Relevance' (Lempert, 1977: 1056–7) that there are good reasons for not urging jurors to be formal Bayesians and that the descriptive utility of Bayes' theorem for modelling juror reasoning is an empirical question. In later work (Lempert, 1988: 65) I noted that the empirical evidence indicated that ordinary people reasoning about even simplified legal problems were not intuitive Bayesians.

Being in the same section of the orchestra does not, however, mean that Allen and I agree on everything. Allen, for example, disagrees with me or misunderstands me when I argue that the human ability to chunk information allows us to treat an array of evidence as one piece for Bayesian analytic purposes. This is neither inconsistent with a Bayesian approach, which recognises that evidence may be decomposed in varying degrees (Schum and Martin, 1982; Kadane and Schum, 1996), nor any great feat. Consider that in Allen's non-Bayesian world, a trial fact finder assimilates all the evidence presented and reaches a judgment about the verdict that taken together, the evidence portends, given the burden of proof. Although this judgment takes the form of an all or nothing assessment such as 'guilty' or 'not guilty,' the fact finder could presumably report its conclusion as a probability that the burden of proof has been met.[6]

What I suggested in the casually written paragraph from an electronic communication that Allen quotes is that in cases where DNA identification evidence is crucial,

4 In 'A Reconceptualization of Civil Trials', the article by Allen (1986), which my article, 'The New Evidence Scholarship: Analyzing the Process of Proof' (Lempert, 1988) engaged, Allen's most trenchant criticisms of Bayesianism were, unlike this article, analytic rather than empirical.

5 In the hands of experts operating at their leisure, Bayes' theorem may be a tool that helps make sense of empirical evidence despite great complexities. See the marvelous book by Kadane and Schum (1996).

6 People are used to speaking in probabilistic terms. Thus someone may say, 'I think the odds are 100 to 1 the defendant is guilty,' another might say 'there's three chances in four the driver was drunk'. Moreover, people seem to have no difficulty providing subjective probabilities of what different legal burdens of proof mean (see Simon and Mahan, 1971) or in giving researchers probabilistic estimates of the probative weight of evidence or of sets of items of evidence (Schum and Martin, 1982).

the fact finder, before considering the implications of the DNA evidence, should estimate on the basis of all the other evidence the probability that the DNA proponent's burden of proof has been met.[7] This estimate provides a prior probability that can then be adjusted in a Bayesian manner when the rarity of the DNA match is captured in a likelihood ratio rather than presented in frequentist terms. There is nothing illegitimate in this argument or particularly difficult about the task assigned to the fact finder. If a fact finder can acceptably estimate probabilities of guilt for all items of evidence taken together, as it must in Allen's world, surely it can do the same for all items except one. Indeed, the Bayesian approach I recommend in the material Allen extracted from our e-mail conversation is one he admits is valid.

As for what Allen suggests is a cavalier treatment of what in his view is an intractable situation of conditional non-independence, this is seldom a problem with the topic of my message, DNA evidence. The probability of a DNA match which is given to the jury is usually conditionally independent of other evidence in the case, since it is the probability that a random person would have DNA matching the evidence sample and not the probability that the defendant would have matching DNA. Where it is not conditionally independent, the implications of the conditional non-independence of a match should be taken into account whether match probabilities are presented as frequencies or as Bayesian likelihood ratios.

But enough quarrelling with Allen. What is interesting lies in our agreement not in our disagreement and in our ability to identify how we can agree on so much when Allen has, before this piece, been tenacious in his criticisms of Bayesian approaches to evidence law, while I have welcomed Bayes' theorem as an aid to clear thinking. I believe that our considerable agreement is possible because much evidence law, like the uses of Bayes' theorem that Allen criticises, lives in an *as if* world. By this I mean some law is written and courts often decide cases *as if* the frailties of human existence and the substantive complexities of actual trial evidence did not exist. Thus, evidence law often presumes that its fact finders have abilities that exceed the capacities of most humans. Allen has his eye on the actual world of trials. I, and to a large extent Friedman, are writing for the more abstract *as if* world of much evidence law.

In the law's *as if* world, jurors are able to follow a judge's instructions to disregard inadmissible evidence inadvertently presented to them or to consider admissible evidence only for a single narrow permitted purpose even when some other evidentiary implications appear far more obvious than the inference the law allows. For example, jurors are presumed to use information about an accused thief's past convictions to decide if the accused might be prone to lie but not as evidence that she is prone to steal, though research suggests jurors cannot so limit the impact of what they hear (Hans and Doob, 1976; Wissler and Saks, 1985). Judges do jurors one better, for the law conceives of them as able to hear far more inadmissible evidence than jurors typically encounter without being prejudiced by it. Indeed in the law's *as if* world, judges can be trusted to ignore improper evidence even if they have not explicitly recognised its inadmissibility. Both trial jurors and judges in this world are able to distinguish truthful and accurate witnesses from lying or inaccurate witnesses, even though body language and other demeanour cues are often

7 See also Friedman's well thought out argument in this volume that the value of Bayes' theorem in thinking about, or even perhaps at, trials does not depend on the fact finder being able to evaluate each individual item of evidence in Bayesian terms.

ambiguous or misleading.[8] Perhaps most astonishingly, jurors and judges are able to listen to the conflicting testimony of honest scientists, all of whom have needed years of advanced study to master their fields to the degree where the law will allow them to speak, and though the law's fact finders lack expertise and may never have encountered the issue before, they can decide accurately which expert's position has the greater scientific validity.[9]

With fact finders capable of such feats, it is not surprising that in evidence law's *as if* world, they can also be, for some purposes, perfectly rational fact finders whose thought processes are appropriately modelled by Bayes' theorem. In particular, trial judges in deciding issues of pure logical relevance are supposed to conceive of jurors in this way. This, at least, is my reading of Federal Rules of Evidence 401 and 402, and it is the premise on which my article 'Modeling Relevance' rests (Lempert, 1977).

It is no accident that evidence law has created an unreal world in which to operate. Some scholars would see this as inescapable. Niklas Luhmann, for example, has argued that all of society's primary subsystems (for example, the political, the economic, the legal) can only operate on self-defined concepts and that it is the fact of operating only on its own closed model of the world which makes the law a system (Luhmann, 1985a,b,c,d). But it is not ethereal theory that has created this condition. Practically speaking a judge's analytic tasks at both the trial level and on appeal are considerably eased by the acceptability of referencing idealised abstractions of how people treat evidence rather than their actual behaviour. Indeed, since people behave differently and their behaviour is often poorly understood, the law can find itself in trouble when it proceeds on an image of behaviour designed to approximate actual behaviour rather than some idealised vision of how rational people act. Thus, where courts attempt to make law that realistically appraises human behaviour, their initial decisions are often followed by ill-disguised retreats that lead to bodies of precedent that are neither behaviourally nor analytically consistent.

In *Lockhart* v *McCree* 476 US 162, 106 S Ct 1758 (1986), for example, the Supreme Court withdrew the invitation it had extended in *Witherspoon* v *Illinois* 391 US 510 (1968) and refused to consider the possibility that death-qualified jurors might have pro-prosecution biases on questions of guilt and innocence. After *Lockhart*, the law's world is one in which death-qualified jurors are treated for legal purposes as if they were no more likely to favour the state's case than a random group of people drawn from the relevant community and not subject to death qualification, despite substantial evidence to the contrary.[10]

In *Bruton* v *United States* 391 US 123 (1968) the law uncharacteristically eschewed its *as if* world of jurors who do as judges tell them for the empirical world in which jurors

8 For example, a large body of work indicates that eyewitness confidence has at best a slight relationship to eyewitness accuracy, but confident eyewitnesses appear far more credible than less confident ones (Wells and Loftus, 1984). On demeanour evidence generally, see, for example, Wellborn (1991).

9 See, for example, the *W.R. Grace & Co.* case discussed in Lempert (1993). This and my other textural examples should not be read as suggesting that in evidence law's *as if* world, the capacity of jurors is always overestimated. Indeed, the same jurors who can resolve knotty scientific questions that leave experts divided cannot be trusted adequately to discount statements attributed to people not in court, nor can they avoid pro-plaintiff biases if they learn the defendant before them is insured.

10 Consider as a body the studies dismissed by the majority in *Lockhart* v *McCree* 476 US 162, 106 S Ct 1758 (1986) at notes 4–12. See especially Cowan *et al.* (1984), Fitzgerald and Ellsworth (1984) and Haney (1984).

are unable to disregard the portion of one codefendant's confession that implicates the other. But *Bruton* led to tangles which it has taken a stream of cases, some of which return to idealised images of fact finders, to unravel. I point this out not to criticise the court, but rather to defend myself and my argument that Bayes's theorem, which admittedly presupposes an ideally rational fact finder, is a viable model for thinking about relevance as the law conceives it and other issues as well. The unreality of the law's assumptions about fact finders' capacities that are reflected in the Bayesian model is not an aberration; rather it is the stuff of which much law is made.

Conclusion

If all Allen means to do in his article is to argue that jurors do not reason in a Bayesian fashion and to conclude from this that Bayes' theorem is not a viable model of how fact finders think, then I agree with him. I also agree with Allen that ordinarily trial fact finders should not be encouraged to quantify and process the evidence they have heard along Bayesian lines. Apart from special situations that Allen recognises, jurors neither make nor can they be expected to make the kinds of estimates and combinations of estimates which a Bayesian fact finder would make to evaluate evidence. But if this is Allen's point, it is a small one, and I do not think it addresses most of the uses which the so-called Bayesian enthusiasts have made of the Reverend Bayes' theorem.

If Allen has the broader goal of moving evidence scholarship in the direction of attempting to gain a better understanding of how jurors and judges actually reason about evidence and reach decisions, I applaud his efforts as will, I am sure, the many social scientists who have made this project their life's work. If the latter goal underlies Allen's article, that is, if he wishes to challenge evidence law's *as if* world through greater understanding of how legal fact finders process information and make decisions, he is embarking on and is inviting others to join him in a truly radical challenge to received law and, indeed, to ways of thinking about law. In doing so he points to exciting paths for evidence scholars to explore. The effort, if taken seriously, can pose a fundamental normative challenge to much of evidence law and so has the potential for reformulating much of this body of rules. The idea of moving evidence scholarship in this direction is not new, for it can be traced back at least to Munsterberg (1908). What is noteworthy is for as keen an analytic lawyer as Allen to move his attention from the analytic to the empirical[11] and to suggest that it is in the empirical world where answers to profound questions about the shapes of trials and the meaning of proof lie. Yet maybe it is not strange that Allen's work has taken an empirical turn. As a social scientist who has long been interested in the empirical aspects of trial fact finding, I can testify that focusing on the empirical is what oboes like most to do.

11 Contrast Allen's attack on Bayesian approaches to trials in Allen (1986), with the empirical case against Bayesianism he makes in his current article.