Extending the confusion about Bayes

Bernard Robertson, Barrister, Wellington
G A Vignaux, Victoria University of Wellington and
Charles E H Berger, Netherlands Forensic Institute

"The probability that an animal with four legs is a cow
is not the probability that a cow has four legs."

In *R* v *T* [2010] EWCA Crim 2439 (26 October 2010), the Court of Appeal indicated that "mathematical formulae", such as likelihood ratios, should not be used by forensic scientists to analyse data where firm statistical evidence did not exist. Unfortunately, when considering the forensic scientist's evidence, the judgment consistently commits a basic logical error, the "transposition of the conditional" This indicates that the Bayesian argument has not been understood and extends the confusion surrounding it. The judgment also fails to distinguish between the validity of the relationships in a formula and the precision of the data. We explain why the Bayesian method is the correct logical method for analyzing forensic scientific evidence, how it works and why "mathematical formulae" can be useful even where firm statistical data is lacking.

At the scene of a crime, a footwear mark is found. A suspect's shoes are examined. They are of the same brand as the footwear that left the mark and one shoe has some wear features which correspond with features in the footwear mark. What can an expert say about whether the suspect's shoe left the footwear mark at the scene?

To answer this question we have to go back to basics. The need to do so is illustrated by the judgment of the Court of Appeal in *R v T* [2010] EWCA Crim 2439

(26 October 2010) which is causing disquiet in the forensic scientific world.[1] This judgment is based on a distinction between evidence based on quantifiable data and less quantifiable data and on aversion to the use of what the Court labels a "mathematical formula" when the data is not readily quantifiable.

What can the expert say?

To illustrate what an expert can say about evidence let us use the example of paternity evidence. An expert today takes a sample and examines the DNA, but a generation ago, the blood type. There are only two things that the expert can and needs to establish. The first is the probability that the child would have the determined blood type if the defendant was its father. In some cases, the evidence would exclude the defendant, in other words, this probability would be zero. In other cases, certainty could not be obtained, and the established probability of the evidence would be larger than zero. The second thing the expert should then establish is the probability that the child would have this blood type if someone else was the father. This would depend on the frequency of the blood type in the population (of possible fathers).

What the court has to assess, consciously or otherwise, is the probability that the defendant was the father. The expert cannot assess this probability, since it requires the combination of all the evidence, not just that of the blood types. The probability that the expert can assess is different: the probability of the evidence when considering the alleged and alternative fathers. Intuitively, we can accept that if the evidence is more probable if one hypothesis is true than if the other hypothesis is true, the evidence makes the first hypothesis more probable and scientifically this turns out to be correct.

The degree to which the evidence supports one hypothesis relative to another is the strength of the evidence. A fundamental theorem demonstrates that this strength of the evidence is determined by how many times more probable the evidence is if one hypothesis is true than if the considered alternative is true.[2] This gives a single

1. "Expressing evaluative opinions: a position statement" signed by numerous leading forensic scientists in (2011) 51 Science and Justice (in press).
2. For a highly accessible explanation see: I.W. Evett "Interpretation: a personal odyssey" in: Aitken C.G.G., Stoney D.A., eds. *The use of statistics in forensic science* (New York: Ellis Horwood 1991), 9-22.

figure, which is generally known as the likelihood ratio (in paternity cases it is also known as the paternity index). It is the task of the court, not the expert witness, to combine this with the other evidence in the case.[3]

Confusion

Unfortunately, for reasons beyond the scope of this article,[4] scientists established a convention that they would give a "probability of paternity" to the courts. This probability was computed by multiplying the paternity index (the strength of the evidence) by so called "neutral" prior odds. That is, without taking the evidence into account, the two hypotheses were assumed equally probable.

This appears to apply Bayes' Theorem, which states that to arrive at posterior odds for an hypothesis, one multiplies the likelihood ratio by prior odds which are determined by all the other evidence one has considered.

The paternity evidence convention, however, takes two illegitimate steps. First, it arrogates to the expert the task of arriving at a final probability which is the task of the court. Secondly, the choice of the prior odds[5] was arbitrary when it should have been based on considering the other evidence in the case, which, again, is a matter for the court.

Once this convention was established, scientists routinely gave evidence of a "probability of paternity". This naturally confused the courts as they could not see how to combine this probability with the other evidence in the case — and indeed there is no proper way of doing so. This dilemma led courts into making puzzling statements such as:

---

3. The way this is done is explained in: Bernard Robertson and G. A. Vignaux, *Interpreting Evidence: Evaluating Forensic Science in the Courtroom*, 1st ed. (John Wiley & Sons, 1995), ch 5.
4. For the full story see Kaye, D "The probability of an ulitmate issue: the strange case of paternity testing" (1989) 1 Iowa LR 75.
5. For an example in handwriting analysis, see: N. Köller et al *Probability Conclusions in Expert Opinions on Handwriting*, 2004, Wolters Kluwer, Munich. See also: I.M. Ellman, D. Kaye, "Probabilities and proof: can HLA and blood group testing prove paternity?" (1979) 54 NYULR 1131, and A. Biedermann et al, "Equal prior probabilities: Can one do any better?" (2007) 85 For Sci Intntl 172.

> *The concept of "probability" in the legal sense is certainly different from the mathematical concept: indeed it is rare to find a situation in which these two usages co-exist, although when they do, the mathematical probability has to be taken into the assessment of probability in the legal sense and given its appropriate weight.*[6]

The forensic scientists giving this evidence were following a procedure laid down by convention; they could not have justified it on any logical grounds. They were combining (by multiplying) the judgment they were entitled to make from their knowledge, research and experience, of how much more probable the evidence was under one hypothesis than under the other, with an arbitrarily selected prior probability. By giving evidence only of a "probability of paternity" they were unwittingly concealing what they were doing. The problem identified by Ormrod LJ, in the quotation above, was not due to any inherent distinction between different kinds of probability but to the way the evidence was given.[7] This practice was not unique to paternity evidence but became the usual way in which scientific evidence was given until the advent of DNA profiling. Unfortunately, as we can see from *R v T*, courts came to expect experts to give evidence in the form of a conclusion about whether two marks had a common source.[8] As a result the courts are now confused when the evidence is given in the logical manner.

Numbers and/or verbal scales

When DNA profiling became the standard method of testing for paternity (and identity) the intense scrutiny to which it was submitted caused forensic scientists to reconsider the way they gave evidence. It has now become routine for forensic scientists to give their evidence in the way we have explained. In paternity cases, they explain how probable the DNA profile found is if the defendant was the father and how probable it is if someone else was the father and then state the likelihood ratio which is a simple matter of dividing one figure by the other.

---

6. *Re JS (a minor)* [1981] Fam 22, 29; [1980] 1 All ER 1061, per Ormrod LJ.
7. For discussion of more issues arising from a narrow view of probability on the part of many lawyers see Robertson and Vignaux "Probability-The Logic of the Law" (1993) 13 OJLS 457.
8. See "Correspondence" (1996) 36 Science & Justice 290-292 and (1997) 37 Sci & J 64-65, following D.A. Rudram "Interpretation of scientific evidence" (1996) 36 Sci & J 133-138.

There is some debate, however, as to whether the expert should explain the effect of the evidence by some verbal convention, instead of or in addition to giving the likelihood ratio. The defence can cross-examine the expert if it is desired to go into the figures, but may not always wish to do this. The expert might explain the effect of the evidence verbally in ways such as:

- "whatever odds you assess that the defendant is the father on the basis of the other evidence in the case, this evidence multiplies those odds by *n*"; or
- "this evidence moderately / strongly / very strongly supports the hypothesis that the defendant is the father".

One reason for adopting verbal conventions is a perception that jurors, and for that matter judges, do not feel confident in handling numbers and react negatively to "mathematical formulae". Another is that precise figures may not be available and the verbal conventions enable the expert to express the relative strength of the evidence in words.

In paternity cases, well-researched data in the form of blood type and DNA databases are available. *But the structure of the argument is dictated by logic not by the nature of the data.* The value of any piece of evidence lies in its power to discriminate between hypotheses: in how much more probable the evidence is if one hypothesis is true than if the other hypothesis is true.

Different impressions from the same shoe

This is equally true when the outcome of a comparison is not simply 'same' or 'different'. When marks left by the same footwear are compared, there is "within source variation", they will never be identical. For such types of evidence (comparative examinations) the outcome of the comparison is the degree of similarity.

The task for the expert, therefore, is first to compare a crime scene mark with a reference print and form an opinion as to how well the evidence (the degree of similarity) supports the hypothesis that both impressions were made by the same shoe. This may be affected by the quality of the mark at the scene and the passage of time.

It may be possible to exclude that hypothesis altogether because the basic patterns of the marks are different, or because the earlier mark has some major scar on it that is not on the second mark. In all other cases, the expert could say simply that

one cannot exclude the possibility that the two marks were made by the same shoe. This however, does not convey any evidential value. Nor does the statement that "this shoe could have made the mark" which is simply another way of saying that it cannot be excluded.[9]

Second, the expert considers the probability of the evidence if the marks were made by different shoes. This will be affected by the proportion of shoes with the observed basic pattern and wear marks. The evidential value of the comparison is defined by how much more probable the observed degree of similarity is when the marks were left by the same shoe than when the marks were left by different shoes. As the degree of similarity increases, so will the support for the hypothesis that the same shoe left both marks.

So the logical structure is the same as with DNA, and indeed with all evidence: the expert first considers how probable the evidence is if the two impressions came from the same shoe and then how probable it is if the two impressions came from different shoes.

This is always the correct structure, even if in some cases the evidence is so strong that it is not credible that there could be another shoe that would leave an impression that is just as similar. In principle, the point still is that the evidence is so strong that it overwhelms any other possible evidence in the case, not that the expert can say on the basis of the impressions alone that they were definitely made by the same shoe.

Fingerprint evidence

The difficulty in explaining this also stems from the way fingerprint evidence has classically been given. Fingerprint evidence has strong similarities to footwear evidence in that the impressions being compared will have been made under different conditions and so will not be identical even if made by the same finger. In recent research,[10] it has been argued that fingerprint evidence can be evaluated logically and can even be quantified in terms of likelihood ratios.

---

9. Pace R v T at [73] where the Court regards the statement "could have made the mark" as more precise than "moderate (scientific) support.
10. See e.g.: C Neumann et al., "Computation of Likelihood Ratios in Fingerprint Identification for Configurations of Any Number of Minutiae" (2007) 52 JFS 54.

Fingerprint examiners adopted a convention that, if the fingerprints did not exclude a suspect, they would only give evidence of identity when the degree of similarity was so great that there was virtually no possibility of another finger making a mark just as similar. Their evidence was in the form that, in their judgment the two impressions were made by the same finger, backed up by reference to the similarities. One effect was that in cases where the similarities supported common origin but were deemed insufficient, courts were deprived of that evidence.[11]

The International Association for Identification (IAI) in 1979 adopted a resolution in which it stated that "any member... who provides oral or written reports, or gives testimony of possible, probable or likely friction ridge identification shall be deemed to be engaged in conduct unbecoming such member". This resolution was rescinded in 2010 because it was "not consistent with advancement since their passage". The resolution from 2010 states that the "use of mathematically based models to assess the associative value of the evidence may provide a scientifically sound basis for supporting the examiner's opinion".[12]

Unfortunately, familiarity with the way fingerprint evidence has historically been given has contributed to the tendency of the courts to misunderstand the logical structure of forensic scientific evidence and even to misunderstand the evidence when it is given correctly. The most common error found stems from the so-called prosecutor's fallacy, which, incidentally, does not necessarily favour the prosecution.

The Prosecutor's Fallacy

This fallacy is technically the "transposition of the conditional".[13] It is illustrated by the sentence highlighted at the beginning of this article. When we are

---

11. "At the present time, in most jurisdictions, an opinion of certainty is the only acceptable opinion when dealing with the individualization of a fingerprint impression. This is not a rule that has been laid down by the courts (who will deal with any opinion offered) but rather by fingerprint examiners. There is no scientific basis for the rule. It is simply one of the aberrations that have developed within the discipline of fingerprint identification."
Harold Tuthill, *Individualization: Principles and Procedures in Criminalistics* (Lightning Powder Company, Inc., 1994).
12. International Association for Identification Resolution 2010-18, Adopted July 16, 2010 in Spokane. http://onin.com/fp/IAI_resolution_2010-18.pdf
13. W. Thompson, E. Schumann, "Interpretation of statistical evidence in criminal trials: the prosecutor's fallacy and the defense attorney's fallacy" (1987) 11 Law and Human Behaviour 167; P. Diaconis, and D. Freedman, "The persistence of cognitive illusions" (1981) 4 The Behavioral and Brain Sciences 333; I.W. Evett "Avoiding the transposed conditional" (1995) 35

dealing with something with which we are familiar, like cows, we can easily see that the two probabilities are not the same. The probability that an animal has four legs given that it is a cow is very high. Clearly this is not the same as the probability that an animal is a cow given that it has four legs.

When we are dealing with matters less intuitively familiar, however, it is easy for anyone to make this mistake. This is illustrated in *R v T* at [33]. The Court describes likelihood ratios as follows:

> *(i) The ratio of two probabilities - the probability of the evidence given that a proposition is true divided by the probability of the evidence given that the alternative proposition is true. ... In the present case it was expressed as the probability that the Nike trainers owned by the appellant had made the marks discovered at the scene divided by the probability that the Nike trainers had not made the marks.*

We can now see that the last sentence does not follow from the first, in fact it commits this error of transposing the conditional. There is a further problem with the final clause, discussed below, but leaving that for the moment, this sentence should have said:

> *In the present case, it was expressed as the probability that the marks discovered at the scene would have been observed if the Nike trainers owned by the appellant had made the marks divided by the probability that the marks would have been observed if those Nike trainers had not made the marks.*

This is only one of a number of examples of this very error throughout the judgment in *R v T*. In fact this misunderstanding appears in the very first unredacted paragraph, para [15] where the Court says:

Science & Justice 127; and B. Robertson and G.A. Vignaux *Interpreting Evidence: Evaluating Forensic Science in the Courtroom* (John Wiley and Sons (UK)) (1995) ch 6.

> *The real issue … was the use of likelihood ratios informing an evaluative opinion on the degree of likelihood that a mark had been made by a particular item of footwear.*

It is not wholly clear whose "evaluative opinion" the Court is referring to here. If it is referring to the conclusion to which the jury must come, then this statement of the issue is correct. This is not the issue with which the Court subsequently grapples, however, and the natural reading is that the Court is referring to the witness's evidence. The witness, however, did not purport to give an opinion on the likelihood that the mark had been made by a particular shoe but only as to how strongly the evidence supported that proposition.

Another important instance appears at [66] and [67]. The Court referred to the US verbal scale, which uses expressions such as "probably made", "could have made" and "probably did not make". It had previously discussed (at [31]) the FSS verbal scale which uses expressions such as "moderate support / strong support / very strong support" for the hypothesis of common origin. The Court then records the witness as saying that "the terminology used in the United States and that used in England and Wales were simply different means of expressing a verbal scale of conclusions". This is not correct. They say different things. The English and Welsh scale expresses how strongly the evidence supports the hypothesis. The US scale purports to express how probable the hypothesis is, which, as we have seen, the expert cannot assess. This error appears consistently through the judgment.[14]

To repeat, the value of an item of evidence lies in its ability to distinguish between two hypotheses, one of which, in this case, is that the suspect's shoe left the mark. Ideally, the alternative hypothesis should be precisely stated in positive rather than negative terms. Instead, in *R v T* the expert referred to a negatively-expressed alternative hypothesis that the suspect's shoe "did not make the mark". This led to trouble when the court came to consider the appropriate database, as that is determined by a positively-expressed alternative hypothesis relevant to facts of the case.

---

14. Other examples are at [30], [31], [50], [67] and [69].

At [42] and [43], the Court considered two possible databases. The witness appears to have used the Forensic Science Service database, in which the sole pattern concerned was relatively common. The Court also discussed a database consisting of frequencies of the distribution of shoes by manufacturers. Which of these would, conceptually, be the correct database depends on the alternative or defence hypothesis. Plausibly, this should be that another person left the mark and was a criminal, rather than a randomly selected member of the population. Since the shoe type concerned seems to be favoured by suspects (viz its relatively common appearance in the FSS database), this, in the present case, produced evidence more favourable to the accused than the use of a database drawn from manufacturers' figures on shoes sold to the general population.

The Court's evaluation

The Court in *R v T* commenced its evaluation by referring (at [46]) to cases in which it said that the "Bayesian approach" had been "robustly rejected" in respect of non-DNA evidence. These cases include *R v Denis Adams* [1996] 2 Cr App R 467, *R v Adams (No 2)* [1998] 1 Cr App R 377; *R v Doheny and Adams* [1997] 1 Cr App R 369. The Court does not refer to *R v Deen* CA 21 December 1993, in which Lord Taylor CJ accepted the use of likelihood ratios and held that the witness had been wrong to state a conclusion that the suspect had left the mark. The Court refers to only one article on *Adams* (at [91]).[15]

Adams was an unusual case. DNA evidence implicated the accused. Other evidence favoured the accused's case. The defence called a professor of statistics to explain in Bayesian terms how the jury could combine the favourable evidence with the DNA likelihood ratio to arrive at a combined figure for the evidence and then at a probability for the prosecution case which would be, the defence hoped, below what was required for "proof beyond reasonable doubt". The accused was convicted and appealed. The professor's evidence was not a matter raised in the appeal and was not the subject of expert evidence or argument. Nonetheless, the Court felt able to make a number of comments about the use of likelihood ratios, and the succeeding cases

15. Two of the present authors vigorously criticised *Adams* at the time, see Robertson and Vignaux "Bayes Theorem in the Court of Appeal" (1997) 69 *The Criminal Lawyer* 4, text available at: tinyurl.com/denisadams

simply picked up what was said in *Denis Adams* and repeated it. Since the relevant matters were not argued, it is respectfully submitted that *Denis Adams* is not authority, as the Court in *R v T* says it is, for the rejection of "the Bayesian approach".

One of the criticisms of *Denis Adams* is that the Court gave no guidance as to how jurors were to combine evidence if they were not to use the logical Bayesian methods. In the same way, the Court in *R v T* refers opaquely to "experience" and "judgment". These are black boxes. The witness in *R v T* multiplied together the likelihood ratios he assessed for the various aspects of the evidence from the footwear mark to arrive at a composite likelihood ratio expressing the strength of the evidence for the footwear mark as a whole. Part of what the witness seems to have done wrong, in the eyes of the Court, was to expose the reasoning that lies behind such "experience" or "judgment". In particular, at [79], the Court refers to the view that all probability assessments have some level of uncertainty and that this uncertainty is accommodated in the likelihood ratio.

The Court went on to say (at [80]):

> *We cannot agree with this in so far as it suggests that a mathematical formula can be used. An approach based on mathematical calculations is only as good as the reliability of the data used.*

The second sentence is truistic. The first sentence refers to a "mathematical formula" but, as we have seen, it is a matter of formal logic rather than mathematics. The correct way to arrive at a composite likelihood ratio is to multiply the likelihood ratios for the various items of evidence together.[16] This applies whether each likelihood ratio is based on precise or imprecise data or, for that matter, on "experience" and "judgment". Saying that the expert should not use this "mathematical formula" to assess the composite likelihood ratio is like saying that if one is just estimating by eye the area of a field, one is not allowed to multiply estimates of its width and length together. Clearly it is the correct procedure: there is no uncertainty in the relation between length, width and area, only in their values. If

16. Assuming the evidence is independent and the same hypotheses apply.

the Court were to say that the expert was not to use a logical procedure, rather than a "mathematical formula", the flaw in its reasoning would be obvious.

The Court illustrated the truth of the second sentence of [80] by showing that defence counsel had obtained other figures which actually turned out to be much more damning of the accused. The Court subsequently (at [108]) gave four reasons why the conviction was unsafe. These are commented on one at a time:

> *i) The calculations which Mr Ryder had used in reaching his view were not before the jury; those figures were more favourable to the appellant than the figures put before the jury. If evidence of the full figures had been put before the jury then, applying the test in Pendleton [2001] 1WLR 72, [2002] 1 Cr App R 441, it might reasonably have affected the decision of the jury to convict.*

What happened (see [42]-[44]), was that the witness testified that the evidence gave moderate support to the hypothesis that the accused's shoe had made the mark. Defence counsel then cross-examined on the basis of another set of figures far more unfavourable to the accused. Prosecuting counsel decided not to re-examine as the defence had not undermined the witness's evidence but had strengthened it. Had any re-examination taken place, the witness would, in addition to giving his own original figures, have been bound to accept that the figures put to him had been reasonable and that he had deliberately erred on the side of caution. It is hard to see that this would undermine the evidence in the eyes of the jury.

> *ii) The process by which the evidence was adduced lacked transparency. This is no personal criticism of Mr Ryder, as he was simply following practice. However, it is simply wrong in principle for an expert to fail to set out the way in which he has reached his conclusion in his report.*

We would agree that expert witnesses should include their methodology in their reports. It seems, however, that the Court would have been content with opaque references to "experience" and "judgment".

> *iii) In the light of the strong criticism by this court in the 1990s of using Bayes theorem before the jury in cases where there was no reliable statistical evidence, the practice of using a Bayesian approach and likelihood ratios to formulate opinions placed before a jury without that process being disclosed and debated in court is contrary to principles of open justice.*

We have questioned the validity of the criticism of the Bayesian approach in the 1990s cases. Whatever good reasons there might be for not expressly urging the jury to apply Bayes' Theorem in their deliberations, they do not apply to the process by which a scientist arrives at a conclusion.

At [90], the Court said that "outside the field of DNA (and possibly other areas where there is a firm statistical base) ... Bayes theorem and likelihood ratios should not be used". Here, however, the Court states that it was the use of likelihood ratios without that approach being revealed in open court that was contrary to the principles of open justice. If that was the basis on which the conviction was quashed, then the prohibition on Bayesian reasoning at [90] is obiter, however firmly it may have been expressed.

> *iv) The practice of using likelihood ratios was justified as producing "balance, logic, robustness and transparency", as we have set out at paragraph 54. In our view, their use in this case was plainly not transparent. Although it was Mr Ryder's evidence (which we accept), that he arrived at his opinion through experience, it would be difficult to see how an opinion of footwear marks arrived at through the application of a formula could be described as "logical", or "balanced" or "robust", when the data are as uncertain as we have set out and could produce such different results.*

It is surely a feature of a "logical, balanced or robust" approach that if one applies it to different figures it will produce different results. While in this case the process may not have been made transparent, transparency is indeed one of the advantages of the use of likelihood ratios as it exposes all the factors considered, unlike a reference to "judgment" or "experience".

Discussion

This evidence could have been better given. The alternative hypothesis should have been positively expressed, either that some unknown shoe had made the mark (in which case the appropriate database would have been the manufacturers' data) or that the shoe of some unknown burglar made the mark, in which case the crime-related Forensic Science Service database would be appropriate. We would agree that the expert should make the reasoning clear in the reports, but there is no reason why, in examination-in-chief, the expert should go further than expressing the strength of the evidence either in numerical or verbal terms. If the defence chooses, it can then cross-examine on the figures and reasoning process reported. It is unclear however why the defence should choose to shoot itself in the foot by cross-examining on the proposition that figures less favourable to the accused could have been used.

At least two points can be derived from the judgment:

(i) evidence can be given evaluating the significance of footwear mark evidence; but

(ii) "mathematical formulae" are not to be used in reasoning to an evaluation of that significance, it must be based only on "judgment" and "experience".

As to the first, it is made clear earlier in the judgment that the Court failed to recognise the difference between a statement that "the evidence strongly supports the hypothesis that this shoe made this mark" and "my conclusion is that this shoe probably made this mark". The first is a statement of the strength of the evidence, which is within the province of the expert. The second is a conclusion which can only validly be made by also assessing the strength of all the other evidence in the case, which is the province of the jury. The Court purports to be concerned to keep experts

within appropriate bounds, but then seems to fail to realise that it is inviting experts to draw conclusions they have no basis for drawing.

As for the second point, at [90], the Court appears to lay down a prohibition on the use of likelihood ratios and Bayes' Theorem in assessing evidence outside DNA "(and possibly other areas where there is a firm statistical base)". One of the Court's justifications for this is alleged division of opinion in other jurisdictions over the use of Bayesian techniques. Since then, the Board of the European Network of Forensic Science Institutes covering 33 countries has signed the "Position Statement" and "engages itself to work towards a full implementation [of the Bayesian approach] within the ENFSI laboratories".[17]

The Court does not explain why the logical structure of an argument should differ according to how quantifiable the evidence is and there is no reason why it should. The Court also assumes that there is a clear distinction between cases where a "firm statistical base" is available and where it is not. All probability assessments involve uncertain and incomplete data; conversely there are several areas, such as glass and fibres, where the statistical data available is better than for footwear marks but not so complete as for DNA.

The judgment fails to explain how reasoned judgments are to be arrived at when several factors are involved, other than by appeals to "experience" and "judgment" – black boxes the Court would apparently prefer to remain unanalysed. Weirdly, the Court then proceeds to be prescriptive about how experts are not to arrive at such judgments. The Court is not even referring to how evidence is to be given in court or to how juries are to be directed but to the reasoning process to be followed by the expert in arriving at a conclusion. It seems strange that a court should dictate to scientists how they carry out their science and by what reasoning process they are to arrive at their conclusions. The danger of doing so is illustrated by this very judgment in which the Court has proceeded from several demonstrably false premises. Indeed, it is quite plausible that the reason that the expert refrained from exposing his reasoning was the prior decisions of the Court of Appeal on "the Bayesian approach". If these cases are to be cited as authority then the detailed criticisms of them referred to above must be discussed.

---

17. Position Statement above

Conclusion

The judgment makes two fundamental errors. The first is consistent transposition of the conditional which indicates that the Court has not understood the difference between assessments of the probability of a proposition and of the strength of the evidence for the proposition; the second is confusion between uncertainty in the values of the variables and uncertainty in the their relationship in a mathematical formula. The fact that variables cannot be precisely expressed does not affect the validity of the relationships described by the formula.

Where does this leave trial judges? The New South Wales Court of Appeal in *R v Shepherd* (1988) 39 A Crim R 266 faced a dilemma in that it was supposedly bound by a decision of the High Court of Australia in *R v Chamberlain (No 2)* (1984) 153 CLR 521 which laid down strange rules of reasoning. Roden J responded:

> *If the proposition's validity depends upon logic, and the resolution of a philosophical debate, it seems to me that it cannot properly be characterised as a rule of law. Its validity certainly does not stem from any statutory provision; nor, in my view, can it be argued that its assertion as a proposition of logic in otherwise binding judgments gives it the force of law. If it were shown to be not logically valid, then it could not apply. That would not be the position with a rule of law.*

"[C]onclusions in relation to the manner in which evaluative evidence should be approached by footwear examiners" [16] cannot possibly be said to be rules of law. They are instructions to scientists as to how to think about scientific matters. Such instructions need to conform with the requirements of logic and the Court's conclusions in *R v T* are not logically valid. This judgment, if allowed value as a precedent, has the potential to halt progress in forensic science and to reduce forensic scientific evidence to unanalysable impressionistic conclusory statements by experts.