



Commentary on the case of R v T on footwear mark evidence and the use of likelihood ratios.

Introduction

This Appeal case involved,

"The extent to which evaluative expert evidence on footwear marks is reliable and the way in which it was put before the jury. ...

The appeal raised an issue of some importance in relation to the use of likelihood ratios in the provision of an evaluative opinion where the statistical data available were uncertain and incomplete."

It has important repercussions for a number of issues including policy and practice in UK forensic science.

We (TFI) had examined all of the evidence in this case, including the case files created in the assessment and analysis of the footwear marks. The Appellant's case was presented by Mr James Wood QC who had extensive discussions with Professor Jamieson. The served scientific statements were prepared by Professor Jamieson who did not give evidence at the Appeal; only evidence from the Crown was actually heard.

The general conclusions expressed by Professor Jamieson were,

"There is no clear basis for the strength of evidence derived by Mr Ryder, its reliability, nor for the expertise on which it rests."

"I do not doubt that it is possible that such comparisons can provide useful evidence. I am not disputing Mr Ryder's opinion, but the scientific basis of it. It is my opinion that the state of development of this expertise is insufficient to ascribe any more than a very rough approximation to the probative value of the evidence, and such opinions cannot be considered scientific."

Case files and an aside on expertise

The Court noted,

"After the trial, the papers of Mr Ryder were reviewed by Professor Jamieson of the Forensic Institute in relation to all the expert evidence given at trial. Although Professor Jamieson has no expertise whatsoever in footwear mark comparison, he noted that on one of the working papers there were the words "Interp/conclusion" "mod evidence" with the formula and values which we have set out at paragraph 35. He described this approach as "the Bayesian approach" of using likelihood ratios. He correctly commented that this had not been explored in the course of the trial."



The identification of these notes was only possible by close examination of the case files prepared by the expert. This has been crucial in many of our cases, including this one. The Forensic Science Regulator has provided guidance regarding what should be recorded by the scientist. Some experts consider a brief visit to the lab and a 'flick-through' of all or part of the case file to be adequate. We do not.

This is not the first case that we have considered the general reliability of the science, as opposed to the specific findings of the expert. This is in contrast to the traditional 'fight fire with fire' approach where solicitors generally seek a similarly qualified expert from the same field as the Crown's witness. As stated by Professor Jamieson,

"My qualifications are as a scientist with considerable expertise in science and in the evaluation of scientific evidence; it is these to which my qualifications and experience were addressed. It is unnecessary, and probably desirable, that I am not and do not claim to be a footwear expert in assessing the scientific value of such evidence. I would feel equally comfortable assessing claims for the validity of astrology or psychic phenomena without necessarily being a practitioner."

We have published an article elsewhere on the issue of experience versus expertise.

Use of Likelihood Ratios and evaluative opinion

A likelihood ratio (LR) is a method of comparing the probability of two things by simply dividing one by the other. If a horse is 10-1 and another is 50-1 then the LR is 5 ($=50/10$) that the outcome will be the former horse will win rather than the latter, and $1/5$ ($=10/50=0.2$) that the outcome will be the latter will win rather than the former. Note that the LR compares only those two in this instance. A LR of greater than 1 (>1) favours the 'top line' (numerator) outcome, whereas a LR less than 1 (<1) favours the 'bottom line' (denominator) outcome. It is simply a means of measuring how much more likely one thing is compared to another.

The use of the LR in the evaluation of forensic evidence has been promoted by some statisticians and groups of scientists, especially in the UK where the Forensic Science Regulator, in a submission to the Court, supported the approach. They contend that this is the fair and balanced way to look at the evidence; by comparing the probability of the evidence given the prosecution story or outcome (Pp) with the probability of the evidence given the defence story or outcome (Pd). The LR is then Pp/Pd .

Professor Jamieson stated in this case,

"This approach, termed the Bayesian approach using likelihood ratios (LR), should not be seen as a standard having universal acceptance nor fully explored as yet by courts."



It is unnecessary here to consider the scientific argument regarding the pros and cons of the LR. The argument advanced by the defence was mainly that there were no data to support any such calculation in this case. Professor Jamieson stated;

"The footwear evidence was described as providing only moderate support for the view that the shoes belonging to T made the footwear marks [exhibit number]. There is no reliable statistical support for this assertion, neither as a scientific concept nor in terms of the data used to derive the opinion. ...

There is no scientific basis for assessing the acceptable range of wear for shoes; a factor that must be considered when the marks have clear differences from the shoes claimed to have made them. The inability to exclude shoes as potential sources of the marks reduces the probative value of the evidence. It is opaque how, under such circumstances, any pair of [manufacturer] shoes could be excluded as potential sources. That being the case, the marks could have been made by any [manufacturer] shoe.

It is essential for the population data for these shoes be applicable to the population potentially present at the scene. Regional, time, and cultural differences all affect the frequency of particular footwear in a relevant population. That data was simply not available to Mr Ryder in performing his assessment. If the shoes were more common in such a population then the probative value is lessened. The converse is also true, but we do not know which is the accurate position.

There is no scientific evidence that footwear comparison, as a form of expertise, has any scientific basis in terms of providing reliable matches between recovered marks and footwear recovered months after the marks were made." [words in square brackets are added by us to replace redacted information]

The Court states,

"An approach based on mathematical calculations is only as good as the reliability of the data used. ...

It is evident from the way in which Mr Ryder identified the figures to be used in the formula for pattern and size that none has any degree of precision. ...

More importantly, the purchase and use footwear is also subject to numerous other factors such as fashion, counterfeiting, distribution, local availability and the length of time footwear is kept. ... There is no way in which the effect of these factors has presently been statistically measured ...

we have concluded that there is not a sufficiently reliable basis for an expert to be able to express an opinion based on the use of a mathematical formula. ... We are satisfied that in the area of footwear evidence, no attempt can realistically be made in the generality of cases to use a formula to calculate the probabilities. The practice has no sound basis."

It is important to understand that in our opinion the Court was not prohibiting the use of an LR. The issues in the case were twofold; first, that it had never been made clear to the trial court that such an approach had been used and, second, that there was insufficient data to support any numerical calculation.



The Court went on to make clear that an expert was able to form an evaluative opinion even without statistics, but that it should not be presented as a mathematical calculation such as a LR.

“However there are cases where it would not be right to confine an examiner (where there are solely class characteristics) to opining on whether the mark could or could not have been made. There may be factors that enable him to go further than “could have made” and express, on the basis of such factors, a more definite evaluative opinion. It would not be appropriate for us to express a view on the factors which would properly enable an examiner to express a more definitive evaluative opinion, but they would certainly include an unusual size or pattern. ...

In our judgment, an expert footwear mark examiner can therefore in appropriate cases use his experience to express a more definite evaluative opinion where the conclusion is that the mark “could have been made” by the footwear. However no likelihood ratios or other mathematical formula should be used in reaching that judgement for the reasons we have given.” [underline added]

It appears to us that the reasons were concerned with the lack of reliable data; there was no blanket prohibition on the use of a LR if such data are available. This is explicit in the judgement;

“If there are reliable statistics and data, it would then be necessary to consider how likelihood ratios should be used and how their use should be explained to a jury.”

We can therefore identify evaluative opinion as a global category in which the expert expresses their opinion of the strength of the evidence. Evaluative opinion may be;

1. Comparative, where the expert compares propositions using either
 - a. data, in which case use of a LR is appropriate (e.g. DNA)
 - b. experience, in which case some other term must be used, such as ‘comparative evaluation’ (e.g. footwear mark) (with some conveyance of the significance of the findings)
2. Absolute, where the expert assesses the weight of only one proposition using, again, either
 - a. data (e.g. frequency of glass type)
 - b. experience (e.g. some clinical diagnoses)

For example, if we wish to assess the chance of getting a 5 on the throw of a die then the calculation is based on the (assumed) data that the probability is 1 in 6. This is an example of an absolute (but nevertheless still probabilistic) opinion based on data (2a). If we wish to assess the probability of throwing a 5 as opposed to a 2 or 3 then this is a comparative assessment using data; a LR (1a). The LR is $(1/6)/(1/6) + (1/6) = (1/6)/(2/6) = 1/2$, which is a LR of 0.5; it is half as likely.



We are not here arguing for or against any of these, but attempting to clarify the different types of assessing the weight of evidence.

In regard to type 1b it may be useful to refer to the case of *Reed & Reed* where the Court endorsed the admissibility of evaluative opinion based on the experience of the expert. However, an important component of that judgement was that such opinion, based on experience, was admissible, "when there is a sufficient evidential basis". Whether that basis yet exists has not been established in many areas of forensic practice.

The Court in *R v T*, referring to *R v Reed & Reed* and *R v Weller* state,

"In neither case was there any question of a statistical basis or the use of a likelihood ratio."

In both of those cases the prosecution scientists had used a comparative evaluation using experience (1b) as the justification.

Accreditation, validation, and responsibility

An expert report is normally regarded as the true and complete opinion of the expert; many jurisdictions insist on a statement to that effect to be included in any written expert opinion. It is explicit that the expert's ultimate duty is to the Court. The Forensic Institute policy is that the expert provides only their own opinion, which may be derived after discussion with colleagues, as they will have to defend that opinion. It comes as some surprise to note the Appeal Court in *R v T* stating,

"Mr Ryder's evidence was that he was trained to use this approach and was following practice within the FSS. He cannot be criticised for doing so."

The question that must arise is where responsibility lies if the approach used was wrong, and contrary to that which the expert would use if they were left to use whatever approach they favoured.

Accreditation is a means by which processes within organisations are codified and checked to ensure a consistent quality of service to clients. It is frequently proffered in statements and in Courts as evidence of the quality of opinion. In fact, despite the FSS Ltd being an ISO17025 accredited organisation the Court records,

"It is important to note, however, that, on the evidence we received, not all examiners within the FSS use the approach; some simply use their experience and have scant, if any, regard to databases."

It would appear that the processes are not as consistent as they should be.



If the Courts expect experts to adhere to corporate policy, then there should be no need for experts within such companies to sign reports as, in a similar manner to reports from Public Analysts, the essential opinion is emanating not from the expert but the organisation.

If that is the case, then there is no scope for experts to depart from accredited procedures as they would not then represent the corporate view.

Regulation and quality standards

In response to identified and well-publicised problems in forensic science, the Government created the post of the Forensic Science Regulator. The Court notes,

"He [the Forensic Science Regulator] has been entrusted by the Home Office with ensuring that appropriate quality standards are developed, implemented and used effectively in criminal justice."

Despite that role, and the advice given direct to the Court by the Regulator, the Court goes on to say,

"We do not agree with the observations of the Regulator that a similar approach is justified in all areas of forensic expertise."

This would suggest that there is a conflict between the Court and the Regulator as to what constitutes 'quality' scientific evidence. It is opaque how the system being implemented by the Government can retain credibility when its advice to the very client that it was set up to serve has been rejected at such an early stage.

Of course, it may not be the principle of a Regulator that is flawed, but the practice. The Court notes,

"Considerable importance appears to have been attached to this paper within the community of providers of forensic services [within the UK]. Despite inquiries made by us, it is not clear to what extent, if any, it was subject to wider debate outside the forensic science community." [added]

The restricted set of advisers to the Regulator, and lack of external scientific input to forensic science generally, has already been highlighted by The Forensic Institute. Professor Allan Jamieson, in a letter to the journal *Nature*, stated,

"The UK response to the documented and public failures in forensic science has been to appoint an 'independent' regulator, Mr Andrew Rennison. The regulator, currently an ex-policeman funded by the Home Office, chairs an advisory council whose scientific input comes from within the forensic community and from the suppliers of services to the police. ...

The introspective and isolated position of forensic science within the United Kingdom is further shown by its removal from the science, engineering and manufacturing Sector Skills Council (SSC) and its placement within the Skills for Justice SSC, where it is the only 'scientific' component, thus removing an opportunity for external scientific scrutiny."



The lack of such scrutiny may in part be directly responsible for the conflict of opinion between the Appeal Court and the Forensic Science Regulator.

Scientific opinion

Professor Jamieson states;

"However, I am concerned that, at page 16, Mr Ryder states, "In my view the scientific findings in this case are somewhat unlikely that a person arrested on suspicion of involvement in this incident would coincidentally possess footwear that would correspond to this extent, hence my conclusion that there is moderate degree of scientific evidence to support the view that the footwear FFF from PLACE attributed to T has made footwear marks within PLACE.

In my opinion, it is entirely unsatisfactory for a scientist to do other than consider the probability that such a match would be found by chance; and perhaps compare the likelihood of another shoe possessing similar characteristics (if, and only if, relevant data was available)."

The judgement states,

"It is essential, if the expert examiner of footwear expresses a view which goes beyond saying that the footwear could or could not have made the mark, that the report makes clear that this is a view which is subjective and based on his experience. For that reason we do not consider that the word "scientific" should be used, as, if that phrase is put before the jury, it is likely to give an impression to the jury of a degree of precision and objectivity that is not present given the current state of this area of expertise."

We suspect that this opinion will affect other areas of expertise currently wearing a scientific cloak.

One feature of science is the practice of publishing and sharing data to enable other scientists to scrutinise and either verify or challenge conclusions from the data; that should be a requirement even more so in a forensic environment. We continue to campaign for full disclosure of scientific data. The Appeal Court appears to have a similar concern;

"There is also the further difficulty, even if it could be used for this purpose, that the data are the property of the FSS and are not routinely available to all examiners. It is only available in a particular case to an examiner appointed to consider the report of an FSS examiner."

Perhaps part of any quality standard should be a requirement to fully disclose any data used to underpin a forensic scientific opinion.

Conclusion

There is likely to be concern among some scientists that the decision is an outright rejection of Bayes or LR's in everything but DNA profiling. We do not think that this is so. Our interpretation of the judgement is that;

- The basis of any evaluative opinion must be made clear to the Court



- Use of terminology that suggests a degree of scientific justification for the opinion can only be used when there is such scientific justification
- The judgement also highlights, or alludes to, a number of problems that can be readily solved with the correct adjustments to the current system.
 - Wider consultation is necessary to inform good scientific practice
 - Claims of a procedure to be scientific must be justified
 - Part of that justification includes enabling full access to any data or other information that underpins the opinion

END

NOT TO BE SERVED