



Neutral Citation Number: [2010] EWCA Crim 2439

Case No: 2007/03644/D2

IN THE HIGH COURT OF JUSTICE
COURT OF APPEAL (CRIMINAL DIVISION)
ON APPEAL FROM THE CROWN COURT

Royal Courts of Justice
Strand, London, WC2A 2LL

Date: 26/10/2010

Before :

LORD JUSTICE THOMAS
MR JUSTICE BEATSON
and
MR JUSTICE KITCHIN

Between :

Regina
- and -

Respondent

Appellant

Mr J A D Wood QC and Mr R Thomas for the Appellant
Mr P Wright QC and Mr A Nuttall for the Respondent

Hearing dates: 28 February, 11 and 12 May and 7 July 2010

**Redacted Judgment Approved by the court
for handing down
(subject to editorial corrections)**

**If this Judgment has been emailed to you it is to be treated as 'read-only'.
You should send any suggested amendments as a separate Word document.**

Lord Justice Thomas:

1.

[REDACTED] The appellant was tried for that murder
[REDACTED] The sole issue was identification. The appellant was
convicted [REDACTED]
[REDACTED] He appeals with the leave of the Single Judge on
three main issues:

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

[REDACTED]
[REDACTED]

The appeal [REDACTED] 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.

[REDACTED]

The factual evidence

[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

The course of the appeal

10. After his conviction, trial counsel advised that there were no grounds for appeal. The appellants contended that he had not been properly represented at trial and sought the advice of new counsel. The Registrar granted a Representation Order for further counsel and new solicitors. Mr Wood QC and Mr Thomas were instructed as were new solicitors in support.

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

I THE EXPERT EVIDENCE ON FOOTWEAR MARKS

15. The real issue in relation to footwear marks, as it emerged in the course of the appeal, was the use of likelihood ratios in forming an evaluative opinion on the degree of likelihood that a mark had been made by a particular item of footwear. This was not an issue at the trial [REDACTED].
16. As the way in which the issue emerged is of some importance and the use of likelihood ratios requires analysis.

[REDACTED] We shall first set out the way the evidence on the footmarks was gathered, [REDACTED]
[REDACTED]

- ii) We will then describe the emergence of the use of likelihood ratios in providing evaluative evidence in relation to footwear marks in England and Wales and make reference to the practices elsewhere (paragraphs 48-67).

- iii) We will then set out our conclusions in relation to the manner in which evaluative evidence should be approached by footwear examiners (paragraphs 69-99).
- iv) We finally set out our conclusions in relation to the footwear marks evidence in the appeal (paragraphs 100-109).

(1) The evidence and the course of the trial and the appeal

(a) The footwear marks found at the scene of the murder

17. Mr Ryder of the FSS attended at the scene of the murder. He was by academic training a Chemist and had been employed by the FSS since 1989. He had very extensive experience of footwear marks.

He attended at the scene of the murder and the evidence of the footwear marks at the scene was recovered by him.

19. He subsequently carried out a comparison with pieces of footwear with which he was provided; these pieces included footwear from the trainers found in the appellant's house after his arrest.

(c) The specific matters covered in Mr Ryder's reports

21. Mr Ryder carried out an analysis of the footwear mark evidence which was recorded in detailed records and notes. His conclusions were set out in three reports served prior to the trial. In his first report he summarised his general approach to the assessment of the significance of a footwear mark:

“When assessing the significance of any correspondence or difference between a shoe and a mark resulting from a comparison, the likelihood of obtaining that correspondence or difference is considered against two alternative propositions:

- 1. The shoe in question has made the mark it has been compared with
- 2. the shoe has not made the mark

When the results of an examination are not conclusive, the strength of assessment of the scientific evidence is normally expressed as a level of support for one or the other of the above two propositions. In attempting to convey the level of this support, the most appropriate expression is selected from among the following progressive verbal scale: very limited, limited, moderate, moderately strong, strong, very strong, extremely strong.”

This approach and way in which the verbal scale related to likelihood ratios was at the heart of the appeal.

22. His examination revealed:

- i) The type of sole pattern and configuration of the pattern (and hence size) of the trainers were the same as the marks made at the scene of the murder. The pattern type was one of the most frequently encountered.
- ii) The wear on the trainers recovered was greater than that on the trainers that had made the marks. His view was that that [REDACTED] additional wear could well be explained by use in the intervening period. However he could not be definite about this, as the amount of wear that occurred would depend upon the use of the trainers.
- iii) Some of the marks on the floor had particular features which might have resulted from damage to the trainers; the trainers recovered from the appellant showed no such damage. There were three explanations for the features which would not exclude the possibility that the marks had been made by the trainers recovered:
 - a) The marks were the result of an artefact on the surface of the floor.
 - b) As the trainers were worn for a [REDACTED] period [REDACTED], the particular features could have worn away.
 - c) The marks could have been caused by a small stone which had subsequently been dislodged.

[REDACTED]

24. He expressed his overall conclusion:

“Overall it is my view somewhat unlikely that the observed correspondence would have been obtained as a result of mere coincidence had the recovered footwear not made the marks in question.

... there is at this stage a moderate degree of scientific evidence to support the view that the [Nike trainers recovered from the appellant] had made the footwear marks..”

Mr Ryder’s other reports expressed similar conclusions. They were checked by a second examiner in accordance with FSS practice.

25. The reports contained no statistical information or reference to use of a likelihood ratio or the formula used in calculating it, though as we explain at paragraph 35, statistics, likelihood ratios and a formula were in fact used. There was cross examination on other statistics at trial, as we explain at paragraph 43.

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

- (e) *The underlying approach of [REDACTED] experts to the significance of the evidence*

[REDACTED]

29. The evidence before us was that experts on footwear marks in England and Wales arrive at their conclusions by an examination of the marks and the footwear by reference to four factors:

- i) The sole pattern. This is referred to in US practice, to which we refer at paragraphs 64 and following, as a characteristic of the footwear. There are many different sole patterns; there are over 1200 patterns of Nike trainers.
- ii) The size. This is again referred to in the United States as a characteristic of the footwear.
- iii) The wear.

- iv) Any damage.
30. Over the years, footwear experts in England and Wales used their skill to determine whether the marks had (i) been caused by a particular shoe, or (ii) definitely had not been made by a particular shoe, or (iii) could have been made by a particular shoe. At a time, which remains unclear to us, they developed in the UK a verbal scale for expressing views as to the degree of likelihood of the third of the three possibilities – whether the marks “could have been made by a particular shoe” – no support, limited support, moderate support, moderately strong, strong, very strong and extremely strong. Sometimes the word “scientific” was placed before the word “support”; Mr Ryder did so in his report in this case.

(f) *The use of likelihood ratios by Mr Ryder to express his conclusion*

31. As we will explain in more detail at paragraph 49 below, a few years before the trial, it was considered by some providers of forensic services that they should try to achieve a greater degree of consistency in the way in which individual examiners of footwear expressed their conclusions where the mark “could have been made by a particular shoe”. It was considered that the use of experience alone and expression of the conclusion on the verbal scale was insufficient. It would be better to use available data to calculate a likelihood ratio to reflect the degree of likelihood that the marks were made by a particular shoe and to provide a numerical scale equivalent to the verbal scale:

<i>Value of likelihood ratio</i>	<i>Verbal equivalent</i>
>1-10	Weak or limited support
10-100	Moderate support
100-1000	Moderately strong support
1000-10,000	Strong support
10,000-1,000,000	Very strong support
>1,000,000	Extremely strong support

32. It is clear that likelihood ratios have been used in other areas of expertise by forensic experts when expressing their conclusions. We are solely concerned in this appeal with the use in relation to footwear mark evidence.
33. The likelihood ratio, as used, was explained to us in a number of different formulations; two will suffice.
- 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. This was the definition given to us by Mr Ryder and is derived in these exact terms from a paper by the Association of Forensic Science Providers published in 2009 to which we refer at paragraph 52 below. 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.
- ii) A measure of how likely it is to obtain a piece of evidence given a proposition, compared to how likely it is to obtain the same piece of evidence given an

alternative proposition. Once a numerical value is calculated, the strength of support for a proposition can be expressed.

34. It is clear from the evidence that emerged during the appeal that in expressing the view that there was “moderate scientific support” that Mr Ryder used not only his own experience of footwear marks to express his opinion on the verbal scale we have set out in paragraph 21, but also used the numerical scale of likelihood ratios in the left hand column of the table in paragraph 31. He then calculated an overall likelihood ratio. In doing so, he followed the practice which was being adopted.
35. In this particular case he followed that methodology by the use of a formula based on each of the four factors to which we have referred at paragraph 29. He ascribed to each a likelihood ratio; the likelihood ratios were then multiplied together to provide an overall likelihood ratio that could then be applied to the scale set out at paragraph 31. His methodology was summarised in the formula set out in a contemporaneous note of 19 December 2006.

$P \times C \times W \times D$ with values then ascribed of $5 \times 10 \times 2 \times <1 = \sim 100$

36. He explained the formula and the values ascribed in a statement made in June 2010 and in his oral evidence in July 2010.
- i) **P** represented frequency of pattern: Mr Ryder used the internal database of the FSS to examine the frequency of pattern. This recorded the number of shoes received by the FSS (in contradistinction to the number distributed within the UK, figures to which we shall refer at paragraph 42). The FSS database comprised approximately 0.00006% of all shoes sold in a year. The pattern was, according to FSS database, one of those most frequently encountered – 20% were of this pattern in 2005-2007, though that figure did not reflect the frequency of use in the population at large, as we explain at paragraph [103.ii] below. The likelihood based on the FSS database figure of 20% for factor **P** was 1/0.2. As a matter of arithmetic this was 5.
 - ii) **C** represented frequency of size; he considered that the size was 11. The database of the Shoes and Allied Trade Research Association was used. This indicated that approximately 3% of the shoes of the population would have size 11 footwear. The likelihood ratio for this factor would therefore on this basis be 1/0.333. He moderated this figure by allowing for manufacturing practices that might occur which resulted in sole units being slightly smaller or larger than the trainer in question, the effect of additional wear and the fact that they might be counterfeit. He considered on this basis that he could exclude 90% of footwear. He therefore substituted the more conservative figure of 0.1 for 0.333; he added in his oral evidence to us that this was conservative because he might have been able to exclude some size 11’s on the basis of mould configuration. Thus the likelihood ratio for factor **C** was 1/0.1. As a matter of arithmetic this was 10 and this was used in the formula.
 - iii) **W** represented amount of wear. Mr Ryder considered that the wear on the trainers meant that he could exclude half of the trainers of this pattern type and approximate size/configuration. He therefore calculated the likelihood ratio for

W as $1/0.5$. As a matter of arithmetic this was 2 and so in the formula W was 2.

- iv) D represented amount of damage. He considered that he could exclude very few pairs of shoes that could not previously have been excluded by the other factors. The figure would have been close to 1 and he therefore calculated the likelihood ratio for D as $1/1$. He used <1 in the formula to indicate that the damage to the shoes made it less likely that they had made the marks thereby reducing the likelihood ratio.

37. Multiplying the likelihood ratios for each of the four factors together produced an overall likelihood ratio of $1/0.01$ or an overall ratio of 100. He expressed this as “~100” in his notes of 19 December 2006.

(g) *The omission of these calculations from the expert reports*

38. In Mr Ryder’s reports for the trial to which we have referred at paragraph 21, there was no reference at all to any of these statistics, the formula or to the use of a likelihood ratio or to the scale of numerical values set out in paragraph 31. The conclusion in his first report, which was supported by the statistics, formula and resulting likelihood ratio, was expressed solely in terms of the verbal scale which we have set out at paragraph 21; this was dated one day after the notes in which he had recorded his calculation. Mr Ryder’s explanation for the omission was that it was not standard practice for the detail relating to statistics and likelihood ratios to be included in a report. He made clear that the data were not available to an exact and precise level and it was only used to confirm an opinion substantially based on his experience and so that it could be expressed in a standardised form.

[REDACTED]

Although none of the counsel concerned had any actual recollection, there is nothing to suggest that [REDACTED] expert made known to any of the lawyers instructed at the trial the use of likelihood ratios or the figures used in calculations which we have described. On the contrary, all the evidence before us points to none being made known. Nothing was said in evidence about them. [REDACTED]

(h) *The evidence at trial*

41. In his evidence in chief Mr Ryder explained the methodology of examining the trainers and the marks and gave detailed evidence about the marks found at the scene of the murder. His conclusion to the jury was the same as in his report:

“In my opinion there is a moderate degree of scientific support for the view that [the Nike trainers] made those marks”

42. When he was cross-examined, he was asked about figures in the UK over 7-8 years for the distribution of Nike trainers of the same model as that found at the appellant's house; some figures had been supplied to him by the defence lawyers the day before. Mr Ryder gave evidence that there were 1,200 different sole patterns of Nike trainers; the pattern of Nike trainers that made the marks on the floor was encountered frequently and had been available since 1995; distribution figures for the pattern were only available from 1999. In the period 1996-2006 there would have been 786,000 pairs of trainers distributed by Nike. On those figures some 3% were size 11. The pattern could also have been made by shoes distributed by Foot Locker and counterfeits of Nike shoes for which there were no figures. In answer to the suggestion that the pattern on the Nike trainers found at the appellant's house was of a common type, he said:

“It is just one example of the vast number of different shoes that are available and to put the figures into context, there are around 42 million pairs of shoes sold every year so if you put that back over the previous 7 or 8 years, sports shoes alone, that multiplies up to nearly 300 million pairs of sports shoes so that particular number of shoes, produced which is a million, based on round numbers, is a very small proportion.”

He made clear that the pattern was the one that was encountered most frequently in the laboratory, but he did not give the actual figures used by him as we have set out at paragraph 36.i). His explanation to us for not mentioning them was that that figures represented what came into the FSS laboratories and not for the population as a whole, even though the figures in the database which he used in his formula were more favourable to the appellant.

43. Apart from detailed questions in relation to the specific features of the marks and wear, Mr Ryder was also asked about his conclusion which we have set out at paragraph 24 and the verbal scale to which we have referred at paragraph 30. He was asked if it was possible that another pair of shoes could have left the marks and that possibility could not be excluded; his response was:

“That is a possibility but consideration has to be given to the likelihood of a pair of shoes of this pattern, and of this particular size and with the appropriate degree of wear.”

When asked if the evidence in relation to footwear stood alone, all it would show was there were shoes marks which provided moderate scientific support that they could have been made by the trainers recovered some 15 weeks later from the appellant's home, he responded: “If it stood alone, yes”.

(k) *The summing up*

44. The judge in the course of summing up the evidence of Mr Ryder, emphasised the need for the jury to consider the correspondence between the prints and the trainers in terms of pattern type, size and degree of wear, emphasising in respect of wear that it could not be assessed precisely because it depended on the amount of time the trainers had been worn. He reminded them of Mr Ryder's conclusion of “a moderate degree of scientific evidence to support” for the view that the marks had been made by the trainers recovered from the appellant's home. Finally he returned to the figures

which Mr Ryder had given in the course of cross examination as we have set out at paragraph 42, concluding:

“To put the whole thing in context Mr Ryder pointed out that 42 million pairs of sports shoes were sold every year which would, I suppose, if one were to take 1999 to 2006 inclusive, be up to and not far short of three hundred million shoes. So in proportion terms these shoes are really a small proportion of the total sold, a quarter of a percent all of shoes of all adult sizes. Bear in mind however that 786,000 is still a lot of trainers and that you have to allow for the Foot Locker and fake shoes as well.”

(1) *The development of the appellant’s appeal in relation to the use of likelihood ratios*

45. 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.

46. In response, Mr Ryder produced numerous statements in the course of which it became evident to the appellant’s counsel that likelihood ratios had been used. It was submitted to the court the approach adopted was a Bayesian analysis which this Court had robustly rejected for non DNA evidence in a number of cases: *R v Dennis Adams* [1996] 2 Cr App R 467, *R v Adams (No 2)* [1998] 1 Cr App R 377; *R v Doheny* [1997] 1 Cr App R 369.

47. After enquiries had been helpfully made by Mr Wright QC on behalf of the Crown, we concluded that, in view of the challenge made and the acceptance by Mr Ryder that he had used a likelihood ratio calculation, we should hear his evidence. The Crown also tendered two further statements. The first was that of Mr David Baldwin, the principal scientist in the FSS responsible for marks and setting standards; he was also chairman of the European Network of Forensic Science Institutes’ expert working group on marks. The second was that of Mr Justin Lewis, the principal scientist at the FSS responsible for Case Assessment and Interpretation. We heard oral evidence from all three, who gave their evidence in a manner that greatly assisted the court. Each was an impressive witness. In addition, at the invitation of the Court, the Forensic Science Regulator submitted a note to the court on which both the Crown and the appellant had the opportunity to comment. We are very grateful to the Regulator for his considerable assistance.

(2) **The development of the practice of using likelihood ratios in footwear mark evidence**

48. We therefore turn to set out our findings, based on the evidence given to us, as to the development of the use of likelihood ratios in England and Wales and the practice in relation to footwear mark examination elsewhere. The evidence in relation to England and Wales (supplemented as it was by the Forensic Science Regulator) was detailed.

That in relation to other countries, see paragraphs 62-68 below, was not as full as would have been desirable.

(a) *The development of likelihood ratios in England and Wales*

49. As we have mentioned at paragraph 31, some forensic practitioners and examiners in England and Wales developed the use of the numerical scale which we there set out. On the evidence before us this development occurred in the late 1990s and was based on the approach to expert evidence on DNA. It was thought appropriate to transfer the approach in relation to DNA to other areas of forensic evidence.
50. In 1998 a paper entitled “*A Bayesian approach to interpreting footwear marks in forensic casework*” by Evett, Lambert and Buckleton (the first two authors being employed by the FSS) was published (38 *Science and Justice* 241). It put forward the contention that the calculation of a likelihood ratio in respect of footwear marks would provide a formal framework for the interpretation of such evidence and the degree of probability that a mark had been made by a particular shoe; the article sets out a number of complex formulae. Some providers of forensic services adopted this approach as they found it helpful to use likelihood ratios to calculate a numerical value which when used against a scale, allowed them to express opinions on degrees of probability in the way we have described.
51. A presentation to a European meeting of shoe print and tool mark examiners in Berlin in 2001 explained the development at that time; the presentation is recorded in a paper entitled “*Principles of interpretation – application of the likelihood ratio in marks cases*” by Jackson, Champod and Evett, all of whom worked for the FSS. The paper began:

“The application of Bayesian reasoning, and the central importance of the likelihood ratio, has been extensively developed over the last three years within a wide range of casework carried out by the [FSS].”

The paper helpfully describes Bayes Theorem verbally as:

“Posterior odds (of an uncertain event, C) = Prior odds (of that uncertain event, C) multiplied by the likelihood ratio (provided by the evidence, E)”

And the likelihood ratio as:

“Probability of the observations (or evidence, E) given that the event (C) were true and given the truth of the background information (I), divided by the probability of the observations (E) given that the event were not true (C) and given the truth of the background information (I).”

The paper explained how this was applied in practice to footwear mark evidence.

(b) *The adoption in 2009 of standards which employed the likelihood ratio approach*

52. The number of forensic scientists and examiners within the UK who accepted this approach grew; it is clear, however, that a number did not adopt it. Nonetheless the approach was adopted by the Association of Forensic Science Providers, a body made up of leading forensic service providers in the UK and Ireland, in a paper entitled “*Standards for the formulation of evaluative forensic science expert opinion*” which was published in (2009) 49 *Science and Justice* 161-164. The evidence provided to us was that this had been discussed at the Forensic Science Advisory Council (FSAC) chaired by Mr Andrew Rennison, the Forensic Science Regulator earlier in December 2008. He has been entrusted by the Home Office with ensuring that appropriate quality standards are developed, implemented and used effectively in criminal justice. Considerable importance appears to have been attached to this paper within the community of providers of forensic services. 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.
53. The paper, though it post dates the trial of the appellant, contains the clearest statement of the approach adopted by Mr Ryder and considered by Dr Cooper. It contains the following definitions of “Evaluative Opinion”:

“An opinion of *evidential weight (evaluation of likelihood ratio)* based upon case specific propositions and clear conditioning information (framework of circumstances) that is provided for use as evidence in court. An *Evaluative opinion* is an opinion based upon the estimation of a likelihood ratio”.

It sets out the standards to be followed which include agreeing the key issues and identifying the relevant prosecution proposition and the relevant defence proposition. It continues:

“Establish if the scientific examination of items submitted can assist by carrying out a pre-assessment. This will involve:

- Identifying the potential outcomes of examinations/tests. ..
- Assigning estimated probabilities to each outcome given the probabilities and relevant information.
- Estimating likelihood ratios for each outcome and using them to inform the examination strategy and tests carried out.”

After setting out the subsequent procedures, the standards conclude:

“The conclusions in the report will be related to the propositions addressed and the estimated likelihood ratio.

The strength of evidence will be expressed either by a value of the likelihood ratio or using the verbal scale related to the value of the likelihood ratio. This verbal scale will be adopted by all AFSP organisations.”

The paper set out the scale which we have set out at paragraph 31 above. The notes make clear

“the opinion or conclusion should express the degree of support for one proposition or the other depending on the value of the likelihood ratio.”

54. The paper sets out what are described as guiding principles which justify this approach – the duty of the expert to base his opinion on four principles – balance, logic, robustness and transparency.

(c) *The argument for the use of the likelihood ratio or Bayesian approach*

55. Mr Lewis explained to us, as is apparent from the 2001 paper by Jackson, Champod and Evett, that the approach set out in the paper amounted to the adoption of what has been termed the “*Bayesian approach*”. It uses likelihood ratios to assess the evidential weight of scientific evidence, based on the framework of the background circumstances. The background circumstances and the key issues are used to formulate the alternative propositions – one for the prosecution case and one for the defence case. The evidential weight of scientific evidence can be expressed either on the verbal scale or the numerical scale to which we have referred; the numerical scale is the ratio of the likelihood of obtaining the scientific evidence if the prosecution proposition were to be true compared to the likelihood of obtaining the scientific evidence if the defence evidence were to be true. Although the background circumstances do not influence the likelihood of the evidence for each proposition, they may influence the way the propositions are formulated.

56. It was claimed that the use of a likelihood ratio enabled evidence to be evaluated by the expert in a balanced way after consideration of the prosecution and defences cases; it lowered the risk of the significance of scientific evidence being either overstated or understated.

57. Mr Lewis explained that in relation to footwear the first task of the examiner was to decide whether the mark could have been made by the shoe. If it could have been made, then what the FSS tried to do was to use the likelihood ratio to convey to the court the meaning of “could have been made” and how significant that was.

58. The FSS considered that, using the data they had collected on footwear, it could provide an assessment of “could have made”. Mr Lewis accepted that a criticism made of the use of likelihood ratios was that sufficient data were not available. He considered, however, that the collection of footwear pattern frequency data by the FSS was extensive and was sufficient to justify the use of a likelihood ratio calculation. Applying the analogy of DNA, the evidence of the FSS was that some assessment could be made of the significance of a match, though it was not, in contradistinction to DNA, an area where a fully numerical computerised assessment could be made. Indeed, as Mr Lewis accepted, numbers were not put into reports because there was a concern about the accuracy and robustness of the data, given the small size of the data set and factors such as distribution, purchasing patterns and the like. It was therefore important that the emphasis on the use of a numerical approach was to achieve consistency; the judgment on likelihood was based on experience.

59. 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. 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.

(d) The view of the Forensic Science Regulator

60. As is apparent from the history we have set out, the Regulator considered that the approach set out in the 2009 paper had a number of advantages, particularly a structured approach to the provision of an opinion which would make clear what factors had been considered and how the factors had been evaluated; this assisted in consistency. It also enabled the opinion to be reviewed.
61. Though he suggested that it was not logical to adopt the position that the Bayesian or likelihood ratio approach could be used in some areas, but not in others, he accepted in his submissions to us that the lack of data in some fields might make the application of the model more difficult and would affect the way it was applied.

(e) The evidence as to the position in the Netherlands, New Zealand and Switzerland

62. It is apparent that the adoption of the approach we have described was led by forensic practitioners and examiners in England and Wales. The evidence of Mr Baldwin was that an approach based on likelihood ratios is in use in the Netherlands Forensic Institute, Switzerland and Slovenia. It was not clear, however, how extensive that use is. In New Zealand the approach is apparently used, on the evidence given to us, in relation to DNA, glass, paint and fibres.
63. Mr Baldwin also told us that the European Network of Forensic Science Institutes had approved a statement undertaking to agree a standard for the interpretation of forensic evidence, but that there was a reluctance fully to adopt the Bayesian approach. The 2006 report of the working group on shoeprint and tool mark examinations of the European Network of Forensic Science Institutes, however, made it very clear that the classical approach of experts on footwear marks was to phrase in a verbal scale their conclusions as to the scale of probability that the sole made the marks. There were only some who advocated the Bayesian approach.

(f) The evidence on practice in the United States and Australia

64. The evidence in relation to the United States given by Mr Baldwin was that the vast majority of forensic examiners of footwear within the United States do not evaluate their findings using likelihood ratio calculations. His understanding was that the US courts require an expert to consider simply the question of whether the shoe made the mark; the courts did not allow experts to take account of motive or opportunity. This is, he considered, in contradistinction to England and Wales where “conditioning information” is used to formulate the alternative propositions that are central to the use of a likelihood ratio. Experts on footwear in the US also have concerns about the accuracy and the reliability of the data. They consider the same factors as examiners in England and Wales but do not use the data to calculate a likelihood ratio. The reasons for the rejection of the Bayesian approach are explained by Mr W Bodziak in a paper delivered to a European meeting of examiners in the early 2000s entitled *Traditional Reporting of Footwear Examination Results by the FBI and other Laboratories in the United States*.

65. The very different approach in the US is set out in an FBI paper by Michael B Smith (helpfully provided to the court by the FSS). It was published in July 2009 and is entitled *The Forensic Analysis of Footwear Impression Evidence*. This paper makes clear that a distinction is drawn between class characteristics (those that result from manufacture such as design and size) and identifying characteristics (such as objects attached to the sole such as small stones and damage caused by cuts). In the majority of examinations resulting in an association between the mark and the shoe, the conclusion most often reached is that an association based on class characteristics enables the examiner to say that the mark could have been made by the shoe or could equally have been made by another shoe having the same design and physical size. The author explained:

“The correspondence of combined class characteristics is valuable information, and its importance should not be understated. Shoes come with thousands of different outsole designs and numerous sizes for each design. Thus even a popular shoe outsole design represents only a small subset of all manufactured shoes. Based on information provided by footwear manufacturers regarding shoe designs and size ranges, the estimated frequency of any particular shoe design in a specific size is much less than 1 per cent of the total shoe population. Although training and experience provide an understanding of the general occurrence of combined class characteristics, footwear examiners do not specify a percentage or likelihood that a shoe of a particular size and design made an impression. Precise quantitative assessments are not provided because accurate information is lacking regarding the exact number of shoes produced in a particular design, size and geographic distribution as well as how many shoes of that design and size remain in use. Often class characteristics have been mistakenly regarded as having little or no value. However, any shoes outsole in a specific design represents a very small portion of all of the shoes manufactured in that design.”

The paper set out figures for purchases worldwide and in the United States. In contradistinction to class characteristics, identifying characteristics might enable the examiner to make a positive identification; mathematical models might, it is suggested, be of assistance in respect of these. Wear is, in general, to be treated as a class characteristic, though if the wear is significant it could amount to an identifying characteristic.

66. The scale of evaluation for footwear (and tyre) impressions used in the United States is different to that used in England and Wales. We were provided with the following scale which we were told is widely used in the United States: (i) identification – a definite conclusion that the shoe made the mark; (ii) probably made - a very high degree of association – correspondence in characteristics with a critical feature missing; (iii) could have made – a significant association of multiple class characteristics; (iv) inconclusive - limited association of some characteristics; (v) probably did not make unsuitable - a very high degree of non-association; (vi)

elimination – definitely not made; (vii) unsuitable – lack of sufficient characteristics for a meaningful comparison. The importance of this scale as a comparator is that it is more precise in differentiating between class and identifying characteristics; it does not express degrees of probability where there are no identifying characteristics.

67. Mr Ryder’s evidence to us was 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. Both, he said, expressed a judgment based on the skill and experience of the examiner in interpreting the observations. “Moderate support” corresponded to “could have made” in the US scale.
68. It was Mr Baldwin’s evidence that in Australia many examiners followed the US approach, but that some were looking at developing an approach based on using likelihood ratio calculations.

(3) Our conclusions in relation to evaluative evidence given by a footwear mark examiner

69. As we have explained the real issue as it emerged in the course of the appeal was whether a footwear mark examiner who had concluded that the mark could have been made by the defendant’s shoe could express an evaluative opinion of the probabilities it was made and, if so, the basis on which that opinion could be formed and expressed.
70. The principles for the admissibility of expert evidence were summarised recently in *Reed & Reed* at paragraphs 111 – 112: the court will consider whether there is a sufficiently reliable scientific basis for the evidence to be admitted, but, if satisfied that there is a sufficiently reliable scientific basis for the evidence to be admitted, then it will leave the opposing views to be tested in the trial before the jury.

(a) The scope of the footwear examiner’s role

71. It was accepted on behalf of the appellant that an examiner of footwear marks could embark on an examination that covered:
- i) An assessment of how the mark at the crime scene could best be visualised such as by enhancement or recovery.
 - ii) Preparation of test impressions from the footwear suspected to have made the mark.
 - iii) A comparison of the impression from the footwear and the crime scene mark.
 - iv) An assessment by the expert of the degree of match or non match in terms of pattern, size, mould detail, degree of wear and damage.
 - v) Subject to the strength of the evidence as to use etc., the assessment should take into account the date it was likely the marks were made and the date of the seizure of the footwear.
72. It was not accepted that the expert could interpret the degree of match. It was submitted on behalf of the appellant that in the case of footwear evidence, an

examiner should not be permitted to use the evaluative approach permitted in the case of facial mapping, but should, where there were no identifying characteristics and only class characteristics, only say whether the mark could or could not have been made by a particular shoe. It was argued that an expert in the case of facial mapping could comment on the verbal scale only because there were only indentifying characteristics in relation to a face.

73. We do not accept this broad submission. It seems to us that in some cases an expert examiner of footwear marks can indeed go no further than to express the opinion that the shoe could or could not have made the mark for the reasons powerfully advanced in the US materials to which we have referred. An opinion that a shoe “could have made the mark” is not in our view the same as saying that “there was moderate [scientific] support for the prosecution case”. The use of the term “could have made” is a more precise statement of the evidence; it enables a jury better to understand the true nature of the evidence than the more opaque phrase “moderate [scientific] support”. We have put the word “scientific” in parenthesis, as the use of this term is a matter to which we return at paragraph 96 below.
74. 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.
75. However, it is important to emphasise that the examiner is giving his opinion on the matters within his expertise – namely the footwear, the marks and, if relevant, scenes of crime evidence; it is not his function to evaluate the other evidence in the case.
76. We therefore turn to whether in the present state of knowledge it is permissible to use mathematical formulae and likelihood ratios based on statistics to arrive at that evaluative opinion in footwear mark cases. We do not agree with the observations of the Regulator that a similar approach is justified in all areas of forensic expertise. Each area requires a separate analysis because of the differences that there are in the nature of the underlying data.

(b) *The parameters of the DNA cases*

77. It is clear that in DNA cases, there has been for sometime a sufficient statistical basis that match probabilities can be given. The approach in the United Kingdom has been, in the light of *Doheny*, to use match probabilities. It is important to emphasise that the evidence is not directed to whether DNA came from the suspect, but the probability of obtaining a match that came from an unknown person who is unrelated to the suspect but has the same profile. In Australia, match probabilities and likelihood ratio calculations are made on the basis of a sound statistical database: the difference in method is illustrated in *R v Karger* [2002] SASC 294 at paragraphs 66-67. However it is clear that the likelihood ratio is permitted as a means of expressing a statistical opinion in DNA cases: see *Kargar* at paragraph 151, *R v GK* [2001] NSWCCA 504; see *R v Berry & Wenitong* [2007] VSCA 202 and in particular the

judgment of Redlich JA at paragraphs 28-30 . The approach also seems to have been adopted in New Zealand: *Lapper v R* [2005] NZCA 259 at paragraphs 19-22.

78. However, no case was drawn to our attention which suggests that a mathematical formula is appropriate where it has no proper statistical basis. We therefore turn first to examine the reliability of the statistics and data in relation to footwear. 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.

(c) *The robustness of the statistics and data on footwear: use of mathematical formulae should not be considered on the present state of the data*

79. The paper by Jackson, Champod and Evett (to which we referred at paragraph 51) rejected the suggestion that hard data were needed to evaluate a likelihood ratio. It did so on the basis that:

“Certainly the more data that is available the more informed would be the scientist’s probabilities. But ultimately, all the probabilities are subjective and based on a combination of personal experience and the available data. The likelihood ratio accommodates the uncertainty.”

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 acceptance of a mathematical approach to the calculation of a match probability in DNA cases is based on the reliability of the statistical database, though an element of judgment is required. It is therefore necessary to examine the evidence on the reliability of the data in relation to footwear.

81. 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. The figure for pattern could never be accurately known. For example, there were only distribution figures for the UK of shoes distributed by Nike; these left out of account the Footlocker shoes and counterfeits. The figure for size again could not be any more than a rough approximation because of the factors specified by Mr Ryder. Indeed, as Mr Ryder accepted, there is no certainty as to the data for pattern and size.

82. 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. A particular shoe might be very common in one area because a retailer has bought a large number or because the price is discounted or because of fashion or choice by a group of people in that area. There is no way in which the effect of these factors has presently been statistically measured; it would appear extremely difficult to do so, but it is an issue that can no doubt be explored for the future.

83. It is important to appreciate that the data on footwear distribution and use is quite unlike DNA. A person’s DNA does not change and a solid statistical base has been developed which enable accurate figures to be produced. Indeed as was accepted by Mr Ryder, the data for footwear sole patterns is a small proportion of what is in use

- and changes rapidly. As was accepted by Mr Lewis it would for these reasons be dangerous to use a straight statistical model.
84. Use of the FSS's own database could not have produced reliable figures as it had only 8,122 shoes whereas some 42 million are sold every year. As we shall explain at paragraph 103.v) below, the likelihood ratio calculated by using figures for the population as a whole is completely different to that calculated using the figures used by Mr Ryder based on the FSS database. 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.
85. The most in our judgment that can be derived from a comparison of pattern and size is that it can form part of the judgment based on how common the pattern or size is. It is impossible to see on the present state of information how any mathematical figure can be properly calculated to express a more definite evaluative opinion— there are far too many variables and uncertainties in the data.
86. In accordance with the approach to expert evidence we have set out at paragraph 70, 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. There are no sufficiently reliable data on which an assessment based on data can properly be made for the reasons we have given. An attempt to assess the degrees of probability where footwear could have made a mark based on figures relating to distribution is inherently unreliable and gives rise to a verisimilitude of mathematical probability based on data where it is not possible to build that data in a way that enables this to be done; none in truth exists for the reasons we have explained. 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.
87. It is of course regrettable that there are, at present, insufficient data for a more certain and objective basis for expert opinion on footwear marks, but it cannot be right to seek to achieve objectivity by reliance on data which does not enable this to be done. We entirely understand the desire of the experts to try and achieve the objectivity in relation to evidence of footwear marks, but the work done has never before, as we understand it, been subject to open scrutiny by a court.
- (e) *The use of likelihood ratios and the directions to the jury*
88. Our conclusion renders it unnecessary to decide how likelihood ratios should be used in expert footwear mark evidence and how the use of likelihood ratios should be explained to a jury. We can therefore mention the issues briefly.
89. In *Adams*, expert evidence was given, without objection, in a DNA case as to the use of statistical evidence and Bayes theorem to deal with the other non DNA evidence in the case. Rose LJ, giving the judgment of the court explained at page 481 that whilst Bayes theorem might be an appropriate tool for statisticians to establish a mathematical assessment of probability, “apparently objective numerical figures used in the theorem may conceal the element of judgment on which it entirely depends”. It was inappropriate for use in jury trials for a number of reasons; the jury's task was to assess the evidence by common sense and their knowledge of the world and not by

reference to a formula. It would be impractical, as different members of a jury might have different views as the constituent numbers to be used in the formulae. He concluded:

“Quite apart from these general objections, as the present case graphically demonstrates, to introduce Bayes Theorem, or any similar method, into a criminal trial plunges the jury into inappropriate and unnecessary realms of theory and complexity deflecting them from their proper task.”

That conclusion was strongly endorsed in *Doheny* by Phillips LJ (as he then was) in giving the judgment of the court and in *Adams (No 2)* by Lord Bingham CJ.

90. It is quite clear therefore that outside the field of DNA (and possibly other areas where there is a firm statistical base), this court has made it clear that Bayes theorem and likelihood ratios should not be used. We should, however, draw attention to *R v George* [2007] EWCA Crim 2722 and to work done by Dr Evett in relation to the strength of the evidence of firearm discharge residue in that case (see paragraphs 14-22). It is apparent, reading the report in the light of what has emerged in this appeal, that Dr Evett was using a likelihood ratio approach, as advocated by him in the papers to which we have referred. However Lord Phillips CJ (in giving the judgment of the court) did not consider the merits of the approach in the judgment or how it was consistent with the views expressed in *Adams*, *Doheny* and *Adams (No2)*. It appears simply to have formed part of the background to the court’s consideration of the overall evidence in the case.
91. It is not necessary for us to consider the matter further and consider how likelihood ratios and Bayes theorem should be used where there is a sufficient database. If there was a sufficient database in footwear cases an expert might be able to express a view reached through a statistical calculation of the probability of the mark being made by the footwear, very much in the same way as in the DNA cases subject to suitable qualification, but whether the expert should be permitted to go any further is, in our view, doubtful. The literature on the extent to which Bayesian reasoning should be permissible in such circumstances in a jury trial is extensive: see for example the comment by Professor Ormerod on *Adams* [1996] Crim L R 898, Professor Glanville Williams “*The Mathematics of Proof*” [1979] Crim L R 341 at 346-350, Professor Redmayne *Doubts and Burdens: DNA Evidence, probability and the Courts* [1995] Crim LR 464, and Chapter 4 of *Expert Evidence and Criminal Justice, Phipson on Evidence*, 17th edition (2010) paragraphs 34-36 to 34-38 and Roberts and Zuckerman’s *Criminal Evidence* at pages 123-132. It is important to note as happened in *R v Abadom* [1983] 1 All ER 364 (where the refractive index of glass was being considered) that an expert can give an opinion using a statistical database by simply using that database and expressing an opinion by reference to it, without recourse to the type of mathematical formula used in this case or to any form of Bayes theorem.
- (e) *An evaluative opinion in the absence of a database*
92. However, the fact that there is no reliable statistical basis does not mean a court cannot admit an evaluative opinion. As we have set out at paragraph 70, it can do so, where there is some other sufficiently reliable basis for its admission. For the reasons

we have given at paragraph 73, there may well be factors in a footwear marks case where an examiner can give an evaluative opinion. In other areas of forensic evidence this happens; for example in *Reed & Reed*, evaluative expert evidence was admitted where the issue related to evaluation of the possibilities of secondary or tertiary transfer of DNA: some evidence in relation to experiments was available but there was no statistical or database evidence. In *R v Weller* [2010] EWCA Crim 1085, evaluative evidence in relation to the possibilities of transfer of DNA was also admitted on the basis of work done by forensic experts where no database existed; the issue on appeal was again whether there was a sufficiently reliable basis for the admission of such evidence. In neither case was there any question of a statistical basis or the use of a likelihood ratio.

93. In *R v Atkins and Atkins* [2009] EWCA Crim 1876 the use of a scale to express an expert opinion was approved by this court in the context of facial imaging. It is important to note that in that case it was made clear that the scale was one based solely on the judgment of the expert without any figures or mathematical formula. Hughes LJ said in giving the judgment of the court:

“An expert who spends years studying this kind of comparison can properly form a judgment as to the significance of what he has found in any particular case. It is a judgment based on his experience. A jury is entitled to be informed of his assessment. The alternative, of simply leaving the jury to make up its own mind about the similarities and dissimilarities, with no assistance at all about their significance, would be to give the jury raw material with no means of evaluating it. It would be as likely to result in over-valuation of the evidence as under-valuation. It would be more, not less, likely to result in an unsafe conclusion than providing the jury with the expert's opinion, properly debated through cross-examination and, if not shared by another expert, countered by contrary evidence.”

The court was again holding that evaluative evidence could be given; there was no issue as to the use of a likelihood ratio, as it was clear that there was no statistical database.

94. As Hughes LJ also made clear in *Atkins v Atkins* at paragraph 23, evidence of this kind needs to be approached with caution and cross examination played an essential role as he explained at paragraph 29:

“The absence of a statistical database is something which will undoubtedly be exposed in cross-examination. The witness may expect to be asked to explain how, if no-one know how often ears or noses of the shape relied upon appear in the population at large, it is possible to say anything at all about the significance of the match; his answers may be satisfactory or unsatisfactory but will be there to be evaluated by the jury, which will have been reminded by the judge that any expert's expression of opinion is that and no more and does not mean that he is necessarily right. Similarly, the expert may be expected to be tested upon the extent to which he has not only

looked for similarities, but has actively sought out dissimilarities. Those are but the simplest of the questions which plainly need to be asked of anyone offering evidence of this kind. Cross examination will also be informed by the fullest disclosure of his method, generally, and of his working notes in the particular case being tried.”

95. 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.
96. 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.

(f) Transparency and pre-trial hearings

97. The importance of an expert complying with his duties under Part 33 of the Criminal Procedure Rules has been emphasised by this court in *Reed & Reed* and in *R v Henderson* [2010] EWCA Crim 1269. As was made clear by Mr Ryder in his evidence to us, it is also important that besides being balanced, clear and logical, it is essential that an expert report is transparent. Where the mark could have been made by the footwear, the factors that enable the expert to express a more definite evaluative opinion must be set out, including any data on which reliance is placed.
98. The report can be considered by the court at a pre-trial hearing and, if there is a challenge to reliability of the basis on which an evaluative opinion is being given, the court can make directions as to the resolution of the issue of its admissibility.
99. The justification advanced in the evidence in this appeal for not including in the reports the use of the formula and statistics was that it might confuse the jury. No doubt this was a reaction to the perceived consequences of the views of this court expressed in *Adams* and subsequent cases. The justification advanced can, however, be no justification, as a court must know what is being done. The report is in any event not put before the jury. If the way in which the opinion on the footwear mark evidence had been reached in this case had been put into the report and been available to the Recorder [REDACTED], then in the light of *Adams*, we have no doubt that the argument that has taken place on this appeal would have taken place at the trial. The decision of this court in *Abadom* as long ago as 1982 explained the importance of referring to all the material so that the cogency and probative value of the conclusions can be tested and evaluated by reference to it.

(4) Our conclusion in relation to the footwear mark evidence in the appeal

(a) *Was there fresh evidence admissible under s.23?*

100. It was submitted by the appellant that the evidence of Mr Ryder, Mr Baldwin and Mr Lewis which we have set out should be admitted under s.23 of the Criminal Appeal Act 1968. From that fresh evidence, it was clear that (1) the statistical evidence was inherently unreliable; (2) the jury had not been told the basis on which Mr Ryder had reached his conclusion; they had acted on figures (which the appellant's counsel had put before the jury and on which the judge had summed up) in relation to pattern and size (configuration) that were much more unfavourable to the appellant than the figures on which Mr Ryder had reached his conclusion; (3) the process had not been transparent. The conviction could not, in the light of that evidence, be regarded as safe.

[REDACTED]

102. [REDACTED] we consider we should admit the evidence, as it is now clear that the basis on which [REDACTED] Mr Ryder [REDACTED] had reached the [REDACTED] view as expert [REDACTED] for the Crown [REDACTED] was not before the court at trial. It is plainly in the interests of justice that we consider the safety of the conviction in the light of all the evidence. We therefore turn to analyze the significance of what emerged in the course of the appeal.

(b) *A comparison of the evidence before us and the evidence at the trial*

103. As we have set out at paragraph 35, we have concluded that none of the lawyers instructed in the case nor the trial judge were made aware of the use of likelihood ratios or the statistics actually used in the preparation of Mr Ryder's reports.

- i) The figures which Mr Ryder had used for his report were not led in evidence. Counsel for the Crown was unaware of the conservative figures used by Mr Ryder. As he very frankly told us and as is also apparent from his re-examination, he considered that the figures put by the defence to Mr Ryder in the course of cross examination strengthened the evidence against the appellant.
- ii) The statistics for frequency of pattern actually used by Mr Ryder - 1 in 5 of the shoes coming into FSS laboratories - were not in his report and were not put before the jury. He simply referred to the pattern as being "one of the pattern types most frequently encountered".
- iii) His report did not mention and the jury were not told that his opinion had been based on the more conservative figure of 10% for size rather than the 3% mentioned at trial.
- iv) The jury were not told that he could exclude half of the trainers because of wear and that the damage pointed to it being less likely that the trainers had caused the mark.

- v) The jury were given two important statistics based on the population as a whole:
- a) for pattern type of 0.25% or 1 in 400; this was given by the judge, as set out at paragraph 44.
 - b) for size of 1 in 33; this was elicited in cross examination as set out at paragraph 42.

These figures were very different to the figures actually used by Mr Ryder in forming the opinion set out in his report.

- vi) It is clear that, if the matter had been left to the jury without any reference to figures, the overall opinion of Mr Ryder was much more favourable to the appellant than one which appeared from the figures put before them.

104. Counsel for the appellant illustrated the consequences by application of the formula used by Mr Ryder to the figures put before the jury:

- i) If the figures used in the evidence for pattern type (1 in 400) and for size (1 in 33) are multiplied together, taking the first part of the formula ($P \times C$), the result would be a likelihood ratio of 1 in 13,200 as opposed to the likelihood ratio of 1 in 50 used by Mr Ryder, based on his more conservative approach. If the figure for wear used by Mr Ryder was then applied to the result of 1 in 13,200 ($P \times C \times W$), then the likelihood ratio would be 26,400 – in the verbal scale “very strong support.”
- ii) If the likelihood ratio of 100 calculated by Mr Ryder for the purpose of his first report had been explained to the jury, that likelihood ratio on its own, as Mr Ryder made clear, implied that 1 in 100 trainers could have made the mark, not the figure arrived at by the calculations using the figures given to the jury.
- iii) It was also clear that if the formula was used for most footwear where there was a full mark, it would be highly unlikely that there would be a value of less than 10 because there were very few patterns that would have a value of less than 10 for P . Applying the verbal scale, this would mean that there would usually be “moderate support”. Indeed Mr Ryder made clear that if the formula produced a figure above 1, it would provide “weak support” for the prosecution case.

105. On the other hand, as Mr Ryder made clear, he had not relied only on the formula in arriving at his opinion. He had simply used the formula in accordance with the then emerging practice, namely as an aid to standardise opinions across the board and evaluate cases relative to each other. His opinion was reached on the basis of his experience and judgment.



[REDACTED]

[REDACTED]

(c) *The conviction was not safe*

108. Nonetheless, we cannot say that the conviction was safe:

- 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.
- 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.
- 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.
- 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.

109. It is, we hope, highly unlikely that the process by which expert evidence was formulated and adduced in this case will ever be repeated; the dangers to which the Crown referred are therefore unlikely to arise in future cases. We have referred at paragraphs 97-99 to the importance of the requirements of transparency. For the future, strict compliance with these obligations should ensure that in each case where an expert provider of a forensic opinion seeks to develop a new way of arriving at an opinion, that new way can be examined in open court applying the ordinary principles for the admissibility of expert evidence. There is, we think, no need for any new process. The principles for admissibility and provision of expert evidence are clear, but courts cannot apply them if they are not made aware of the way in which an expert has reached his opinion through the provision of a full and transparent report.

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]



Conclusion

130. We therefore allow the appeal [redacted] and quash the conviction.