



COMMENTARY

Interpreting Evidence of Torture

BERNARD WN ROBERTSON¹ AND
CHARLES EH BERGER^{2,*}

¹ Barrister, Wellington, New Zealand; Department of Law, AUT University, Auckland, New Zealand

² Netherlands Forensic Institute; Institute for Criminal Law and Criminology, Leiden University

*c.e.h.berger@law.leidenuniv.nl

ABSTRACT

The Istanbul Protocol provides a scheme for giving evidence of signs of torture. This scheme does not conform with the principles of logical inference, revolving as it does round the concept of ‘consistency’. The shortcomings of the Protocol are explained using the evidence given in the recent case of *KV(Sri Lanka)* and the logical approach to such evidence explained.

KEYWORDS: Consistency, Istanbul Protocol, *KV(Sri Lanka)*, Likelihood ratios, Probability, Torture

The ‘Istanbul Protocol’ or to give it its full name, the ‘Manual on the Effective Investigation and Documentation of Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment’, sets out guidelines for interpreting medical evidence of torture. The Protocol is said by the UK Supreme Court in *KV(Sri Lanka) v Secretary of State for the Home Department*¹ to be the product of a great deal of work by doctors, lawyers, and human rights experts. Unfortunately, they did not seem to include any forensic scientists familiar with modern approaches to interpretation and the Protocol contains fundamental flaws which undermine its usefulness. This is illustrated by *KV*.

KV, a Sri Lankan of Tamil ethnicity, came to the UK and claimed asylum. He alleged that he had been tortured by Sri Lankan government forces in the course of detention on suspicion of association with the LTTE (‘the Tamil Tigers’). He had five long scars on his back and two shorter scars on his right arm. It was accepted that these were the product of branding with a hot metal rod. He contended that they were evidence of the torture. The UK Upper Tribunal (Immigration and Asylum

¹ *KV(Sri Lanka) v Secretary of State for the Home Department* [2019] UKSC 10.

© The Author(s) 2019. Published by Oxford University Press. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Chamber) in effect concluded that the scars represented wounding Self-Inflicted By Proxy ('wounding SIBP'), in other words inflicted by another person at KV's invitation in an attempt to manufacture evidence in support of a false asylum claim.² The Tribunal dismissed his appeal against the refusal of asylum and the Court of Appeal dismissed his further appeal by a majority.³ In doing so, the Court of Appeal made controversial observations about the limit of the role of a medical expert in contributing to the evidence referable to a claim of torture. These comments were considered to raise a point of general public importance on which KV was granted permission to appeal to the Supreme Court. Other aspects of the case, however, raise issues of general importance relating to the 'Istanbul Protocol' and to the interpretation of medical evidence generally.

It is always important to keep in mind what ultimately has to be proved. In the case of an asylum-seeker it is that there is a well-founded fear of persecution in the home state and that owing to such fear the asylum-seeker is unable or unwilling to avail himself of the protection of that state.⁴ One can speculate that the scars could be from an older and unrelated incident, or were the result of involvement in organised crime. In this case however, the persecution alleged was by the government and the question of whether there was a well-founded fear of persecution was treated as equivalent to the question whether the torture alleged had occurred. The question whether the torture had occurred achieved even greater significance as the evidence relating to the general train of events was equivocal and in the end the question whether the torture had occurred was regarded as determinative of the question whether KV had been persecuted at all.⁵

THE ISTANBUL PROTOCOL

The evidence considered by the Court of Appeal was the evidence of a doctor who had been trained in dealing with cases of torture. He observed, and photographs shown to the Tribunal confirmed, that the scars on KV's back were long, narrow, and parallel and their edges were precise. The scars on the arms, however, were not parallel and their edges were blurred. From this, the doctor inferred that the wounds on the back had been inflicted while KV had been unconscious but he had been conscious when the wounds on the arms were inflicted. Even if he had been restrained, reflex actions would have caused the edges of the scars to be irregular. The doctor surmised that the wounding to the arms may have caused KV to pass out and he had remained unconscious while the wounds to the back were inflicted. This was in accordance with KV's own evidence.

What complicated this picture was that the doctor tentatively said that it would have taken ten minutes to inflict such wounds. The question then arose whether a victim could remain unconscious for that length of time while further pain was being inflicted without having been anaesthetised. The doctor thought this might be the case if KV had been generally unwell, malnourished and so on. He summarised his

2 *KV v Secretary of State for the Home Department* [2014] UKUT 230 (IAC).

3 *KV v Secretary of State for the Home Department* [2017] EWCA Civ 119, [2017] 4 WLR 88.

4 International Refugee Convention 1951, art 1(2).

5 *KV(Sri Lanka) v Secretary of State for the Home Department* [2019] UKSC 10 at [8].

evidence by saying that his clinical findings were 'highly consistent' with KV's account (at [14]).⁶

In giving that summary the doctor was following guidance laid down in Annex 1 of the Istanbul Protocol. The Manual provides as follows:

187. . . . For each lesion and for the overall pattern of lesions, the physician should indicate the degree of consistency between it and the attribution given by the patient. The following terms are generally used:

- a. Not consistent: the lesion could not have been caused by the trauma described;
- b. Consistent with: the lesion could have been caused by the trauma described, but it is non-specific and there are many other possible causes;
- c. Highly consistent: the lesion could have been caused by the trauma described, and there are few other possible causes;
- d. Typical of: this is an appearance that is usually found with this type of trauma, but there are other possible causes;
- e. Diagnostic of: this appearance could not have been caused in any other way than that described.

188. Ultimately, it is the overall evaluation of all lesions and not the consistency of each lesion with a particular form of torture that is important in assessing the torture story . . .

The first issue with para 187 of the Protocol is the meaning of 'degree of consistency'. To physicians and scientists 'consistent' means 'not inconsistent' and 'inconsistent' means 'logically incompatible' with. Anything which is not inconsistent is consistent. When lawyers and police hear 'consistent with', however, they tend to understand it as meaning 'supporting the hypothesis of'.⁷

Whether A is consistent with B is a binary decision. Two matters are either consistent or they are inconsistent. The Protocol bears this out because under each heading consistency simply means that the lesions could have been made by the trauma described. The meaning of 'degree of consistency' is not clear. It could be regarded as referring to the probability of observations given an hypothesis: in this case, the probability that these lesions would be found if KV's account is true. Even that does not tell us its value as evidence.

The value of an item of evidence is determined by its ability to distinguish between hypotheses. That in turn is determined by how probable the evidence is under each hypothesis. The ratio of those probabilities (the Likelihood Ratio) is the measure of the strength of the evidence.⁸ The instruction in the Protocol, however, is only to indicate the 'degree of consistency between [the observation] and the attribution given

⁶ *ibid* [10].

⁷ See discussion and references in B Robertson, GA Vignaux and CEH Berger, *Interpreting Evidence* (Wiley, 2nd edn, 2016) 59–60.

⁸ *ibid* chs 2 and 3; F Taroni et al, *Data Analysis in Forensic Science* (Wiley 2010) 48.

by the patient'. This is only half the story. The evidence could be equally probable under some other hypothesis. The decision-maker would then have to consider how probable each of the two hypotheses was considering all the other evidence. This is illustrated in the present case where the Tribunal considered that there were only two real possibilities: torture and wounding SIBP. The evidence might then be considered highly probable under the attribution given by the 'patient' but that story may a priori be highly unlikely.

Let us now examine each of the 'terms . . . generally used'.

Not consistent: this appears to mean that the probability of finding this evidence given the hypothesis is zero. This will mean that the likelihood ratio will be infinitely in favour of any alternative hypothesis, as long as the probability of finding the evidence is more than zero for that alternative hypothesis. Conversely, if the probability of finding this evidence given the hypothesis is not zero but very small, the evidence will still support this hypothesis if the probability of finding the evidence under a relevant alternative hypothesis is even smaller.⁹

Consistent with: this says little more than that the lesions could have been caused by alleged trauma but they could also have been caused by something else. So, the idea of alternative hypotheses is raised but without any discussion of their role. There is no reference to an assessment of the probability of the lesions given the applicant's story, merely that the lesions could have been caused by the trauma described. The rubric then refers to there being many other possible causes. But it is not the number of possible causes that matters. What matters is the probability of the lesions given those hypotheses. Perhaps the intuition is that if there are many other possible causes there will be one under which the evidence is highly probable, but that is what should be considered.¹⁰

Highly consistent: this sounds as if it might mean that the evidence is highly probable under the applicant's hypothesis but it does not. It means only that the lesions could have been caused by torture. Again, the number of alternative hypotheses is irrelevant. It would not matter, for example, that there were only a small number of alternative hypotheses if the evidence were highly probable given one of them or if that alternative hypothesis had a high prior probability.

Typical of: this comes closer to consideration of the probability of the evidence given the applicant's hypothesis but goes on merely to require that there be alternative explanations. Again, what matters is the probability of the evidence given an alternative. It is also unclear whether this is stronger or weaker evidence than 'highly consistent'.¹¹

⁹ This, incidentally, demonstrates one of the problems with significance tests. A significance test would reject a hypothesis under which the evidence is very unlikely. But if the evidence is even more unlikely given the relevant alternative hypothesis the evidence will actually support the hypothesis that was rejected. See Taroni et al (n 8) 53 and Robertson, Vignaux and Berger (n 7) 148–50.

¹⁰ See ET Jaynes, *Probability Theory: The Logic of Science* (CUP, 2003) 103, where it is shown that the relative probabilities of the two most probable hypotheses will approximate true odds (as if the hypotheses were exhaustive) and other hypotheses are likely to be of little importance unless new evidence changes the assessments of probability.

¹¹ A point made by the Upper Tribunal [2014] UKUT 230 (IAC) at [276] n 2.

Diagnostic of: this refers to evidence which could only occur given the hypothesis. It is the converse of ‘not consistent’ in that the probability of the evidence given any other explanation will be zero. Such evidence will seldom occur in real life apart from analysis of pathogens. This phrase tends to underplay the extent to which diagnosis is often a process of inference from more equivocal evidence.

Ultimately, it is the overall evaluation of all lesions . . . that is important in assessing the torture story: This is correct: all the evidence should be combined and its combined effect considered, rather than examining each item of evidence in turn and ‘accepting’ or ‘rejecting’ it. Furthermore, ‘assessing the torture story’ requires more than just medical observations. They can only provide evidence when torture is compared to an alternative hypothesis. The value of that evidence must then be added to that of the other evidence and then to the prior probabilities of the alternative stories.¹²

The Supreme Court at [17] was content that the doctor’s evidence that his findings were ‘highly consistent’ with the applicant’s account had been framed in accordance with para 187(c) of the *Manual*. At [14], however, the Court records the doctor’s conclusion that his clinical findings were ‘highly consistent’ with the applicant’s account and that the other hypothesis of wounding SIBP was unlikely. This is not given as a quotation and maybe a misunderstanding of what the doctor said, but somewhere there was confusion. The ‘highly consistent’ statement was a statement about the evidence. The comment that the other hypothesis was unlikely is a statement about the probability of the hypothesis, something which the Court rightly said the expert witness was not entitled to give.

THE EXPERT’S BELIEF

That issue arose because a submission was made that in cases of alleged torture expert witnesses are entitled to express the view that they believed that the person had suffered torture and that that belief constituted independent evidence to be taken into account by the decision maker.¹³ The Supreme Court pointed out that *R(AM)* was a case where the expert had described the evidence as ‘diagnostic’ which the Court said was tantamount to belief in the explanation given. The Court considered that there was no room for such an expression of belief or otherwise in instances (b), (c), and (d) in the Protocol.¹⁴ Even when the expert referred to (a) or (b), credibility was still a matter for the decision-maker and might be the key issue.

It is submitted that the Court was correct. Historically, at common law, the Ultimate Issue Rule prevented experts from giving an opinion as to the matter the court had to decide. In recent decades this rule has come under fire and has been retreated from.¹⁵ It is in fact a requirement of logic. In addition to the reason given by the Court there are two reasons why the expert witness should not express such a belief. The first is that examining a single piece of evidence can produce only a likelihood ratio which is the relative probability of the evidence under each hypothesis. To form

12 In keeping with Bayes’ Theorem, see Robertson, Vignaux and Berger (n 7) ch 6 and Appendix 2.5 – 2.8 and Taroni et al (n 8) 43.

13 Citing *R(AM) v Secretary of State for the Home Department* [2012] EWCA Civ 521 at [29] and [30].

14 *KV(Sri Lanka) v Secretary of State for the Home Department* [2019] UKSC 10 at [25].

15 See, eg, the (New Zealand) Evidence Act 2006, s 25; US Federal Rules of Evidence, r 704(a). See, contra Taroni et al (n 8) 59 and Robertson, Vignaux and Berger (n 7) 50.

a belief as to the probability of the hypothesis one needs to combine the likelihood ratio with one's assessment of prior odds based on the other evidence and background knowledge. This is clearly the province of the decision-maker or tribunal of fact. The second reason is that the expert's opinion will be based on the observations the expert has reported and the decision-maker has heard; if it were taken into account, then, this would mean double counting the observational evidence. These comments are true of all expert evidence. The Supreme Court was therefore correct to say at [25] that there was no room in the Protocol for expression of belief or otherwise in the account given. But the Court at [14] appears to acquiesce in expression of the view that the hypothesis of wounding SIBP was unlikely.

SCOPE OF PERMISSIBLE EVIDENCE

The issue on which leave to appeal was granted concerned the permissible scope of an expert witness's evidence. The medical witness had said that the lesions were diagnostic of the application of hot metal rods and that they were 'highly consistent' with the applicant's account of torture. The majority of the Court of Appeal, without having invited any argument on the subject, observed that in making the second comment, the witness had trespassed into the role of the Tribunal and had effectively been saying that he believed the applicant's account. The response of the Supreme Court was:

[20] . . . decision-makers can legitimately receive assistance, often valuable, from medical experts who feel able, within their expertise, to offer an opinion about the consistency of their findings with the asylum-seeker's account of the circumstances in which the scarring was sustained, not limited to the mechanism by which it was sustained. . . . when [the doctor] proceeded to correlate his findings of a difference in the presentation of the scars on the back and those on the arm with KV's account of how the alleged torture had proceeded, he was giving assistance to the tribunal of significant potential value; and it never suggested that he lacked the expertise with which to do so.

The Supreme Court backed this up by citing several previous cases from British courts and tribunals and the European Court of Human Rights and other authorities in which it had been said that medical witnesses could give an opinion on the consistency of their findings with the account given by the asylum-seeker and not merely with the physical mechanism by which the wounds had been inflicted.¹⁶

This should have been considered by reference to first principles. Authority is not required for the admission of evidence which is relevant and probative. Such evidence is admissible unless there is authority for its exclusion. The question therefore simply is whether the witness has expertise which enables him or her to give evidence which is of probative value. It would be preferable, of course, that the witness expressed the level of support the evidence provided for the account given as compared with some appropriate alternative hypothesis.

16 KV(Sri Lanka) (n 5) [21]–[23].

COMPLEMENTARITY AND CHAMBERLAIN

Alternative hypotheses were considered when the Court dealt at [9] with a comment by the Upper Tribunal relating to the possible alternative hypotheses:

[364] . . . In relation to the medical evidence, we have found that whilst it assisted in eliminating some possible causes, it left us with only two that were real possibilities: that the appellant was tortured as claimed; that his scarring was SIBP. Of these two real possibilities, we have found, on analysis, that the former claim does not withstand scrutiny. Certainly we cannot say in his case that the evidence inexorably points to SIBP, but given that we have concluded it is left as the only real possibility that we have not been able to discount, taking the evidence as a whole, we are satisfied that he has not shown his account is reasonably likely to be true.

The Supreme Court then said at [10]:

One should respectfully place a question-mark against the tribunal's disclaimer in para 364 of any conclusion that the evidence inexorably pointed to wounding SIBP. If your inquiry into the disputed circumstances of a past event leads you to conclude that there are only two real possibilities and if you then proceed to reject one of them (indeed in this case to reject it in terms which could not be more absolute: see para 365), you are necessarily concluding that the other real possibility represents what happened.

If the hypotheses of torture and of SIBP are regarded as the only two possibilities, and as mutually exclusive, then one must be false and the other true. Given that we do not have perfect information, we have to talk in terms of probabilities that one or the other happened. The higher we assess the probability of one story, the lower we must assess the probability of the other. The probability assessments are complementary.¹⁷

The possibility that a hypothesis that has not been considered may be true could explain what Gibbs CJ and Mason J in the High Court of Australia meant in the '*Dingo Baby Case*':

. . . once the possibility that one of the children killed Azaria is rejected, as it was by common agreement at the trial, only two possible explanations of the facts remain open - either a dingo took Azaria, or Mrs Chamberlain killed her. Therefore, if the jury were satisfied beyond reasonable doubt that a dingo did not kill the baby, they were entitled to accept the only other available

¹⁷ This is formally known as the axiom of complementarity. If we write B as the complement of A, denoted by $\sim A$, the complement rule states that:

$$P(A|I) + P(\sim A|I) = 1.$$

This simply means that it is certain that either A or $\sim A$ (not A) is true. Each probability is conditional upon all our information ("I"). Hence, if we can assess the probability that A is not true $P(\sim A|I)$, the probability that A is true ($P(A|I)$) must logically be $1 - P(\sim A|I)$. See Robertson, Vignaux and Berger (n 7) 186; Taroni et al (n 8) 18.

hypothesis, that Mrs Chamberlain was guilty of murder. However, it would have been unsafe for a jury to approach the case by asking 'Are we satisfied that a dingo did not do it?' because that would have diverted attention from the evidence that bore on the critical issue - whether Mrs Chamberlain killed the baby.¹⁸

This appears self-contradictory. Given the complement rule, if the assessed probability of one explanation increases, the probability of the other decreases. One should therefore be able to prove one by disproving the other as the first part of the quotation suggests. That the Judges were not happy with doing so in the *Dingo Baby Case* probably reflects a lurking doubt that there may have been a third possible hypothesis.

The Supreme Court's statement in *KV* is much more correct than that of the High Court of Australia above. It is a welcome statement of the applicability of the complement rule to court decision making. The questioning of the Tribunal may be merited in this case but, as a general proposition, the Court's last sentence is questionable because it uses the language of rejection (and assumes its converse, acceptance). We should not, however, talk in terms of acceptance and rejection of evidence but of combining evidence having considered its relative probability under appropriate hypotheses.

PRIOR PROBABILITIES

Prior probabilities assumed importance when the Supreme Court discussed the dissenting judgment of Elias LJ in the Court of Appeal. At [33], the Court considered evidence of underlying 'base rates', i.e. it considered that it was commonplace that torture had taken place in Sri Lanka while wounding SIBP was almost unknown. At [34] the Court considered other evidence specific to this case that tended to favour the hypothesis of torture rather than wounding SIBP. The Court at [35] approved the conclusion of Elias LJ at CA [101]:

In my view, very considerable weight should be given to the fact that injuries which are SIBP are likely to be extremely rare.

In other words, even if the likelihood ratio provided by the evidence of the lesions is in favour of the proposition of SIPB, the a priori odds are so low that very considerable weight of evidence is required to produce posterior odds in favour of SIBP.

CONCLUSIONS

It is respectfully submitted that the Supreme Court judgment in *KV* correctly identifies and deals with issues on which The Istanbul Protocol is unclear or silent. The central issue, however, is the Protocol's unhelpful use of the language of 'consistency'. On that issue, the Court does not question the clear words of the Protocol but accepts them as the product of the work of doctors, lawyers and human rights experts. The language of 'consistency' does not, however, direct physicians to the correct questions. This article argues that:

18 *R v Chamberlain (No 2)* (1984) 153 CLR 521, 51 ALR 225 at [74].

- the first question is not the consistency of observed lesions with an hypothesis but the probability of observing such lesions under that hypothesis;
- the second question is not the number of possible alternative causes but the consideration of an alternative hypothesis and the probability of observing the lesions under that hypothesis;
- the ratio of the probabilities of observing the lesions under either hypothesis (the likelihood ratio) gives the value of the evidence as support for one hypothesis relative to the other;
- the likelihood ratio for the evidence must then be combined by the decision maker with other evidence and with an assessment of prior odds to arrive at an assessment of odds in favour of, or against, the hypothesis.