The Continuing Saga of the *Chamberlain* Direction: Untangling the Cables and Chains of Criminal Proof

DAVID HAMER*

I INTRODUCTION

In *Chamberlain* v *The Queen (No 2)*[1] the High Court analysed the structure of circumstantial proof and endorsed principles laid down by the South Australian Supreme Court in *The Queen* v *Van Beelen*[2]. In the leading judgment in *Chamberlain (No 2)*, Gibbs CJ and Mason J said that 'the jury cannot view a fact as a basis for an inference of guilt unless at the end of the day they are satisfied of the existence of that fact beyond reasonable doubt.'[3] In that case, which concerned the alleged murder of a baby, a majority held that the jury should not have based the inference of guilt upon evidence that traces of baby's blood were found in the accused's car, in view of the doubt raised by the defence about the blood tests.[4]

However the principles advanced in *Chamberlain (No 2)* became the subject of much uncertainty among trial judges and state courts of criminal appeal.[5] How should one distinguish between a 'fact' and an 'inference'? Was this statement consistent with the cumulative nature of circumstantial proof? When should a *Chamberlain (No 2)* direction be given?[6]

In *Shepherd* v *The Queen (No 5)*[7] the High Court advanced a comprehensive restatement of the principles of circumstantial proof. The relevant distinction is not between facts and inferences, but between the cable inference, where each strand of argument increases its strength, and the chain inference, which can be no stronger than its weakest link. The criminal standard has application only to the indispensable links of chain inferences.[8] In that case there were a number of strands of evidence implicating the accused, and so it was

* BSc LLB(Hons). David Hamer is a PhD student at the University of Melbourne. This paper is work in progress towards his dissertation. The author gratefully acknowledges the comments and criticisms on an earlier draft of this paper, by Dan Hunter, Richard Johnstone and Andrew Palmer of the Faculty of Law, the University of Melbourne, and Edwin Coleman of the Philosophy Department, the University of Melbourne.

1 (1984) 153 CLR 521.

2 (1973) 4 SASR 353.

3 (1984) 153 CLR 521, 536.

4 See infra fnn 59 — 82.

5 Eg *The Queen* v *Maleckas* (1991) 1 VR 363, see fn 82 infra; cf W A N Wells, *Natural Logic, Judicial Proof and Objective Facts* (1994) 67–68; A Ligertwood, *Australian Evidence* (2nd ed, 1993) 71; *Shepherd* v *The Queen* (1990) 170 CLR 573, 575–576 (per Mason CJ).

6 This paper focuses on the underlying correctness of the reasoning in *Chamberlain (No 2)*, rather than the precise form of the direction or the occasions when it should be given; cf *Van Beelen* (1973) 4 SASR 353, 374; *Chamberlain (No 2)* (1984) 153 CLR 521, 537–8 (per Gibbs CJ and Mason J).

7 (1990) 170 CLR 573.

8 Id 579–580 (per Dawson J; Mason CJ, Toohey and Gaudron JJ agreeing; McHugh J dissenting.)

inappropriate to apply the criminal standard of proof at any point prior to the ultimate assessment of guilt.[9]

While the new principles laid down in *Shepherd (No 5)* were more appropriate to the cumulative nature of circumstantial proof, the court failed to clear away the debris flowing from *Van Beelen* and *Chamberlain (No 2)*. Rather than overruling the earlier decisions the High Court suggested that they were consistent with the chain-cable distinction.

The contention of this paper is that the earlier judgments do not conform with the principles of *Shepherd (No 5)*. The contrast between the two approaches can be understood in terms of two irreconcilable theories of probability: conventional mathematical probability theory and Jonathan Cohen's non-mathematical inductive theory. As a consequence of the High Court's failure to overrule *Van Beelen* and *Chamberlain (No 2)*, the uncertainty in this area of law continues. A number of recent decisions dealing with the consciousness of guilt inference, while paying lip-service to *Shepherd (No 5)* actually embody the principles of *Van Beelen* and *Chamberlain (No 2)*.

II *SHEPHERD (NO 5)*: THE STRUCTURAL DISTINCTION BETWEEN CABLE INFERENCES AND CHAIN INFERENCES

A. *Shepherd (No 5)*

In *Shepherd (No 5)*, the High Court considered the question of whether, as the defendant contended, the facts underlying an inference of guilt needed to be proved beyond reasonable doubt. Giving the leading judgment, Dawson J denied that this was the effect of *Chamberlain (No 2)*. However, he suggested that, in certain cases, it may be worthwhile for the trial judge to draw the jury's attention to what he termed 'intermediate facts'. This will be the case where the intermediate facts constitute:

> indispensable links in a chain of reasoning towards an inference of guilt. Not every possible intermediate conclusion of fact will be of that character. If it is appropriate to identify an intermediate fact as indispensable it may well be appropriate to tell the jury that that fact must be found beyond reasonable doubt before the ultimate inference can be drawn. But where — to use the metaphor referred to by *Wigmore on Evidence* . . . — the evidence consists of strands in a cable rather than links in a chain, it will not be appropriate to give such a warning.[10]

This distinction had been foreshadowed in Deane J's dissenting judgment in *Chamberlain (No 2)*.

> If, for example, the case against an accused is contingent upon each of four matters being proved against him, it is obvious that each of those matters must be proved beyond reasonable doubt. Indeed, it would be appropriate for the presiding judge to emphasise to the jury in such a case that even a

[9] Infra fn 12.

[10] Id 579; cf id 576 (per Mason CJ); the citation is *Wigmore on Evidence* vol 9 (Chadbourn rev, 1981) para 2497, 412–414.

> minimal doubt about the existence of each of those matters would be greatly magnified in the combination of all. On the other hand, if the guilt of an accused would be established by, or a particular inference against an accused could be drawn from, the existence of any one of two hundred different matters, each of which had been proved on the balance of probabilities, it would be absurd to require that a jury should disregard each of them unless satisfied, either in isolation or in the context of all of the facts, that any particular one of those matters had been proved beyond reasonable doubt.[11]

The application of these principles in *Shepherd (No 5)* was straightforward. The case concerned an appeal against a conviction for conspiracy to import heroin. The prosecution relied on evidence of three types. First, two undercover police officers testified that they heard the defendant taking instructions on the running of the operation from a co-conspirator. Secondly, several accomplices gave evidence implicating the accused. Thirdly, there was evidence of financial transactions which, the Crown argued, involved income from the operation. Dawson J concluded that the criminal standard had no application to this underlying cable inference structure.

> The only proper course for the jury to adopt was to consider all the evidence together. [It was unnecessary] for the jury to reach any particular intermediate conclusion of fact in making an inference of guilt on the part of the applicant, other than the obvious one, tantamount to an inference of guilt, that the applicant was engaged in a combination of the kind alleged against him.[12]

B. Metaphors and simple mathematical demonstrations

The cable and chain metaphors give a clear picture of the distinction between the two inference structures and offer support to the rule proposed in *Shepherd (No 5)*. Why should each strand of a cable inference have to bear the entire weight of the criminal standard? Surely some of the weight is being carried by the other strands (figure 1). However where the prosecution relies solely on a chain inference, each link is essential to the inference holding, and each must bear the entire weight (figure 2).

Prior to *Shepherd (No 5)*, Eggleston and Roden J had alluded to the same distinction. However they relied, not on the metaphorical distinction, but on simple mathematical demonstrations. Eggleston, for example, said that the proposition that a fact must be proved beyond reasonable doubt in order to support an inference of guilt beyond reasonable doubt:

> does not accord with the dictates of probability theory, and may result in the rejection of inferences which are fully justified. If a coin is tossed ten times, the probability that at least one toss resulted in a head is one minus the probability that all the tosses resulted in a tail. Assuming the coin is fair, that is to say, that the probability of a tail on each toss is 1/2, the probability that at least one head was tossed is $1-(1/2)^{10}$, that is to say, 0.99902, or

[11] (1984) 153 CLR 521, 627; affirmed *Shepherd (No 5)* (1990) 170 CLR 573, 584–585 (per Dawson J).
[12] Id 586.

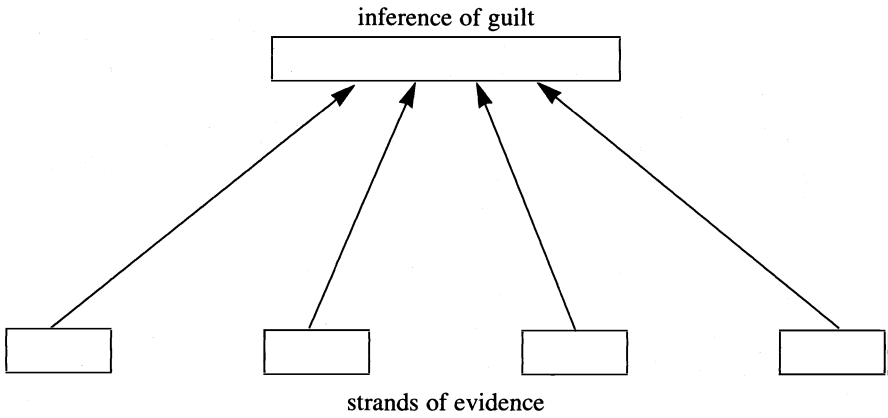

Figure 1: the cable inference structure

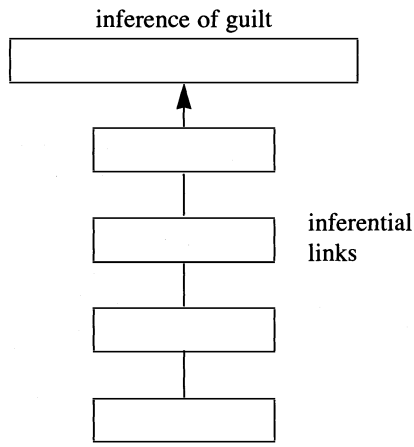

Figure 2: the chain inference structure

> more than 99.9 per cent. So that although we do not know the outcome of any one toss, we can be reasonably certain that at least one head was tossed.[13]

Roden J demonstrated how these calculations would apply to a case where the prosecution's success depended upon the establishment of any one of three primary facts:

> If each one of those three . . . has a 90 per cent probability of being correct, then the degree of probability that at least one of them is, . . . is 99.9 per cent.[14]

[13] R Eggleston, 'Focusing on the Defendant', (1987) 61 *ALJ* 58, 63; cf R Eggleston, *Evidence, Proof and Probability* (2nd ed, 1983), 121–122.
[14] *Shepherd (No 4)* (1988) 85 ALR 387, 394.

Roden J contrasted this situation with one where:

> in the assessment of the jury, all three need to be established before the inference can be drawn . . . If each of the three primary facts has only a 90 per cent probability of being correct, then the degree of probability of the inference is less than 73 per cent. Add a fourth 90 per cent probable primary fact as necessary to the inference and the degree of probability of the inference being correct falls to below 66 per cent.[15]

C. Mathematical modelling

What relevance, it might be asked, do games of chance and mathematics have for forensic proof? As Roden J himself acknowledged:

> The jury function is not performed in such a sterile, theoretical arena, . . . [The] proper performance of that function depends, not only on logic, but also on experience and knowledge and understanding of human behaviour. Degrees of probability and degrees of proof with which juries are concerned are rarely capable of expression in mathematical terms.[16]

Actually there are two related objections here. One draws attention to the typically non-mathematical form of evidence. But while the strength of evidence is rarely susceptible to an objective numerical measure, probability theory may still be applicable. As Ramsey and his successors have demonstrated, via a betting analogy,[17] personal probability judgments can be given a subjective numerical measure conforming to the mathematical rules. On this basis, mathematical models may be constructed which illustrate and offer guidance on human reasoning with uncertainty.[18]

The other objection is that mathematical models tend to be oversimplified, ignoring many complex issues that arise in actual trials. This objection is also misplaced. The purpose of modelling in this instance is not to mirror reality in all its detail, but to present a useful abstraction from it. The models are idealisations which may seldom exist in the real world. But it is this very removal

[15] Id 393. These are applications of the mathematical conjunction and negation rules. Given the probabilities of three independent events, P(E1), P(E2) and P(E3), the probability of all three occuring is given by P(E1&E2&E3)=P(E1) × P(E2) × P(E3), by the conjunction rule. Substituting in P(E1) = P(E2) = P(E3) = 0.90, and then adding the additional term P(E4) = 0.90 gives the results in the second example.

For Roden J's first example we also need the negation rule: P(^E1) = 1 — P(E1). (Read '^' as 'not'.) Now we need to calculate P(E1 or E2 or E3) = 1 — P(^E1& ^ E2&^ E3) = 1 — (1 — P(E1))(1 — P(E2))(1 — P(E3)). The second term on the right hand side was expanded, first using the conjunction rule, and then the negation rule. Again, substituting in the value of 0.90 gives the answer.

[16] Id, 392; cf *Wigmore on Evidence* Vol IA (Tillers ed, 1983) 1034–6.

[17] FP Ramsey, *The Foundations of Mathematics and other Logical Essays* (1931), 156–98; cf B de Finetti, 'Foresight: Its Logical Laws, Its Subjective Sources' (1937), in H E Kyburg, and H E Smokler, *Studies in Subjective Probability* (1964) 99–158; L J Savage, *The Foundations of Statistics* (1954).

[18] Cf P Tillers and E Green, E, (eds) *Probability and Inference in the Law of Evidence: The Uses and Limits of Bayesianism* (1988); various articles in 13 (1991) *Cardozo L Rev.*

from reality that advances our understanding of it.[19] A scribbled map on a piece of paper may be a gross oversimplification of the earth's surface, but it can still guide you to your destination. As the discussion below of *Chamberlain (No 2)*, *Shepherd (No 5)* and subsequent cases will demonstrate, a map of inference structures is sorely needed.

Of course we should proceed cautiously. There are dangers in too enthusiastically embracing a normative theory that is far removed from current practice. Counter-intuitive systems will be difficult to implement, their prescriptions may be intermittently ignored, and they may end up being counterproductive.[20]

However this is not the case here.[21] The line of authority flowing from the judgment of Deane J in *Chamberlain (No 2)* and culminating in the chain-cable distinction developed in *Shepherd (No 5)* demonstrates an implicit appreciation of the operation of the mathematical laws of probability.

D. A Bayesian model of the cable inference structure

A general model of the cable inference structure can be constructed from Bayes's theorem,[22] an increasingly popular, though controversial, tool among evidence scholars.[23]

Bayes's theorem can be understood most readily in its odds-likelihood form. This states that the odds of a proposition, given an item of evidence, is equal to the product of the prior odds of the proposition and the likelihood ratio for the evidence. Symbolically,

[19] 'Every explanatory or heuristic model in natural science has both positive and negative analogies with that which it models. If it had no positive analogy it would be irrelevant, and if it had no negative analogy it would be just another instance of the puzzling process that needs instead to be modelled.' L J Cohen *The Dialogue of Reason: An Analysis of Analytic Philosophy* (1986) 220; cf L J Cohen, *The Implications of Induction* (1970) 4.

[20] R J Allen, 'On the significance of batting averages and strikeout totals: a clarification of the "naked statistical evidence" debate, the meaning of "evidence", and the requirement of proof beyond a reasonable doubt' (1991)65 *Tul L R* 1093, 1110; C R Callen, 'Second-order considerations, weight, sufficiency and schema theory: a comment on Proessor Brilmayer's Theory' (1986) 66 *BUL Rev* 715, 721; L L Lopes and G C Oden, 'The Rationality of Intelligence', in E Eells and T Maruszewski (eds), *Probability and Rationality* (1991).

[21] There appears little agreement among psychologists on the extent to which human judgments accord with mathematical theory. If the predominant view in the 1970s and 1980s was that humans were irrational and non-mathematical (D Kahneman, P Slovic, A Tversky, (eds), *Judgment under Uncertainty: Heuristics and Biases* (1982)), the tide may be turning (G Gigirenzer and DJ Murray, *Cognition as intuitive statistics* (1987)); cf LJ Cohen *The Dialogue of Reason* (1986), 157–164.

[22] Cf R Eggleston, *Evidence, Proof and Probability* (2nd ed, 1983) 206–207. Models built directly from the conjunction rule have been flawed: id 202–206; Cohen, *The Probable and the Provable* (1977) ch10.

[23] G G Aitken and D A Stoney (eds) *The Use of Statistics in Forensic Science* (1991); cf supra fn 18.

$$O(H|E) = O(H) \times L(E) \ldots (1)^{24}$$

O(H) represents the prior odds of the hypothesis, that is, before the evidence is considered. It is the ratio between the probability of the hypothesis, P(H), and the probability of the negation of the hypothesis, P(~H)[25]. Similarly, O(H|E) represents the posterior odds, that is, after the evidence has been considered. It is the ratio between the probability of the hypothesis given the evidence, P(H|E), and the probability of the negation of the hypothesis, given the evidence P(~H|E). L(E) is known as the likelihood ratio. It is the ratio between the probability of finding the evidence if the hypothesis is true, P(E|H), and the probability of finding the evidence if the hypothesis is false, P(E|~H). If we accept, as many have, that evidence is relevant to a proposition if it changes its probability,[26] then the likelihood ratio may be understood as measuring the relevance of the evidence.[27] If the likelihood ratio is greater than one, the evidence is favourably relevant to the hypothesis. If the likelihood ratio is less than one, the evidence is unfavourably relevant to the hypothesis. If the likelihood ratio equals one, the evidence is irrelevant to the hypothesis.

The composition of the likelihood ratio may appear counter-intuitive. It measures relevance in terms of P(E|H) and P(E|~H), the probability of the evidence given, guilt and innocence respectively. Perhaps one would expect relevance to be measured in terms of P(H|E) and P(~H|E), the probability of guilt and innocence respectively, given the evidence.[28] After all, the jury is 'given' the evidence and is ultimately concerned about the degree to which guilt or innocence has been proven. However, a simple example will show that proof is a different matter from relevance, and that the likelihood ratio is an appropriate measure of the latter.

Consider a murder case in which the prosecution provides evidence that the accused was one of four beneficiaries under the will of the wealthy deceased. Clearly this evidence would be favourably relevant to the prosecution's case. And yet, the probability of guilt, on this evidence alone, would be less than the probability of innocence. At strongest the figures might be P(H|E) = 1/4 <

[24] This can be derived from the conjunction rule for conditional probabilities. (Read 'P(H|E)' as, 'the probability of H given E, and '~' as 'not').
$$P(H\&E) = P(E) \times P(H|E) \ldots (1)$$
And also $P(H\&E) = P(H) \times P(E|H) \ldots (2)$
Now if we let (1) = (2), and put P(H|E) on the right hand side, by itself, we get
$$P(H|E) = P(E|\sim H) \times P(H)/P(E) \ldots (3)$$
We can run through the same steps to derive P(~H|E)
$$P(\sim H|E) = P(E|H) \times P(\sim H)/P(E) \ldots (4)$$
Then dividing (3) by (4) gives the odds form of Bayes's theorem.

[25] So to accept odds of 2 to 1 that it will rain today is to assign a 66.7 per cent probability to rain.

[26] G James, 'Relevancy, Probability and the Law' (1941) 29 *Calif L Rev* 689, 699, quoted in R Eggleston, *Evidence, Proof and Probability* (2nd ed, 1983) 80; cf P Gärdenfors, 'On the Logic of Relevance', in J Dubucs, *Philosophy of Probability* (1993) 35–54.

[27] R Lempert, 'Modeling Relevance', 75 *Mich L Rev* 1021, 1025–1026; D Schum and A Martin, 'Formal and Empirical Research on Cascaded Inference in Jurisprudence' (1982) 17 *Law & Soc Rev* 105, 108.

[28] Eg, Williams mistaken construction of the likelihood ratio in terms of P(H|E) and P(~H|E): G Williams, 'The Mathematics of Proof — II' (1979) *Crim L Rev* 340, 347.

$P(\hat{}H|E) = 3/4$. And so, while the terms $P(H|E)$ and $P(\hat{}H|E)$ may represent the proof-value of the evidence, they do not adequately measure its relevance.

Now let us consider the likelihood ratio. Consider first motive evidence which relates to the defendant alone. $P(E|H)$ would, perhaps, be close to one — if the defendant was guilty we would expect to find that she had a motive. $P(E|\hat{}H)$ on the other hand, would ordinarily be much smaller — if the defendant was innocent we would not expect to find motive evidence implicating the defendant alone. The likelihood ratio in this case would be far greater than one, and applying Bayes's theorem, the posterior odds of guilt would become proportionally higher than the prior odds.

Compare, now, the situation posed above where the motive evidence relates, not just to the defendant, but to three others as well. $P(E|H)$ would, perhaps, be unaffected and remain close to one — the defendant's guilt would have little bearing on whether others also had a motive. However $P(E|\hat{}H)$ would be higher than in the previous case — given the defendant's innocence, we would have a higher expectation of finding motive evidence of this kind, which implicates others as well. The likelihood ratio would be greater than one but lower than in the previous case; the evidence would still be favourably relevant to the prosecution, but less so.

How then could additional strands of evidence, such as opportunity and means be incorporated? To model the operation of such a cable inference, equation (1) could be applied a number of times, once for each strand of evidence, E_1, E_2, E_3 and so on. The prior odds, $O(H)$, would be multiplied by the likelihood ratio for the first piece of evidence, $L(E_1)$, to give $O(H|E_1)$. We would then take $O(H|E_1)$ as the prior odds, and, applying (1) again, multiply it by the likelihood ratio for the second piece of evidence, $L(E_2)$, to give $O(H|E1\&E2)$, and so on.

However we can get a clearer picture of a cable inference by reducing these operations to a single step:

$$O(H|E_1\&E_2\&E_3) = O(H) \times L(E_1) \times L(E_2) \times L(E_3) \ldots (2)^{29}$$

Now if the three pieces of evidence are all relevant and favourable to the prosecution case, each likelihood ratio will be greater than one. From the structure of (2) we may observe that the cable inference will be stronger than any of the individual strands, and will increase in strength as the pieces of evidence increase in number and in strength. This endorses the statements of Deane J in *Chamberlain (No 2)* and Dawson J in *Shepherd (No 5)*, that it would be inappropriate to impose the criminal standard of proof on the individual strands of a cable inference.

Nevertheless the model offers an incomplete picture of the structure of criminal proof. In particular, no distinction is drawn between testimony about an event, the event itself, and any inferences that may be drawn from

[29] If the separate strands support the hypothesis of guilt independently the likelihood ratios are defined as above. For the case of dependence L(E1) can remain defined as previously, $L(E2) = P(E_2|H\&E_1) / P(E_2|\hat{}H\&E_1)$, and $L(E_3) = P(E_3|H\&E_1\&E_2) / P(E_3|\hat{}H\&E_1\&E_2)$.

the event.[30] The discussion was in terms of 'strands of evidence', and these details were glossed over. However they form the conceptual basis of the chain inference structure modelled next.

E. A non-Bayesian model of a chain inference

The structure of Bayes's theorem makes it inappropriate for the modelling of chain inferences in the present context.[31] Bayesian analysis focuses on the development of a likelihood ratio[32] which ultimately is to be multiplied by the prior odds of the hypothesis, to give the posterior odds. This raises the question, what are the prior odds?[33] A prior probability of zero means that the accused's guilt is impossible; no amount of favourable evidence will produce a non-zero posterior probability. Yet is it consistent with the presumption of innocence to assign any non-zero figure to the prior probability of guilt? To take as the prior probability the reciprocal of the population of possible offenders is to assume that an offence has been committed, and raises difficult questions about the selection of an appropriate population.[34]

It is beyond the scope of this paper to offer a detailed solution to the prior probability problem. It does not threaten the purely heuristic model of cable reasoning advanced above. Indeed, a corollary of the model is that, as the evidence accumulates, the choice of prior odds becomes less important.[35] However, the question of assigning prior odds cannot be so easily dismissed when we come to model the chain inference. Dawson J in *Shepherd (No 5)* suggested that the criminal standard should apply to the '*indispensable* links in a chain of reasoning towards an inference of guilt'.[36] We are concerned then with cases where there is *no further evidence* and hence *no prior probability*, rendering a Bayesian analysis inappropriate. The question is not how a chain inference changes the prior odds, but rather the strength of a chain inference *per se*.

To develop an appropriate model, let us first consider the nature of the links of a chain inference. Actually, they fall into two categories. One category

[30] D Schum and A Martin, 'Formal and Empirical Research on Cascaded Inference in Jurisprudence' (1982) 17 *Law & Soc Rev* 105, 111.

[31] Cf Id, 107; D Schum, 'Jonathan Cohen and Thomas Bayes on the Analysis of Chains of Reasoning' in E Eells, and T Maruszewski, *Probability and Rationality* (1991), 99, 132–139.

[32] D Schum, 'Probability and the Processes of Discovery, Proof and Choice', in P Tillers and E Green, *Probability and Inference in the Law of Evidence* (1988) 213, 250.

[33] L J Cohen, *The Probable and the Provable* (1977), 107–113; L Jaffee, 'Prior Probability — A black hole in the mathematician's view of the sufficiency and weight of evidence' (1988) 9 *Cardozo L Rev* 967.

[34] G G Aitken, 'Populations and samples', in G G Aitken and D A Stoney (eds), *The Use of Statistics in Forensic Science* (1991), 51, 56–58.

[35] Miller's contrived chaotic distributions that are highly sensitive to the choice of priors, are far removed from the operation of a cable inference in a typical trial (D Miller, 'Diverging Distributions', in J Dubucs, *Philosophy of Probability* (1993), 55–77). I share Miller's doubts, however, about the stronger convergence thesis that, as the evidence accumulates, different Bayesian reasoners will necessarily converge to the same probability assessment, which is therefore attributed a degree of objective endorsement: Id 71; cf L J Cohen, *An Introduction to the Philosophy of Induction and Probability* (1989), 69–70.

[36] (1990) 170 CLR 573, 579 (emphasis added).

comprises assessments of witness credibility. The strength of such a link is given by the probability that an event, E1, occurred, given a witness's testimony that it occurred[37], T1, $P(E_1|T_1)$. The other category comprises assessments of the cogency of evidence. For example, what is the probability that the defendant is guilty of the murder, G, given that they had an argument with the victim a few hours before the killing, E1, $P(G|E_1)$? The ultimate strength of the inference of guilt given the testimony, $P(G|T_1)$, depends on both the witness's credit, the first link, and the cogency of their testimony, the second link. Of course, chains of greater length and complexity may be constructed, for example, by interposing an additional event, such as the formation of a motive, between the argument and the killing, or by separating out the various issues of credibility, such as observational sensitivity, belief formation, and veracity.[38]

Arguably every testimony will involve a basic chain inference with issues of credit and cogency. The more direct and creditworthy the testimony, the shorter and stronger the chain.[39] More remote or circumstantial evidence will involve longer chains, sometimes depending on more than one testimony.[40] Consider, for example, *Van Beelen*, a murder case in which the prosecution relied heavily upon forensic identification evidence. According to the court's analysis, there were four steps involved in the use of forensic evidence, the first two links involving assessments of credit, and the second two, assessments of cogency (figure 3).

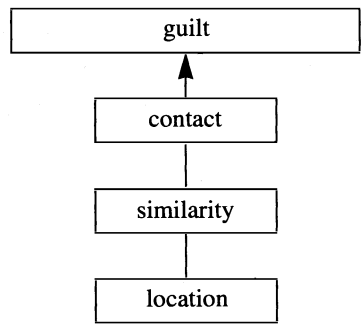

Figure 3: the forensic chain inference in *Van Beelen*

[37] This model takes the testimony as a basic empirical fact. Cf infra, fn 93.

[38] D Schum, 'Jonathan Cohen and Thomas Bayes on the Analysis of Chains of Reasoning' in E Eells and T Maruszewski, *Probability and Rationality* (1991) 99, 107–108.

[39] This is one reason for doubting whether any evidence is truly direct; cf D Hamer, 'The Civil Standard of Proof Uncertainty: Probability, Belief and Justice' (1994) 16 *Syd L R* 506530; M Saks and R Kidd, 'Human Information Processing and Adjudication: Trial by Heuristics' (1980) 15 *Law Soc Rev* 123, 154; L H Tribe, 'Trial by Mathematics: Precision and Ritual in the Legal Process' (1971) 84 *Harv L R* 1329, 1330; infra fn 87.

[40] Cf *The Queen* v *Maleckas* [1991] 1 VR 363, 375 (per McGarvie J).

1. The probability that the police found one set of the trace materials on the victim and the other in connection with the accused, E_1, given the police testimony to that effect, T_1, $P(E_1|T_1)$;
2. The strength and significance of similarities between the sets of trace materials, E_2, given the expert testimony to that effect, T_2, $P(E_2|T_2)$;
3. The probability that the sets of trace materials had a common origin, and that the accused came into contact with the victim, E_3, given the similarities between the sets of trace materials, E_2, and their location, E_1, $P(E_3|E_2\&E_1)$.
4. The probability that the accused was the murderer of the victim, G, given that he came into contact with her, E_3, $P(G|E_3)$.

It should be noted that this chain is structurally slightly different from the argument-motive chain considered above. That chain had a sequential structure, with each link building on the foundation established by the previous link. The finding that the accused had had an argument with the victim a few hours before the killing, formed a basis for the finding that the accused had a motive for the killing, which in turn provided a basis for the finding that the accused killed the victim. The chain inference in *Van Beelen* does not have such an ordered structure. The first link, concerning the location of the trace materials, does not provide a basis for the second link, concerning the similarity between the two sets of trace materials. Rather, the two, independently but conjointly, provide the basis for the third link, the finding of that the accused came into contact with the victim.[41]

Despite this distinction, it appears appropriate for the inference structure such as in *Van Beelen*, to be represented as a chain. Like the simpler, sequential chain based on a single testimony, the ultimate conclusion in this case is dependent upon every link holding. If for example, there was no similarity between the sets of trace materials, or that they were not found in connection with the accused and the victim, the chain would be broken and lose all its strength.

But what is the precise relationship between the strength of the individual links and the overall strength of the chain inference? We have said that the essence of a chain inference is that, for the argument as a whole to succeed, each and every link in the chain must hold. Does this imply that the probability of the chain inference as a whole holding is simply the probability of the weakest link holding? Perhaps this underlies the suggestion of Dawson J in *Shepherd (No 5)* that 'indispensable links in a chain of reasoning towards an inference of guilt ... must be found beyond reasonable doubt.'[42] This interpretation may be formalised as follows.

Let Eg represent the guilt of the accused, the ultimate conclusion flowing from a chain inference built upon a single testimony.[43] Let E_1 represent the giving of the testimony, and E_2, E_3, ... E_i, ... E_g represent the series of successive inferences. Then $P(E_{i+1}|E_i)$ represents the strength of each link in

[41] I am grateful to Edwin Coleman from the Philosophy Department of the University of Melbourne, who helped me to clarify this distinction.

[42] Supra fn 10.

[43] The assumption of a single testimony simplifies the notation, but the analysis will be basically the same where more than one testimony is involved; cf fn 48 infra.

the chain, for i = 1,2, ... g−1. Then the weakest link interpretation implies:

$$P(Eg|E1) = \min_{i=1...g-1} P(E_{i+1}|E_i), \ldots (3)$$

But this interpretation seems too crude. None of the links are certain to hold. This approach conceals our doubts about all of the links, other than the weakest. As Deane J suggested in *Chamberlain (No 2)* 'even a minimal doubt about [each link in the chain] would be greatly magnified in the combination of all.'[44] To similar effect Dixon J commented, 'With any chain of circumstantial evidence ... [the] possibilities of error at all points must be combined and assessed together.'[45] A better approach, taking account of the weakness of all the links, can be modelled using the mathematical conjunction rule.[46] Assuming that the links are conditionally independent of each other,[47] this gives:

$$P(E_g|E_1) = P(E_2|E_1) \times P(E_3|E_2) \times \ldots \times P(E_g|E_{g-1}) \ldots (4)^{48}$$

The two interpretations may differ little in their implications for most cases. Both focus on the strength of the links, rather than their level in the hierarchy. And, given that the criminal standard is the highest standard of proof known to the law, both imply that it should be applied to all the links. In cases involving longer inference chains, however, it may be worthwhile giving the additional cautions of Dixon and Deane JJ.

The models differ more significantly in their implications for civil cases. The weakest link interpretation suggests simply that the civil standard should be applied to each link in the chain. However the conjunctive interpretation would require that each link be proved to some undefined higher standard, the precise level depending upon the length of the chain and the strength of the other links.

[44] (1984) 153 CLR 521, 627.

[45] *Morrison* v *Jenkins* (1949) 80 CLR 626, 644; quoted by R Eggleston, 'The Philosophy of Probability' (1991) *ALJ* 130, 132; cf *US v Ravich* 421 F.2d. 1196 (1970), 1204: 'the length of the chain of inferences necessary to connect the evidence with the ultimate fact to be proved necessarily lessens the probative value of the evidence' quoted in *Wigmore on Evidence* Volume 1A (Tillers ed, 1983), 1034.

[46] Employed by Roden J, supra fn15.

[47] This type of independence should be distinguished from the independence mentioned above in connection with serial and conjoint chain structures. The links of a serial chain, such as the argument-motive chain, may be conditionally independent, in that the strength of the inference of guilt from the presence of motive, P(G|E2), is unaffected by the strength of the inference of motive from the fact of the argument, P(E2|E1).

If we drop the simplifying assumption of independence, the 'chain' inference starts to bear some resemblance to a 'cable' inference with multiple lines of dependence and some links becoming dispensable. Eg, it may be that guilt can be safely inferred from the fact of a witness's testimony, ($P(E_g|E_1)$ approaches 1), even though, or perhaps because, that witness's credibility is severely compromised, ($P(E_2|E_1)$ approaches 0). Cf D Schum and A Martin, 'Formal and Empirical Research on Cascaded Inference in Jurisprudence' (1982) 17 *Law & Soc Rev* 105, 146.

[48] The equation for the chain inference in *Van Beelen*, with its two conjoint foundational links, would be $P(G|T1\&T2) = P(E1|T1) \times P(E2|T2) \times P(E3|E1\&E2) \times P(G|E3)$.

F. On the comparison and combination of cable and chain inferences

We may now draw some comparisons between the chain and cable inference structures. The two are similar in that both derive their strength from their basic building blocks. The stronger the individual strands and links, the stronger, respectively, the overall cable inference and chain inference. However, in other respects, the structures are the converse of each other. For a cable inference, no strand is indispensable, and the addition of even a weak strand will add strength to the cable. By contrast, any lengthening of a chain inference will tend to decrease its strength, and the failure of a single link will mean the failure of the inference as a whole.

This is not to say that the two structures are mutually exclusive. Virtually every testimony will give rise to a chain inference, which ordinarily will be supported by other testimony leading to the same conclusion in a larger cable inference. Conversely, a link in a chain inference may have a cable structure. For example, in *Van Beelen*, the chain of forensic identification evidence leading to the inference of contact and guilt was supported by a separate strand of opportunity evidence. Moreover, the link in the forensic chain inference, establishing contact from the similarity between the sets of trace materials, had a number of strands, one for each pair of matching trace materials (figure 4).[49]

It follows, as Dawson J recognised in *Shepherd (No 5)*, that the criminal standard of proof does not have application to every link of every chain inference, but only '*indispensable* links . . . Not every possible intermediate conclusion of fact will be of that character.'[50] The standard should not be applied to a link where there is an independent strand of evidence, offering

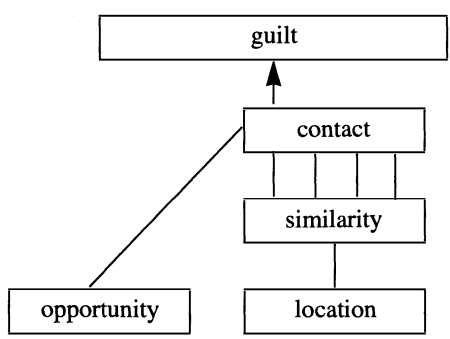

Figure 4: the the proof structure in *Van Beelen*

[49] (1973) 4 SASR 353, 373.
[50] *Shepherd (No 5)* (1990) 170 CLR 573, 579 (per Dawson J) emphasis added; cf id at 576 (per Mason CJ).

support to the same inference or an inference higher in the same structure.[51] In this situation, the link will not need to bear the weight of the criminal standard alone.

A further refinement should be mentioned here. So far, I have spoken in terms of proof beyond reasonable doubt of the criminal offence as a whole. Strictly speaking, however, each element of the offence should be tested against the criminal standard independently.[52] But while it may not be uncommon for defendants to be acquitted due to insufficient proof of a single element, this does not mean that each element will necessarily have its own discrete proof structure. It may be merely that a common proof structure supports the different elements to different degrees.

Consider, for example, how dependent proof of the *mens rea* is upon proof of the *actus reus*, 'for the Devil does not know man's intention'.[53] For example, on a charge of battery, the prosecution may merely tender eyewitness evidence that the defendant struck the victim and invite the jury to infer that the act was deliberate and intentional. Similarly where the *actus reus* of the crime includes a consequence. For example if evidence is adduced that the beating led to the victim's death, and murder is charged, the jury may infer that this component of the *actus reus* was intended on the basis that the accused intended the natural consequences of their act.[54] The defendant may question such inferences, claiming that the act or consequence was accidental or unintended. Motive evidence may then be led to show that the act was indeed intentional. But even then proof of the *actus reus* and *mens rea* will be intertwined: 'the existence of a motive may tend to show either that the person in question did the act *simpliciter*, or that he did it intentionally.'[55]

Beyond this it is difficult to generalise. Much will depend upon the precise definition of the offence, and the nature of the evidence presented. However, for simplicity, in the remainder of the paper I will talk in terms of proof of the offence, as a whole.

[51] Wigmore's chart method may be useful in representing the overall inference structure of more complex cases: J Wigmore *The Science of Judicial Proof* (1937); cf T Anderson and W Twining, *Analysis of Evidence* (1991), 105–155.

[52] Eg, *Shepherd (No 5)* 170 CLR 573, 580 (per Dawson J).

[53] Per Brian CJ (1477) YB 17 Ed IV fo 2, the title of an article by O Briscoe (1970) 44 *ALJ* 23 — 31.

[54] Indeed, for a brief period in the recent history of the English criminal law this inference was held to be irresistible: *DPP v Smith* [1960] 3 All ER 161, [1961] AC 290. This decision provoked considerable controversy. The High Court of Australia elected not to follow *Smith*: *Parker v R* [1963] ALR 524 at 537 (per Dixon CJ), and in the UK, the law was subsequently changed by legislation *Criminal Justice Act* 1967 (UK), s 8.

[55] *Mutual Life Insurance Co of New York v Moss* (1906) 4 CLR 311, 317 (per Griffith CJ).

III *VAN BEELEN* AND *CHAMBERLAIN (NO 2)*: THE SUPPOSED HIERARCHICAL DISTINCTION BETWEEN FACTS AND INFERENCES

A. The *Van Beelen* proposition and its adoption in *Chamberlain (No 2)*

Shepherd (No 5) appeared to constitute a restatement of the law relating to the structure of criminal proof. In previous cases the courts had spoken in terms, not of a structural distinction between cable and chain inferences, but of a hierarchical distinction between facts and inferences. In *Van Beelen* for example, the South Australian Supreme Court held it to be 'an obvious proposition in logic, that you cannot be satisfied beyond reasonable doubt of the truth of an inference drawn from facts about the existence of which you are in doubt.'[56] At another point they said:

> There is a clear distinction between drawing an inference of guilt from a combination of several proved facts, none of which by itself would support the inference, and drawing an inference of guilt from several facts whose existence is in doubt. In the first place the combination does what each fact taken in isolation could not do; in the second case the combination counts for nothing.[57]

In *Van Beelen*, the court applied this hierarchical distinction to the forensic chain inference discussed in the previous section. It held that the trial judge had been in error in directing the jury that the inference of guilt was open despite any doubts they may have held about where the two sets of trace materials had been found or the degree of similarity between them.[58]

Van Beelen was affirmed by a majority of the High Court in *Chamberlain (No 2)*.[59] The most influential judgment was that of Gibbs CJ and Mason J, who held that:

> The jury . . . can draw an inference of guilt from a combination of facts, none of which viewed alone would support that inference. Nevertheless the jury cannot view a fact as a basis for an inference of guilt unless at the end of the day they are satisfied of the existence of that fact beyond reasonable doubt.[60]

The facts of *Chamberlain (No 2)* are notorious.[61] The defendant, Lindy Chamberlain was charged with the murder of her two month old baby, allegedly cutting the baby's throat in the family car. The majority held that evidence that traces of baby's blood had been found in the car should not have been used as a basis for an inference of guilt, in view of the doubts raised by the defence about the prosecution's blood tests.[62]

[56] (1973) 4 SASR 353, 379.
[57] Id 374.
[58] Infra fn 76.
[59] (1984) 153 CLR 521, 536–7 (per Gibbs CJ and Mason J), 570 (per Murphy J),599 (per Brennan J); contra, id 626–627 (per Deane J).
[60] Id 536.
[61] These events spawned a nationwide debate, a best-selling book (John Bryson, *Evil Angels* (1985)), and a movie, based on the book, starring Meryl Streep.
[62] Infra fn 77.

And yet, in *Shepherd (No 5)*, Dawson J, with the agreement of a majority of the High Court, suggested that these propositions and their application in *Van Beelen* and *Chamberlain (No 2)*, were consistent with the structural distinction between chain and cable inferences. For a number of reasons, this claim is untenable.

B. The inconsistency between *Shepherd (No 5)* and *Chamberlain (No 2)*

First of all, Dawson J's claim of consistency with the earlier cases fails to take account of history. In 1983, Eggleston criticised the hierarchical approach in *Van Beelen* advocating instead the structural distinction between cable and chain inferences, though not in these precise terms.[63] In *Chamberlain (No 2)*, Gibbs CJ and Mason J rejected Eggleston's criticisms, explicitly favouring the hierarchical approach advanced in *Van Beelen*.[64] It is only the dissenting judgment of Deane J [65] in *Chamberlain (No 2)* that is consistent with the Eggleston view.[66]

Secondly, and most importantly, Dawson J's interpretation of the earlier cases is at odds with the text of the judgments. He referred to the passage from the judgment of Gibbs CJ and Mason J in *Chamberlain (No 2)*, quoted above, and suggested that:

> It is, I think, quite plain that, in saying that a 'fact as a basis for an inference of guilt' must be proved beyond reasonable doubt, their Honours are referring to an intermediate fact which is a necessary basis for the ultimate inference.[67]

The suggestion is that Gibbs CJ and Mason J were applying the criminal standard to an indispensable link of a chain inference. But there is no support for this interpretation. On the contrary, Gibbs CJ and Mason J describe the operation of a cable inference and suggest that the criminal standard 'nevertheless' applies to the facts underlying that inference. Gibbs CJ and Mason J make no reference to necessary or indispensable facts in this passage. The clear implication is that the hierarchical rule requires the proof, beyond reasonable doubt, of each and every primary fact, whether existing in a chain or a cable inference structure. However the criminal standard is not to be applied to the individual inferences drawn from the primary facts. The strength of all inferences supporting the conclusion of guilt should be considered together.[68]

The hierarchical approach in *Chamberlain (No 2)* draws a sharp distinction, not between cable inferences and chain inferences, but between 'primary

[63] R Eggleston, *Evidence, Proof and Probability* (2nd ed,1983), 121.
[64] (1984) 153 CLR 521, 538.
[65] Supra fn 11.
[66] In *Shepherd (No 5)* Dawson J quoted from the judgment of Deane J in *Chamberlain (No 2)* at length without noting its inconsistencies with the other judgments: (1990) 170 CLR 573, 584–585.
[67] (1990) 170 CLR 573, 581 (per Dawson J).
[68] The quotation from *Van Beelen* (supra fn 57) has the same structure. A description of a cable inference, then the requirement that the underlying facts be proven beyond reasonable doubt.

facts' and inferences. This is also apparent in the following passage from the judgment of Brennan J:

> The primary facts from which the inference of guilt is to be drawn must be proved beyond reasonable doubt . . . An inference of guilt may properly be drawn although any particular primary fact, or any concatenation of primary facts falling short of the whole, would be insufficient to exclude other inferences.[69]

Dawson J struggles to construe this passage consistently with the chain-cable distinction. He makes the implausible suggestion that Brennan J uses the term 'primary fact' with two inconsistent meanings within the one brief passage. According to Dawson J, the first reference:

> is clearly a reference to such intermediate conclusions of fact as are necessary for the drawing of an inference of guilt and is not a reference to each basic fact — each individual item of evidence — upon which those conclusions may be based.[70]

The second reference to a 'primary fact' however, Dawson J does 'not take . . . to mean anything more than a piece of evidence.'[71] In other words Brennan J uses the term to refer, first to the indispensable links of a chain inference, and secondly to the strands of a cable inference. But if Brennan J intended to draw this contrast, surely he would have used different terms. A much more credible interpretation is that Brennan J, like Gibbs CJ and Mason J, considered that primary facts are to be proved beyond reasonable doubt, even where they support a cable inference.[72]

This hierarchical interpretation of *Chamberlain (No 2)* is reinforced by the suggestion of Gibbs CJ and Mason J that the *Van Beelen* proposition applies to civil cases as well as criminal. Following the proposition that the facts underlying an inference must be proved beyond reasonable doubt, they add, 'in a civil case the circumstances must raise a more probable inference in favour of what is alleged.'[73] The suggestion is that foundation of fact must be established beyond reasonable doubt in a civil case also, however the inference need then only satisfy the civil standard of proof. This is consistent only with a hierarchical analysis. As demonstrated above,[74] on the weakest link interpretation of the structural rule, the civil standard would apply to each link in the chain, while on the conjunctive interpretation, some higher indefinite standard would apply.

Finally, Dawson J's reading of *Chamberlain (No 2)* and *Van Beelen* is inconsistent with the way in which the 'proposition of logic' was applied in

[69] (1984) 153 CLR 521, 599; cf id 537–538, 559 (per Gibbs CJ and Mason J).
[70] (1990) 170 CLR 573, 584.
[71] Ibid.
[72] Consider also the following passage from the judgment of Brennan J, not quoted by Dawson J: 'An inference of guilt can safely be drawn if it is based upon primary facts which are found beyond reasonable doubt and if it is the only inference which is reasonably open on the whole body of primary facts.' *Chamberlain (No 2)* (1984) 153 CLR 521, 599.
[73] (1984) 153 CLR 521, 536.
[74] Supra, text following fn 48.

those cases. In neither case was the criminal standard applied to facts that constituted 'an intermediate fact as an *indispensable* basis for an inference of guilt'.[75] The facts may have been links in chain inferences, but these chains were strands of larger cable inferences.

In *Van Beelen*, the court applied the criminal standard to the forensic chain inference even though it was supported by an independent strand of opportunity evidence. Moreover, the criminal standard was imposed hierarchically, on only the first two links — the two sets of trace materials having been found in connection with the accused and the victim, and the similarity between them.[76] There was no suggestion that the criminal standard applied to the *inference* of contact drawn from those *facts*.

In *Chamberlain (No 2)* the High Court, on appeal, was concerned with whether the jury verdict of guilt was unsafe. Gibbs CJ and Mason J were unwilling to base an inference of guilt on the presence of traces of baby's blood in the car:

> The conflicting evidence should have raised a doubt in a reasonable mind, and there is no other evidence that can resolve the doubt before a decision on the verdict is ultimately reached.[77]

It is clear that Gibbs CJ and Mason J did not consider this to constitute an indispensable link in the reasoning leading to the accused's guilt. Gibbs CJ and Mason J, for the same reason, put to one side a *second strand* of evidence relating to a bloody handmark that a witness claimed to have seen on the baby's jumpsuit.[78] (This evidence was inconsistent with the defence account that the baby had been taken by a dingo.) And it is clear, as McHugh J pointed out in *Shepherd (No 5)*,[79] that Gibbs CJ and Mason J considered there to be further *strands* of evidence, since they ultimately held that there was a sufficient basis for the guilty verdict.[80] They formed a majority with Brennan J in dismissing the appeal.[81]

C. The conceptual difficulty in identifying 'facts'

Perhaps one reason for the lingering uncertainty about the meaning of *Chamberlain (No 2)* is a conceptual difficulty in identifying the 'facts' which must be proved beyond reasonable doubt before an inference may be drawn

[75] *Shepherd (No 5)* (1990) 170 CLR 573, 576 (per Mason CJ), emphasis added.
[76] (1973) 4 SASR 353, 380.
[77] (1984) 153 CLR 521, 559.
[78] Id 568.
[79] (1990) 170 CLR 573, 591.
[80] (1984) 153 CLR 521, 569.
[81] Ultimately the forensic evidence was shown to be fundamentally flawed, and the verdicts were set aside in the Morling Royal Commission, in 1987. Some of the 'blood' ultimately turned out to be Dulux Dufin 1081, a sound deadening compound, and other traces were found to be copper dust. (K Crispin, 'An Australian Witch-hunt', *The Weekend Australian* 16 December 1995, 23) This left the question open what had happened to Azaria. The most recent coronial enquiry, in December 1995, concluded that a dingo was not the culprit, and the case continues to be the subject of public debate: cf G Alcorn, 'Azaria: the final verdict draws to an end', *The Age*, 25 November 1995, pA19; J Bryson, 'Dingo Justice', *The Age* 16 December 1995, A13; A McGregor, 'The macho conspiracy against Lindy', *The Weekend Australian* 16 December 1995, 1.

from them. Consider the different approaches taken in the Victorian case of *The Queen* v *Maleckas*.[82]

In *Maleckas* the defendant, charged with murder, admitted that he had killed the victim with a kitchen knife, but argued that the killing was in self defence. The prosecution rebuttal rested upon the contention that the victim was sitting down when the defendant attacked. In establishing this fact, the prosecution relied upon eyewitness testimony and several items of circumstantial evidence: the chair on which the victim was said to have been sitting was broken, and had blood on it; there were other blood stains in that area; the police found the knife on the floor near the chair; and expert evidence that the angle of the knife wounds suggested that the victim had been seated. The question raised on appeal was whether the trial judge was correct in directing the jury that:

> Not every matter of fact has to be proved beyond reasonable doubt . . . [For example] the opinions [of] the doctors as to the position that the deceased man was in when he was stabbed need only be proved on the lesser standard of proof, the balance of probability.[83]

Murphy J held that the direction was not incorrect. *Chamberlain (No 2)* required the jury to be satisfied beyond reasonable doubt that the accused was seated when the attack took place, a conclusion which derived support from a number of pieces of evidence.[84] It did not matter that the expert evidence relating to the knife wound, or any other piece of evidence, was only established on the balance of probabilities.

McGarvie J, on the other hand, considered that the criminal standard had application, not just to the inferred fact that the victim was seated, but to the facts from which that fact was inferred. He considered that the direction was flawed, particularly with respect to the medical evidence,[85] and that, on the contrary, a *Chamberlain (No 2)* direction should have been given. Otherwise, in cases such as this:

> a jury might be expected to be prone to conclude it could be satisfied beyond reasonable doubt that the accused committed the crime if a sufficient number [of items of circumstantial evidence] were established to their satisfaction, though some or all of the items were established to less than . . . beyond reasonable doubt.[86]

The third judge, Brooking J took a different approach again. He considered that *Chamberlain (No 2)* had no application to cases, such as the present,

[82] (1991) 1 VR 363. *Maleckas* was decided prior to *Shepherd (No 5)*, though it appeared later in the reports.

[83] Id, 367.

[84] Id 371.

[85] Id 380–381. There are suggestions in the judgment of McGarvie J, that he, like Murphy J, was centrally concerned that it had been proved beyond reasonable doubt that the defendant was seated. The reason that he imposed the standard to the underlying expert evidence was the risk that the jury would rely on that evidence alone (Id 381). However this might always be a possibility, and the reasoning of McGarvie J boils down to an application of the criminal standard at the lowest level of fact.

[86] Id 380.

where there was direct evidence.[87] But the distinction between direct and circumstantial evidence is highly problematic. First, factual testimony will rarely, if ever, provide an immediate answer to a legal question. In this case, would 'the victim was sitting' necessarily equate with 'the killing was not in self defence'? And secondly, there will always be some scope for questioning the credit of the witness. The eyewitness in *Maleckas* was a former lover of the deceased, and that evening had consumed alcohol, heroin and sleeping pills.[88]

Putting the judgment of Brooking J to one side, which application of the law was more correct? Certainly the decision of Murphy J was more consistent with the chain-cable distinction subsequently developed in *Shepherd (No 5)*. The fact that the victim was sitting constituted an indispensable link in the prosecution case and required proof beyond reasonable doubt. But the criminal standard had no application to the strands of evidence supporting this inference.

But which judgment in *Maleckas* applied the hierarchical rule from *Chamberlain (No 2)* properly? The concepts invoked by *Chamberlain (No 2)* are not sufficiently well defined to provide a definitive answer. Against Murphy J, it could be said that *Chamberlain (No 2)* is authority that an inference could not be drawn without a solid foundation in fact. In *Maleckas*, as in *Chamberlain (No 2)*, the expert evidence did not provide a sufficiently solid basis. On the other hand, it could be said that McGarvie J confused the notions of evidence, fact and inference. Any doubts about the facts to be derived from the expert evidence in this case were not fatal since they were resolved by the other evidence. In *Maleckas* the term 'fact' applied equally well to *the victim was seated* as it did to *the position of the knife wound indicated that the victim was seated.*

D. Roden J's semantic games with 'facts'

In the NSW Criminal Court of Appeal decision, *Shepherd (No 4)*,[89] Roden J developed a disruptive critique of *Chamberlain (No 2)*. He argued that the notion of 'fact' was so indefinite as to render the hierarchical test ineffectual. By adjusting the level of abstraction it would always be possible to identify a fact that has been proved beyond reasonable doubt, and thus satisfy the test.

Consider, for example,[90] a robbery case in which the prosecution relies on three independent witnesses, testifying respectively that they saw the accused preparing for the crime, committing it, and running from the scene of the crime. The jury has lingering doubts about each identification, but considers that at least one of the three must be right, and that any of the testimonies, if reliable, would support a finding of guilt. But would *Chamberlain (No 2)* preclude reliance on any of the identifications? Applying the hierarchical test

[87] Id 382.
[88] Id 370.
[89] (1988) 85 ALR 387.
[90] This example is adapted from the judgment of Roden J: id 392–393, 402.

to each testimony in turn, there would be no fact established beyond reasonable doubt, capable of supporting the inference of guilt.

However at a more abstract level, the jury is satisfied beyond reasonable doubt of the fact that at least one of the three witnesses correctly identified the accused in connection with the crime. Would this provide a sufficiently solid foundation for the inference of guilt? While the defence counsel apparently considered this strategy valid in *Shepherd (No 4)*[91] such a concatenation may be considered inconsistent with the concern expressed in *Van Beelen* and *Chamberlain (No 2)* with the foundations of inference.

Perhaps this concern could be accommodated by selecting a fact at a more empirical level. It might be said that the jury is satisfied beyond reasonable doubt that each of the witnesses saw a person resembling the accused committing acts in connection with the robbery. And from these proven facts it may be inferred beyond reasonable doubt that the accused committed the robbery. This may appear a sensible solution. But would it be subject to the precise form of witness testimony? What if the witness is adamant that they saw the accused, not merely someone resembling the accused?

This difficulty could be overcome by applying the test at a still more basic level, requiring that the jury be satisfied beyond reasonable doubt that each witness did in fact give the testimony appearing in the transcript.[92] But this would render the hierarchical test ineffectual. On this basis *Chamberlain (No 2)* would be satisfied in virtually every case.[93] In addition, the objection may be raised that I am straying into the realm of evidence, recalling that Gibbs CJ and Mason J suggested that facts should be looked at 'in the light of the whole evidence'.[94]

E. McHugh J's extended hierarchical rule

As noted above, McHugh J in *Shepherd (No 5)* disagreed with Dawson J's structural analysis of *Chamberlain (No 2)*. While considering *Chamberlain (No 2)* inconsistent with the cumulative nature of circumstantial proof McHugh felt constrained by authority to follow it.[95] In so doing he highlighted another uncertainty about the meaning of *Chamberlain (No 2)*. According to McHugh J:

[91] Id 402.
[92] Id 393.
[93] The chain inference model advanced above, assumes that it is certain the underlying evidence does exist in a particular form, supra fn 37. A sceptic may question even this assumption. After all the 'very act of perception itself . . . involves a complex set of hidden inferences to which we are predisposed by our experience and by the structure of our mind or thought.' (*Wigmore on Evidence* Vol 1A (Tillers ed, 1983), 1036, fn 16; cf J Bronowski, *The Origins of Knowledge and Imagination* (1978), 13–16.) At the very foundation of the superstructure of inference are the photons and sound waves emanating from the witness and striking the sensory organs of the jury. But for present purposes, it appears unnecessary to extend our analysis below the level of human consciousness.
[94] (1984) 153 CLR 521, 536.
[95] (1990) 170 CLR 573, 592–3. He sought to limit the scope of the authority of *Chamberlain (No 2)* to 'cases concerned with whether a verdict, based on circumstantial evidence, is unsafe or unsatisfactory.' It is no authority on trial judge directions (id 594).

> the majority judgments . . . intended to assert that in a criminal case, any fact — primary or intermediate — relied upon as the basis for drawing the inference of guilt must be proven beyond reasonable doubt.[96]

This is certainly a sensible interpretation of *Chamberlain (No 2)*. First, it makes less important the difficulty, noted above, of distinguishing between evidence, facts and inferences. McHugh J suggests that the criminal standard is to be applied at all levels. Secondly, although a signally strict approach to criminal proof, this is more consistent than to apply the criminal standard arbitrarily at a single level. Consider two murder cases in which the prosecution seeks to establish contact between the accused and the victim. In the first case contact is a primary fact established by the testimony of an eyewitness that the two were seen together. In the second case contact is an intermediate fact inferred from forensic evidence of the type considered in *Van Beelen*. In both cases guilt is to be inferred from the fact of contact, together with a lack of evidence of opportunity on the part of anyone other than the accused. It would clearly be inconsistent to apply the criminal standard of proof to the eyewitness account of contact in the first case, while not applying it to the inference of contact in the second. In both cases the fact of contact plays the same role in providing a foundation for the ultimate inference of guilt.

However, as McHugh J accepted in a different context, 'a case is only an authority for what it actually decides . . . [and not] for a proposition that may seem to follow logically from it.'[97] Though logical, the interpretation of McHugh J regarding the hierarchical rule is at odds with how it was applied in *Chamberlain (No 2)*. Gibbs CJ and Mason J excluded from consideration the '*primary fact* that the blood was foetal blood'[98] and that a bloody handmark had been seen on the jumpsuit,[99] as these had not been proven beyond reasonable doubt. They took account of opportunity and 'other *facts* which were established beyond reasonable doubt'.[100] However they also took account of *inferences* from the established facts that had apparently not reached the criminal standard:

> The blood-stains on the clothing made it *probable* that the baby died as a result of a cut or other incised wound to the throat. The small amount of blood in the tent . . . made it *improbable* that a dingo had seized Azaria while she lay asleep in the tent . . .; the *probable* object of interference [with Azaria's clothes] was to make it appear [they had] been dragged and torn by a dingo.[101]

Admittedly it is difficult to understand why these constituted inferences to which the criminal standard had no application, while the foetal blood, bloody handmark and opportunity were considered primary facts to which

[96] Id 590.
[97] *Quinn v Leathem* [1901] AC 495, 506 (per Earl of Halsbury LC); *Shepherd (No 5)* (1990) 170 CLR 573, 593.
[98] (1984) 153 CLR 521, 559 (emphasis added).
[99] Supra fn 78.
[100] Id 567 (emphasis added).
[101] Ibid (emphasis added).

the criminal standard did apply. This is another illustration of the arbitrary nature of the hierarchical rule.

F. The inductive probability theory of Jonathan Cohen

In the first section, mathematical models were presented which confirmed the structural distinction between cable and chain inferences. In the present section we have seen the potentially arbitrary operation of the hierarchical approach. This raises the question, if the argument in favour of the structural analysis is so clear, how did the hierarchical approach ever gain support? Jonathan Cohen presents an alternative non-mathematical account of probabilistic reasoning which provides a rationale for the rule developed in *Van Beelen* and *Chamberlain (No 2)*. However I will argue that, like these cases, the inductive theory deals inadequately with the cable inference structure.

The mathematical rules of probability were derived from the study of games of chance, such as coin tosses and card games.[102] In such cases probability assessments can be derived as a simple ratio of the number of outcomes of a specified kind to the total number of outcomes, and the mathematical rules follow straightforwardly from this relationship. Cohen's inductive probabilities are not based on the simple enumeration of classes of events. Instead they offer a measure of the variety of circumstances covered by the evidence, which experience has shown to be favourable to the specified outcome. The inductive rules for the negation and conjunction of probabilities which flow from this variative assessment differ markedly from the enumerative mathematical rules.

Cohen's formalisation of the measurement of favourable circumstances is termed the method of relevant variables. In each field of knowledge we can imagine the construction, from experience, of a list of relevant variables. Each relevant variable is 'a set of circumstance-types . . . [where] each of its variants, or circumstance-types, suffices to falsify at least one generalisation in that field.'[103] Moreover, this list of relevant variables is ordered according to the importance of the relevant variable. 'The greater the variety of types of hypotheses that a particular relevant variable is seen to falsify, the more important it will normally be presumed to be' and the higher it will be in the list.[104] A probability assessment in a particular case is a matter of counting, in order of importance, the variants present in the evidence, favourable to the specified outcome.[105] The more favourable variants appearing in the evidence, the higher the inductive probability.

For example, we may know from experience many factors which affect the likelihood of rain.[106] In a particular case, given the presence of dark clouds, the falling barometer, the presence of an offshore wind, and the fact that it is mid-winter, we may consider that we have favourable variants for the first

[102] L J Cohen, *An Introduction to the Philosophy of Probability and Induction* (1989), 13–27.
[103] LJ Cohen, *The Probable and the Provable* (1977), 131–2.
[104] Id 141; cf fn 26, 156.
[105] Id 202.
[106] Id 203–7.

four relevant variables, and so assign an inductive probability four to the hypothesis of rain, $P_I(H|E) = 4$, or 4/n where n is the total, usually unknown, number of relevant variables.

G. An inductive model of the cable inference structure

Like the hierarchical approach developed in *Van Beelen* and *Chamberlain (No 2)*, inductive probability theory offers an inadequate representation of the cable inference structure. According to Cohen,[107] the strands of a cable inference are effectively represented as the favourable variants of the variables relevant to the hypothesis in question.[108]

> So, if we are assessing the inductive probability that a particular man with a motive for murdering the victim was actually the murderer, one relevant circumstance that raises the probability is the fact that this man had a good opportunity to commit the crime.[109]

This offers few insights into the structure of a cable inference. First, it provides only a crude picture of how the inference gains strength as the number of strands increases. It should be noted that an inductive probability assessment provides much less information than a mathematical assessment. For example $P_I(H|E) = 2/4$ states merely that the evidence reports two circumstances favourable to the hypothesis out of four known relevant circumstance-types. Unlike the equivalent mathematical statement, it does not tell us, for example, that rain (H) can be expected in half the cases that the dark clouds and a falling barometer (E) are observed. Nor, given $P_I(H|E_1) = 1/4$ and $P_I(H|E_2) = 2/4$, could we say that rain (H) is twice as likely on the evidence of dark clouds and a falling barometer (E_2), as it is on evidence of dark clouds alone (E_1).[110] Inductive probability operates on a merely ordinal scale, whereas mathematical probabilities may be interpreted as rational or real numbers.[111] As Cohen acknowledges, 'in this sense . . . inductive probability [is] in principle unmeasurable.'[112]

Secondly, Cohen's picture of the cable inference offers no gradation of the strength of the strands. Favourable variants are either present or absent. There is no middle ground. The model is no more than 'a list of the various points that all have to be established, and of the various let-outs that all have to be barred in relation to each element in the crime.'[113] In short, a mere checklist, each item of which is either ticked or crossed.

We might attempt a more detailed analysis, in which the various strands of support were treated, not merely as items on the checklist of relevant

[107] Cohen claimed that mathematical theory could not adequately depict the cable inference, or as he termed it 'corroboration and convergence'. However he did not consider Bayes's theorem: id 93–115.

[108] Id 279–80.

[109] Id 278–9.

[110] That would be to accord equal importance to both relevant variables whereas the presence of dark clouds is a variant of the first, and hence most important relevant variable. Cohen offers no measure of importance.

[111] Id 229.

[112] Id 229.

[113] Id 273.

variables, but as hypotheses with their own measures of strength. The problem is that these disparate measures could not then be combined into a measure of support for the overall case. The various inference strands, motive, opportunity, etc, will have been tested against different lists of relevant variables. Each measure will be *sui generis*, on a scale of its own, and incommensurable with any other measure, 'except in respect of their limiting values.'[114] Only if the various strands had full inductive support, could they enter as favourable variants in the inductive measure of the larger cable inference. Like *Van Beelen* and *Chamberlain (No 2)*, the inductive theory shows insufficient recognition that even weak strands can combine to produce a strong cable inference.

This lack of appreciation of the cumulative nature of circumstantial proof is also reflected in Cohen's treatment of the criminal standard of proof. According to Cohen, proof beyond reasonable doubt equates with 'full inductive support'.[115] The evidence before the court must contain a favourable variant of every variable relevant to guilt. Consider the following illustration provided by Cohen:

> The facts before the court may not suffice to establish whether the defendant had a motive for committing the murder of which he is accused, and he can hardly be convicted on circumstantial evidence if there is no known motive.[116]

Cohen has been misled by the shortcomings in his own theory. Defendants *are* convicted in circumstantial cases without motive evidence.[117] The absence may be a matter of concern to the court but will not necessarily be fatal to the prosecution case.

To model the nature of circumstantial proof we need the greater sophistication of the mathematical theory. The overall strength of a circumstantial case is determined not merely by counting the number of strands to the cable inference, and proof beyond reasonable doubt does not equate with a complete set of strands. It is necessary to also take account of the strength of the individual strands. In some cases, though motive evidence is absent, the other strands will be strong enough that the cable as a whole still satisfies the criminal standard.

H. A Inductive model of the chain inference structure

The mathematical chain inference model was built using the conjunction rule. This rule demonstrates, for example, how uncertainty about the credit of a witness, combines with uncertainty about the cogency of their testimony.

[114] Id, p 268; cf id 143; cf L J Cohen, 'The Logic of Proof' [1980] *Crim L Rev* 91, 96.
[115] L J Cohen, *The Probable and the Provable* (1977), 272.
[116] Ibid.
[117] *Chamberlain (No 2)* diverges from the inductive theory in this respect: (1984)153 CLR 521, 565 (per Gibbs CJ and Mason J); cf *Plomp* (1963) 110 CLR 234, 249–250 (per Menzies J); *Askeland* (1983) 18 A Crim R 103; *Neilan* (1991) 52 A Crim R 303; *Jeffrey* (1991) 60 A Crim R 384.

However, the inductive theory does not allow the development of a corresponding conjunctive model. Because of the commensurability limitation, the inductive conjunction rule applies only where the propositions to be conjoined are 'in the same field of inquiry, [and] open to the same series of tests as one another.'[118] Cohen's theory could tell us the level of inductive support for the proposition that *bee-populations can discriminate between tastes* and *butterfly-populations can discriminate between smells.*[119] But it could not tell us the inductive probability that *we will reach the next petrol station* while *it is still open.*[120]

> From such mutually incommensurable probabilities no transitive inference emerges. Inductive probability-functions evaluate the weight of relevant evidence, and what is relatively weighty for one type of conclusion may not be nearly so weighty for another.[121]

Rarely then, will the inductive conjunction rule have application to chain inferences in criminal trials.[122] It would not, for example, operate in *Van Beelen* to combine the various uncertainties about the police testimony, the expert forensic testimony, the inference of contact, and the inference of guilt.

However another inductive analysis of chain inferences is open which offers an explanation of the rationale behind *Van Beelen* and *Chamberlain (No 2)*. As Ligertwood explains, 'the inductivist . . . cannot find the final stage proved beyond reasonable doubt unless each of the preceding stages are first proved to that same standard.'[123] Cohen suggests that at each stage of the chain inference, the fact or inference must be 'reasonably certain . . . [so that it can] be detached as a known or accepted fact which can provide a premise for further proof.'[124]

Consider again a simple chain inference built on the testimony that the accused had an argument with the victim shortly before the killing. According to Cohen, the jury would need to be convinced beyond reasonable doubt that the argument had in fact occurred, to proceed any further with the inference; $P_I(E_1|T)$, representing the witness's credit, would need to be maximal in order for E_1 to be detached and used for the second inference.[125] Similarly, the jury

[118] L J Cohen, *The Probable and the Provable* (1977), 169. The inductive conjunction rule states: 'If $P_I[S',R] \leqslant P_I[S',R]$, then $P_I[S\&S',R] = P_I[S,R]$.' (Id 221). This constitutes the 'weakest link' approach criticised above, fn 44.

[119] Id 132.

[120] Id 266.

[121] Id 269.

[122] Ligertwood appears to suggest the inductive conjunction rule *is* applicable: 'inductivists argue that, after a fact has been broken down into its essential inferences, the overall probability of guilt can be no higher than the lowest probability of those material and essential facts upon which the case as a whole depends.'(A Ligertwood, *Australian Evidence* (2nd ed, 1994) 69; cf id 29.)

[123] Id 70.

[124] LJ Cohen, *The Probable and the Provable* (1977), 269.

[125] Cohen gives an inductive account of the assessment of credit in which he recognises the many possible sources of doubt (id 251–252). He does not dwell on the implication of these doubts, that the evidence is unusable unless they are all resolved. Note more recently Schum developed an inductive analysis of credit, suggesting that propositions may be detached without full support (D Schum, 'Jonathan Cohen and Thomas Bayes

would need to be sure, from the fact of the argument, that the accused possessed a motive to kill the victim in order to proceed to the final stage; $P_I(E_2|E_1)$ would need to be maximal in order for E_2 to be detached. And since this is a criminal trial, the inference of guilt from the existence of motive, $P_I(G|E_2)$ would also need to be maximal.

This model provides a rationale for the hierarchical reasoning in *Van Beelen* and *Chamberlain (No 2)* where the courts held that the underlying fact must be fully proved from the evidence before an inference could be drawn.[126] Actually the inductive account differs slightly in that the criminal standard would be applied to every link of the chain rather than just the 'primary fact' link. However, as noted in the discussion of the judgment of McHugh J in *Shepherd (No 5)*, logically the hierarchical rule would have this extended operation.

In two further respects, the inductive model is in accord with *Chamberlain (No 2)* and *Van Beelen.* First, the model would have application to chain inferences in civil cases, the only difference being that the last link leading to liability need only be proved on the civil standard.[127] As discussed above, Gibbs CJ and Mason J in *Chamberlain (No 2)*[128] seemed to consider this the logical implication of the *Van Beelen* proposition. Secondly, on the inductive approach, the criminal standard should be applied to each link of a chain inference, even where the chain inference constituted a strand of a larger cable inference.[129] As noted above, the hierarchical rule was applied on this basis in both *Chamberlain (No 2)* and *Van Beelen.*

Ligertwood suggests that 'nothing said by the court in *R v Van Beelen* was intended to interfere with this [cumulative] process of circumstantial proof'.[130] This may be true, but whatever their intentions the inductive logic of the *Van Beelen* proposition was intrinsically inconsistent with the nature of circumstantial proof. In this respect Ligertwood is correct in drawing 'the clear conclusion from the approaches taken in both [*Van Beelen*] and ... *Chamberlain (No 2)* ... that proof is approached inductively, not mathematically.'[131] However he is wrong to suggest that *Shepherd (No 5)* supports this

on the Analysis of Chains of Reasoning' in E Eels and T Maruszewski (eds), *Probability and Rationality*, 99, 122–4). This met with Cohen's endorsement, as far as civil cases are concerned 'contrary to what may have been suggested' by the passages discussed above ('Some comments by LJC'; id 319, 329). However for criminal cases Cohen would still require 'at each stage of a cascaded inference the level of inductive probability that amounts to proof beyond reasonable doubt' (id 328).

126 Compare the similar view of Michael and Adler, discussed by Tillers, *Wigmore on Evidence* Vol IA (Tillers ed, 1983), 1035. Cf supra fn 93, for Tillers response.

127 Cohen discusses his model in connection with the 'inference upon inference' prohibition, as described by Wigmore, which applied in civil cases. Actually, as Eggleston demonstrated, the law was more unsettled than Cohen represented. However it is interesting to note the closeness of Cohen's interpretation of the rule to the reasoning in *Van Beelen* and *Chamberlain (No 2)*: LJ Cohen *The Probable and the Provable*(1977), 69–70; J H Wigmore, *A Treatise on the Anglo-American System of Evidence in Trials at Common Law*, (3rd ed, 1940), Vol I, b 41, 439; R Eggleston, *Evidence, Proof and Probability* (2nd ed, 1983), 237–240, 261; cf *Van Beelen* (1973) 4 SASR 353, 375.

128 Supra fn 73.

129 Cf supra fn 114.

130 A Ligertwood, *Australian Evidence* (1993), 71.

131 Id 73.

conclusion.[132] On the contrary, to take proper account of the cable structure of circumstantial proof, the High Court was required to replace the crude inductive logic of the hierarchical rule with a more sophisticated structural analysis founded in mathematical probability theory.

IV THE CONTINUING CONFUSION AFTER *SHEPHERD (NO 5)*

The foregoing analysis suggests that, contrary to the claims made by the majority in *Shepherd (No 5)*, the principles advanced in that case were fundamentally inconsistent with the earlier cases. However the earlier cases were not overruled. In this context it should not be surprising to find that the decisions since *Shepherd (No 5)* have shown little improvement in the understanding of the structure of proof.[133] Lip-service has occasionally been paid to the structural analysis of *Shepherd (No 5)*,[134] but the results are generally more consistent with the hierarchical logic of *Van Beelen*. Here I consider just one line of cases, dealing with the consciousness of guilt chain inference.[135]

A. The inference from the accused's lies to consciousness of guilt

The argument that evidence of the accused's lies display a consciousness of guilt, has a chain inference structure. The links are as follows:

1. the accused made the alleged statement;
2. the alleged statement was false;[136]
3. the falsehood was deliberate;[137]
4. the lie was motivated by a consciousness of guilt;[138]
5. the accused is guilty.

This chain inference has a peculiar structure. In effect the prosecution is seeking to base an inference of guilt upon what typically is an exculpatory

[132] 'If this interpretation of *Shepherd* is correct, then the High Court accepts the inductive approach to proof described by Cohen.': id 75.

[133] The confusion flowing from the High Court decisions has been facilitated by the text writers. Ligertwood endorses a model that is consistent with the earlier cases and not with *Shepherd (No 5)* (Ibid). *Cross on Evidence* contains a two page quotation from Gibbs CJ and Mason J in *Chamberlain (No 2)*, and barely one third of a page from *Shepherd (No 5)* (Byrne and Heydon, Service 27, September 1995, Butterworths, b9040, 9019–21).

[134] Occasionally the principles from *Shepherd (No 5)* have been applied correctly, see eg, *Marijancevic* (1993) 70 A Crim R 272, 279 (note however the uncritical approval of *Maleckas*); *Rogerson and Paltos* (1992) 65 A Crim R 530, 545; cf infra fn 142.

[135] Cf *Familic* (1994) 75 A Crim R 229 and *Clarke* (1993) 71 A Crim R 58, similar fact evidence cases where the Victorian Court of Criminal Appeal has also misapplied the principles of *Shepherd (No 5)*. Cf *Pfennig* (1995) 69 ALJR 147.

[136] Where the falsity is relied upon by the prosecution to corroborate other evidence, the falsity must be proven other than by inconsistency with the evidence requiring corroboration. This is the fourth of four tests laid down by Lord Lane CJ in *R v Lucas (Ruth)* [1981] QB 720, 724, which were broadly approved by the High Court in *Edwards* (1993) 178 CLR 193.

[137] This is the first of the four tests laid down by Lord Lane CJ in *R v Lucas (Ruth)* [1981] QB 720, 724.

[138] The second and third tests laid down by Lord Lane CJ are that the alleged lie must relate to a material issue, and that other possible reasons for the lie, besides a consciousness of guilt, must be considered and rejected: *R v Lucas (Ruth)* [1981] QB 720, 724.

statement. This has led to concerns about bootstrapping and the appropriate directions to give to juries,[139] but space does not permit consideration of these issues here.

In the present context our interest in the inference arises because it has been the frequent subject of the *Chamberlain (No 2)* direction. In a series of decisions[140] it has been held that, to draw the inference of guilt from the evidence of the accused's lies, the jury should be satisfied beyond reasonable doubt, both that the accused lied and that the lie was caused by a consciousness of guilt.[141]

Following *Shepherd (No 5)* some courts appreciated that these directions were inappropriate where the consciousness of guilt argument was just one strand in a larger cable inference.[142] The Tasmanian case of *Jeffrey* v *The Queen*[143] provides an illustration. The accused was charged with murdering his father at his father's home. The accused admitted that he had visited his father twice on the day in question. His explanation for the second visit was that he had left his jacket behind on the first visit. However witnesses testified that the accused was wearing the jacket between the two visits. The prosecution argued that, in giving a false account of the innocent purpose of the second visit, the accused was displaying a consciousness of guilt. The court rejected the argument that the prosecution need establish the lie, and that it was motivated by a consciousness of guilt, beyond reasonable doubt. The court pointed out the argument was but one strand of a larger cable inference.[144] The prosecution also relied upon forensic evidence that the defendant's clothing had blood stains matching his father's rare blood type, and opportunity evidence.

But in other cases the structural principles of *Shepherd (No 5)* have not been applied correctly. Consider the Victorian case *R* v *Akbulat*.[145] The defendant, charged with murder, had, early in the investigation denied any involvement in the killing. Later, however, he admitted that he had killed the victim, but with provocation. The prosecution argued that his original denial revealed a consciousness of guilt. The defence submitted that, to accept this argument, the jury would need to be satisfied beyond reasonable doubt that the accused's

[139] 'By no torturing of the statement "I did not do the act" can you extract the evidence 'I did do the act".': *Edmunds v Edmunds and Asycough* [1935] VLR 177, 186; cf *R v Lucas (Ruth)* [1981] QB 720, 724; *Jeffrey* (1991) 60 A Crim R 384, 391 (per Cox J; Wright and Crawford JJ agreeing); *Gionfriddo and Fauvre* (1990) 50 A Crim R 327, 332–333(per Crockett and O'Bryan JJ); *Edwards* (1993) 178 CLR 193, at 199–200 (per Brennan J); 209–210 (per Deane, Dawson and Gaudron JJ)

[140] Eg, *Evans* (1985) 38 SASR 344, 347–8 ; *Gionfriddo and Fauvre* (1990) 50 A Crim R 327, 333 (per Crockett and O'Bryan JJ), 339 (per Gray J, diss); *Heyde* (1990) 20 NSWLR 233, 244.

[141] It is possible that such directions were considered necessary, in part, because of the concern about bootstrapping (cf infra fn 150).

[142] Eg, *Neilan* (1991) 52 A Crim R 303; *Jeffrey* (1991) 60 A Crim R 384; *Dellapatrona* (1993) 31 NSWLR 123; *Sandford* (1994) 72 A Crim R 160, 181; *Small* (1994) 72 A Crim R 462.

[143] (1991) 60 A Crim R 384.

[144] Id 396 (per Cox J; Wright and Crawford JJ agreeing).

[145] (1993) 69 A Crim R 75 (Coldrey J, Crockett and Hampel JJ agreeing); cf *Gibb* (31 August 1993, Victorian Court of Criminal Appeal, unreported)

lies were motivated by a consciousness of guilt. The court not only accepted this proposition[146] but considered that it was supported by *Shepherd (No 5)*.[147] The evidence, said the court, was not merely a strand of a cable, but was an implied admission of guilt. However the dichotomy established by the court was a false one. The consciousness of guilt inference was a strand in a larger cable structure. The prosecution also relied upon the testimony of an eyewitness that contradicted the accused's account of the circumstances of the killing.[148]

B. *Edwards*: 'the telling of a lie . . . an indispensable link . . .'?[149]

The High Court had occasion to consider the issue in *Edwards* v *The Queen*.[150] The majority gave a clear statement of how the principles of *Shepherd (No 5)* were to be applied.

> If the lie said to constitute the admission is the only evidence against the accused or is an indispensable link in a chain of evidence necessary to prove guilt, then the lie and its character as an admission against interest must be proved beyond reasonable doubt before the jury may conclude that the accused is guilty. But ordinarily a lie will form part of the body of evidence to be considered by the jury in reaching their conclusion according to the required standard of proof.[151]

In such a case the lie 'does not have to be proved to any particular standard of proof.'[152] There was, in *Edwards*, another strand of evidence, the testimony of the victim, and so the consciousness of guilt inference contained no indispensable links.

Unfortunately, the simplicity and clarity of this statement may have been obscured by Brennan J's dissenting judgment, in which he sought, unsuccessfully, to reconcile the hierarchical logic of the earlier cases with the structural principles of *Shepherd (No 5)*. His judgment is considered in detail below. In any event, subsequent decisions dealing with the consciousness have failed to follow the majority judgment in *Edwards*.

In the Victorian case of *The Queen* v *Heffernan*[153] the accused appealed against an arson conviction in relation to a fire at his business premises. The prosecution relied on three strands of evidence. First, there was motive evidence, consisting of his financial records showing the lack of success of the accused's business. Secondly, there was evidence that traces of mineral

[146] (1993) 69 A Crim R 75, 79, following *Heyde* (1990) 20 NSWLR 233 and *Gibb* (unreported, Victorian Court of Criminal Appeal, 31 August 1993).
[147] (1993) 69 A Crim R 75, 79.
[148] Id, 77.
[149] (1993)178 CLR 193, 210 (per Deane, Dawson and Gaudron JJ).
[150] (1993) 178 CLR 193. From the reported decision it is not clear whether the defence relied upon *Chamberlain (No 2)* to any degree in arguing for the application of the criminal standard to the inference. It appears that the argument may have been based upon a supposed circularity in the chain inference; cf fn 139.
[151] Id, 210 (per Deane, Dawson and Gaudron JJ).
[152] Ibid. Note that McHugh J would have left the question of the applicability of the criminal standard open: id 215, 216.
[153] Unreported, Victorian Court of Criminal Appeal, 20 May 1994.

turpentine had been found in the boot of his car; other evidence indicated mineral turpentine had been an accelerant for the fire. Thirdly, there was evidence that the accused had given an untruthful account of how these traces of mineral turpentine had appeared in his car.

Despite these three strands of evidence leading towards the accused's guilt, the trial judge had directed that the lies and the consciousness of guilt both needed to be established beyond reasonable doubt. Though the appeal court referred to the majority judgment in *Edwards*, it failed to criticise or correct the trial judge's approach.[154] On the contrary, the court relied more heavily on the judgment of Brennan J in *Edwards*, suggesting that it was irrelevant that it was a dissenting judgment.[155] The court said that the trial judge should not have talked so much in terms of 'the probabilities'[156] and should have given a general *Chamberlain (No 2)* direction. However it was considered 'unrealistic to suggest . . . that the jury . . . may not have understood that the [ultimate criminal] standard was to apply to the establishment of the very few disputed facts.'[157]

Unfortunately *Heffernan* is not an isolated example. The Victorian case of *The Queen* v *Angus*[158] concerned an appeal against an assault conviction. The prosecution relied primarily on the evidence of seven eyewitnesses, which, according to the appeal court, amounted to 'a seemingly overwhelming case'.[159] The trial judge suggested that the jury could also infer the accused's guilt from inconsistencies in the various accounts that he had given of the attack. The defence argued that the trial judge direction on this issue was insufficient, while the prosecution argued that there had been no significant departure from *Edwards*. Clearly, in view of the strength of the eyewitness accounts, the consciousness of guilt inference was not indispensable. However the appeal court criticised the trial judge for not giving the jury a *Chamberlain (No 2)* direction.

> It was clearly incumbent upon him carefully to direct the jury that before drawing the adverse inference the jury would need to be satisfied beyond reasonable doubt that the applicants explanation for the lie was false.[160]

C. Brennan J's attempted reconciliation of *Chamberlain (No 2)* and *Shepherd (No 5)*

Brennan J, in a dissenting judgment in *Edwards*, devoted more attention to the topic of cable and chain inferences than did the majority. He sought what

[154] Id 28.
[155] Id 18.
[156] Id 23.
[157] Id 21.
[158] Unreported, Victorian Court of Criminal Appeal, 2 December 1994; see also the cases *R v Diaper* (unreported, Victorian Court of Criminal Appeal, 20 May 1994) and *R v Tiplea* (unreported, Victorian Court of Criminal Appeal, 12 April 1995), in which the court found it unnecessary to determine the correctness of the *Chamberlain (No 2)* direction in relation to evidence of the accused's lies as showing a consciousness of guilt.
[159] Unreported, Victorian Court of Criminal Appeal, 2 December 1994, 13.
[160] Id 20.

this paper suggests is impossible — a reasoned reconciliation between the hierarchical fact-inference analysis and the structural cable-chain analysis.

Brennan J's strategy was to place the two inference structures at different points in a hierarchy.[161]

> The pieces of evidence are 'strands in a cable' tending to establish a material fact, but intermediate facts established by evidence are links in the chain of proof of the fact to be inferred.[162]

This enabled him to affirm structural principles resembling those in *Shepherd (No 5)*: 'the standard of proof applies to links; it says nothing about the strands.'[163] At the same time Brennan J was able to confirm the hierarchy established by *Van Beelen* and *Chamberlain (No 2)*:

> It is logically impossible to be satisfied beyond reasonable doubt that an inferred fact exists without being satisfied beyond reasonable doubt of the existence of those facts from which the inference is drawn. . . . There can be no superstructure erected without the foundations needed to support it.[164]

However the reconciliation is only superficial. The hierarchy of Brennan J is in fact stricter than that envisaged in *Chamberlain (No 2)*. In his scheme there is a place for cable inferences only at the base of the hierarchy, connecting evidence to facts: 'what is erroneous is a direction that the standard of proof governs the jury's evaluation of pieces of evidence.'[165] According to Brennan J, the facts are then connected to the ultimate verdict by chains.

> Standards of proof become relevant to the finding of the facts on which the verdict depends — the facts constituting the elements of an offence . . . and, where those facts are matters of inference, the facts from which the inference might be drawn.[166]

In *Chamberlain (No 2)*, however, Gibbs CJ and Mason J located cable inferences at the lowest level of the hierarchy, between evidence and facts, they also recognised their existence at the highest level, up to the ultimate inference of guilt.

> The jury should decide whether they accept the evidence of a particular fact, not by considering the evidence directly relating to that fact in isolation, but in light of the whole evidence, . . . they can draw an inference of guilt from a combination of facts, none of which viewed alone would support that inference.[167]

The scheme of Brennan J is still less consistent with the structural analysis of *Shepherd (No 5)*. Both chain inferences and cable inferences sit uneasily at

[161] The term 'intermediate facts' in *Shepherd (No 5)* could be taken to suggest some sort of hierarchy (supra fn 10) as could Dawson J's analysis of Brennan J's judgment in *Chamberlain (No 2)* (supra fn 69).
[162] (1993)178 CLR 193, 204 (per Brennan J).
[163] Id 204.
[164] Id 202.
[165] Id 205.
[166] Id 200–201.
[167] (1984) 153 CLR 521, 536.

their allocated levels in his hierarchy, as is clear from his analysis of a typical chain inference:

> If the emergence of an accused from the scene of the murder holding a smoking gun is the only evidence offered in proof of guilt, the case against him fails if there be a reasonable doubt about the identity of the person seen or the place where the fatal shot was fired or that the thing held by the accused was a smoking gun.[168]

The 'material facts'[169] to which the criminal standard is to be applied are readily identifiable here. But the cable substructure is missing. Where is the underlying evidence to which the standard has no application?

The opposite problem arises in Brennan J's analysis of a cable inference. The lower part of the hierarchy is there. It consists of 'a number of facts which are not themselves established beyond reasonable doubt'.[170] (For consistency, Brennan J should have referred to 'items of evidence' rather than 'facts'.) But where are the material facts satisfying the criminal standard, from which the inference of guilt may be drawn? To identify the 'foundation' of fact Brennan J is forced to play Roden J's semantic game.[171] 'The concatenation of probabilities is the fact on which the guilty inference must be based.'[172]

This reveals the artifice in the scheme of Brennan J. The chain inference in the first example and the cable inference in the second play equivalent roles in establishing guilt from evidence. They are better viewed as alternative structures with equal standing, rather than at different levels of a hierarchy.

In fact the hierarchy of Brennan J leads him to contradict the central principle of *Shepherd (No 5)*. As the majority in *Edwards* recognised,[173] the application of the criminal standard to the consciousness of guilt inference depends upon whether it involves indispensable links. For Brennan J, however, the question is determined solely by its location in the hierarchy. If the lies in *Edwards* were tendered merely to provide the necessary corroboration for the evidence of another witness, the inference would be located in the lower part of the hierarchy and 'no question of proof to a particular standard'[174] would arise. On the other hand:

> If the prosecution were to rely on the telling of the lie as an independent proof of guilt, the jury would have to be directed that, in order to convict on that basis, they must be satisfied beyond reasonable doubt that the true inference to be drawn from the accused's conduct is that he has confessed his guilt.[175]

In this case the proximity of the chain inference to the ultimate inference of guilt would put it in the higher portion of the hierarchy and attract the

[168] (1993)178 CLR 193, 202.
[169] Eg, id 202.
[170] Id 203.
[171] Supra fn 89.
[172] (1993)178 CLR 193, 203 (per Brennan J).
[173] Supra fn 151.
[174] (1993)178 CLR 193, 200.
[175] Id 201–2.

criminal standard, even though it contained no indispensable links in the chain of reasoning towards guilt.[176]

It is clear then that, despite the ingenuity of Brennan J's scheme, he did not achieve a reconciliation between *Chamberlain (No 2)* and *Shepherd (No 5)*. On the contrary, his hierarchy obfuscates the true structure of proof.

V CONCLUSION

Wells recently described the troubled history of the *Chamberlain (No 2)* direction in terms of a conflict between the jury's 'natural logic [which] remains intrinsically and categorically the same'[177] and the legal rules of evidence — the 'created or moulded product of legal policy'[178]. He considered the *Chamberlain (No 2)* direction, with no justification in policy, 'intolerably intrusive'.[179]

The analysis in this paper indicates, on the contrary, that humans have not been blessed with a uniform, incontrovertible system of natural logic. The confused lines of authority on criminal proof can be understood as the product of two prevalent but irreconcilable systems of logic. In this situation it is entirely appropriate for judges and legal commentators to intervene and offer the jury guidance on the structure of proof.

Shepherd (No 5) is based on the sound structural distinction between cable and chain inference structures. A conclusion of guilt that rests upon a number of strands of evidence may attain a standard of proof not reached by any single strand. The criminal standard should only be applied to the indispensible links of chain inferences — those underlying facts which are essential to the prosecution's case.

However the High Court in *Shepherd (No 5)* failed to overrule *Van Beelen* and *Chamberlain (No 2)*. These earlier cases are authority for a hierarchical rule which is in breach of the structural principles: whatever the larger proof structure, an inference cannot be drawn unless the factual basis of the inference is established beyond reasonable doubt.

The hierarchical rule of *Chamberlain (No 2)* is arbitrary in its application and inconsistent with the cumulative nature of circumstantial proof. Until the structural principles of *Shepherd (No 5)* are categorically endorsed, and *Chamberlain (No 2)* is clearly overruled, the courts will continue to use evidence inappropriately in drawing the ultimate inference of guilt.

[176] The distinction between the use of the evidence as confession or merely as corroboration may be an artefact of Brennan J's hierarchy. It did not occur to the majority who treated the two concepts interchangeably: id, 209 (per Deane, Dawson and Gaudron JJ).

[177] W A N Wells, *Natural Logic, Judicial Proof and Objective Facts* (1994), 2.

[178] Id 3; cf id 67–78.

[179] Id 78. Wells was clearly influenced by Roden J, who expressed similar views in *Shepherd (No 4)* (1988) 85 ALR 387, 389, 390, 395.