

A CRITIQUE OF THE THEORY OF COMPARATIVE PROPENSITY

PETER M ROBINSON*

The law of propensity evidence is in a state of flux in Australia as various State jurisdictions decide on their responses to recommendations of the Royal Commission into Institutional Responses to Child Sexual Abuse. Controversy persists about the probative value of such evidence, not limited to child sexual assault cases. An influential theory in this area is the theory of comparative propensity, advocated by Professor Hamer, and approved in a qualified way by the Royal Commission. The theory employs a mathematical model based on Bayes' equation to estimate the probative value of such evidence. This article critiques the theory and concludes that it does not reflect the real-world factors that impact the probative value of such evidence.

I INTRODUCTION

The law of propensity or tendency evidence in Australia is in a state of flux, thanks in part to the varied responses of the states to recommendations of the Royal Commission into Institutional Responses to Child Sexual Abuse. The Royal Commission concluded that evidence of prior conduct is generally undervalued.¹ While the Royal Commission formed this conclusion with respect to all types of case, its final recommendations were limited to its precise remit of child sexual assault cases. It recommended admission of tendency evidence in such cases if 'relevant to an important evidentiary issue' in the proceeding — which specifically included the 'propensity of the defendant to commit particular kinds of offences' — unless it is likely to result in unfairness to the defendant, which cannot be alleviated by an appropriate jury direction.²

Addressing the *Uniform Evidence Acts* ('UEA'),³ the Council of Attorneys-General declined to follow this recommendation, instead adopting a Model Bill, which takes the assessment of probative value largely out of the hands of judges. Under that Bill, it is presumed in trials of child sexual offences that tendency

* Lecturer, School of Business and Law, Central Queensland University: p.robinson1@cqu.edu.au. This research was conducted with the support of Macquarie University and Central Queensland University.

¹ *Royal Commission into Institutional Responses to Child Sexual Abuse* (Criminal Justice Report, Executive Summary and Parts II–III, August 2017) 70.

² *Ibid* 72, 128.

³ *Evidence Act 1995* (Cth); *Evidence Act 2011* (ACT); *Evidence Act 1995* (NSW); *Evidence (National Uniform Legislation) Act 2011* (NT); *Evidence Act 2001* (Tas); *Evidence Act 2008* (Vic).

evidence about a defendant's sexual interest in children has significant probative value under s 97 of the *UEA* even if the defendant has not acted on that interest, subject to the court's power to overturn this presumption on sufficient grounds.⁴ However, in considering whether sufficient grounds exist, the Bill precludes (except in exceptional circumstances) consideration of a number of factors that have historically been regarded as important in assessing the probative value of such evidence (eg, dissimilarity and distance in time between the past conduct and the alleged offences, and the generality of the supposed tendency). These changes only relate to child sexual offences, but the removal of the word 'substantially' from the requirement in s 101 that the probative value of evidence must substantially outweigh its prejudicial effect applies to all tendency and coincidence evidence.

At the time of writing, the model reforms have already been enacted in two jurisdictions (New South Wales and the Australian Capital Territory),⁵ but are expected to be implemented in other *UEA* States.⁶ However, the model legislation has already been heavily criticised by Hamer as poorly designed, 'paradoxical and ill-conceived', and propagating 'myths and misconceptions' about the probative value of propensity evidence.⁷

The non-*UEA* states (South Australia, Queensland and Western Australia) are yet to respond to the Royal Commission's recommendations, but it has been suggested that these states are unlikely to follow the *UEA* model, since they have previously been critical of the uniform legislation.⁸

In Queensland, the common law test of *Pfennig's* case — that there must be no 'rational view of the evidence that is consistent with the innocence of the accused'⁹ — still applies, subject to two qualifications contained in the *Evidence Act 1977* (Qld):

- *Hoch v The Queen*¹⁰ is overruled by s 132A such that the possibility of collusion or suggestion is not a ground for exclusion; and

⁴ Council of Attorneys-General, 'Communiqué', *Attorney-General's Department* (Statement, 29 November 2019) <<https://www.ag.gov.au/sites/default/files/2020-03/Council-of-Attorneys-General-communicue-November-2019.pdf>>; Parliamentary Counsel's Committee, *Uniform Evidence Law (Tendency and Coincidence) Model Provisions 2019*, Draft d15, s 97A <<https://pcc.gov.au/uniform/2019/29%20November%202019%20amendments.pdf>>.

⁵ *Evidence Act 1995* (NSW) s 97A; *Evidence Act 2011* (ACT) s 97A.

⁶ Andrew Hemming, 'Is There Any Prospect of a Model Provision for Similar Fact/Propensity Evidence or the Coincidence/Tendency Rules in Australia?' (2020) 44(4) *Criminal Law Journal* 207, 225; David Hamer, 'Myths, Misconceptions and Mixed Messages: An Early Look at the New Tendency and Coincidence Evidence Provisions' (2021) 45(4) *Criminal Law Journal* 232, 233 ('Myths, Misconceptions and Mixed Messages').

⁷ Hamer, 'Myths, Misconceptions and Mixed Messages' (n 6) 252.

⁸ Hemming (n 6) 231.

⁹ *Pfennig v The Queen* (1995) 182 CLR 461, 483 (Mason CJ, Deane and Dawson JJ) ('*Pfennig*').

¹⁰ (1988) 165 CLR 292.

- Section 132B permits the admission of evidence of prior domestic violence on certain charges involving homicide or serious assault.

The *Evidence Act 1929* (SA) echoes common law concepts of impermissible reasoning from general propensity and permissible reasoning about a ‘particular propensity’ directed to a fact in issue, but lowers the admissibility threshold for ‘discreditable conduct evidence’ to ‘strong probative value’, which ‘substantially outweighs any prejudicial effect’. This appears to be a higher threshold than that of ‘significant probative value’ under the *UEA*.¹¹

The Western Australian legislation provides that, to be admissible, propensity or relationship evidence must not only have significant probative value but the public interest in its admission must also, in the minds of ‘fair-minded people’, have priority over the risk of an unfair trial.¹² The issue of probative value is therefore common across jurisdictions, and conflicting views about the probative value of propensity evidence are evident in both common law cases¹³ and cases on identical legislation.¹⁴ Such conflicts no doubt impact policy and induce hesitancy in implementing reform.

The Royal Commission’s conclusion about the probative value of tendency evidence was strongly influenced by submissions from Hamer and earlier work by the late Mike Redmayne in the United Kingdom. Hamer and Redmayne both employed a form of mathematical probability theory, known as Bayesian analysis, to argue that evidence of prior conduct and convictions is substantially more probative than is traditionally appreciated.¹⁵ This approach was a fundamental plank in Hamer’s submissions to the Royal Commission and in the Commission’s conclusion that evidence of prior conduct was often undervalued.¹⁶ However, while the Royal Commission clearly embraced much of the work of Redmayne and

¹¹ *Evidence Act 1929* (SA) s 34P. Cf *UEA* s 97.

¹² *Evidence Act 1906* (WA) s 31A(2).

¹³ See, eg, *Phillips v The Queen* (2006) 225 CLR 303 (‘*Phillips*’); and the strident criticism of it by: David Hamer, ‘Similar Fact Reasoning in *Phillips*: Artificial, Disjointed and Pernicious’ (2007) 30(3) *University of New South Wales Law Journal* 609 (‘Similar Fact Reasoning in *Phillips*’); Jeremy Gans, ‘Similar Facts after *Phillips*’ (2006) 30(4) *Criminal Law Journal* 224; and Annie Cossins, ‘Similar Facts and Consent in Sexual Assault Cases: Filling in the Gap Left by the High Court in *Phillips*’ (2011) 37(2) *Monash University Law Review* 47 (‘Similar Facts and Consent’).

¹⁴ See the conflict between Victorian and New South Wales authorities in *Hughes v The Queen* (2017) 263 CLR 338; Peter M Robinson, ‘Reasoning About Tendency: What Does *Hughes v The Queen* Really Tell Us?’ (2019) 45(1) *Monash University Law Review* 98 (‘Reasoning About Tendency’).

¹⁵ David Hamer, ‘The Significant Probative Value of Tendency Evidence’ (2019) 42(2) *Melbourne University Law Review* 506, 508, 530, 548; Mike Redmayne, ‘The Relevance of Bad Character’ (2002) 61(3) *Cambridge Law Journal* 684; Mike Redmayne, *Character in the Criminal Trial* (Oxford University Press, 2015) 36.

¹⁶ See David Hamer, Submission to Royal Commission into Institutional Responses to Child Sexual Abuse (28 October 2016) 5–6; David Hamer, ‘Proof of Serial Child Sexual Abuse’, in Thomas Crofts and Arlie Loughnan (eds), *Criminalisation and Criminal Responsibility in Australia* (Oxford University Press, 2015), 242, 253–5, cited in *Royal Commission into Institutional Responses to Child Sexual Abuse* (Criminal Justice Report, Parts III–VI, August 2017) 604–7 (‘*Royal Commission*, Parts III–VI’); Redmayne, ‘The Relevance of Bad Character’ (n 15); Redmayne, *Character in the Criminal Trial* (n 15), cited in *Royal Commission*, Parts III–VI (n 16) 604–5.

Hamer in this area,¹⁷ it reserved its opinion on the mathematical calculations that flow from it:

We have not sought to test the validity of Redmayne's calculation of comparative propensity, so we do not place weight on a numerical proof of the relevance of prior convictions. However, it is interesting work and we note that it may warrant further consideration by law reform commissions if they consider these matters at some time in the future.¹⁸

Hamer has expounded his views on comparative propensity in a number of publications, most elaborately in 2019, when he asserted that the 'implications [of the analysis in his article] extend beyond child sex offence cases to criminal cases more broadly'.¹⁹ The theory therefore has the capacity to influence judicial thinking about the probative value of prior conduct evidence in all areas of law and in all forms of factual decision-making — not simply with respect to admissibility. The purpose of this article is to undertake the analysis suggested by the Royal Commission, and more generally, to critique the Bayesian approach to propensity evidence as implemented in the theory of comparative propensity.

This article has the following structure.

Part II introduces the mathematics on which the theory of comparative propensity is based. It first introduces some basic mathematics of probability with associated terminology and symbols, showing (among other things) the relationship between Bayes' equation and the legal concept of coincidence. It then proceeds to explain how Hamer's theory of comparative propensity fleshes out this equation with crime statistics to model the probative value of propensity evidence in terms of a 'likelihood ratio'. It then presents example calculations that lead to implausibly high estimations of the probative value of a record.

Part III considers the importance of the other evidence in the case apart from the propensity evidence. In the Bayesian model, prior odds of guilt are assessed without the propensity evidence, and I explain how these prior odds interact with the likelihood ratio. This creates problems for Hamer's model, because the model disregards the other evidence. I argue that Hamer's statistical approach of treating a hypothetical innocent defendant as if he or she were a random member of the general public fails to address the reality that the other evidence in the case is already likely to contain evidence adverse to the defendant's character (making a record more likely) and an innocent defendant may well have been wrongly charged for the very reason that he or she had a record. The problem is to find an appropriate reference class to statistically model the defendant, and I argue that the general population is not an appropriate class.

Part IV discusses how probative value might be affected by the size of the suspect pool. A limited suspect pool may point to the defendant as a potential

¹⁷ *Royal Commission*, Parts III–VI (n 16) 604–7.

¹⁸ *Ibid* 607.

¹⁹ Hamer, 'The Significant Probative Value of Tendency Evidence' (n 15) 508.

perpetrator without detracting from his or her character. The example is somewhat theoretical, but it is pertinent because Hamer rejects the relevance of the suspect pool on the basis that it involves a separate issue from propensity. This highlights his insistence that the probative value of propensity evidence must be considered without regard to the other evidence.

Part V introduces a related concept of redundancy. Since the other evidence is likely to lead to adverse inferences about the defendant's character and capacity to commit the crime, propensity evidence will only be corroborative of those inferences and therefore partially redundant. This raises questions about whether the Bayesian model can, from a mathematical point of view, properly represent the situation being modelled.

Part VI raises the theoretical question of whether a Bayesian model that disregards the other evidence can ever be appropriate, since Bayes' equation is designed to evaluate the coincidence of events and, if no other event is stipulated, there is no coincidence at all. I further argue that unlikely coincidence rather than predictive tendency is usually the basis upon which propensity evidence may have probative value.

For similar reasons, Part VII questions whether the Bayesian model can in any event be regarded as a model of propensity, rather than coincidence, reasoning. I consider the distinction between propensity and coincidence reasoning and provide examples to show how the probative value of motive evidence, which Hamer uses as an analogy to propensity, is more readily explained by coincidence.

Part VIII then proceeds to compare comparative propensity and coincidence reasoning, arguing that coincidence reasoning has the capacity to solve some difficult cases. However, this raises a further problem — that a similarity of uncommon offences can give rise to an unlikely coincidence with substantial probative value but nevertheless offend the law's prohibition on evidence of rank propensity. This is a policy matter that can only be addressed by the courts, but a clearer picture of the reasoning processes can illuminate the issues.

Part IX briefly considers how the theory of comparative propensity interacts with the presumption of innocence.

Part X concludes.

II THE THEORY OF COMPARATIVE PROPENSITY

The theory of comparative propensity posits that the probative value of propensity evidence can be determined by comparing the likelihood that a

hypothetical *guilty* defendant would have the record in question,²⁰ to the likelihood that a hypothetical *innocent* defendant would have such a record. It would be hard to quibble with this generalisation, but in itself it adds little to the rationale of proof. However, the theory goes further and asserts that this comparison can be mathematically calculated as a ‘likelihood ratio’ by means of a mathematical model in which an equation known as Bayes’ theorem is fleshed out with general population and crime statistics. Within this mathematical model, the likelihood ratio is a multiplier that multiplies the prior odds of guilt (absent the propensity evidence) to arrive at the posterior odds of guilt (after adding the propensity evidence). The size of this multiplier is therefore said to reflect the probative value of the propensity evidence, and evidence will have probative value favouring guilt if the likelihood ratio is greater than one.

A Bayesian Model of the Probative Value of Propensity Evidence

Bayes’ equation is based on a simple proposition, which is expressed in Equation 1.

Equation 1

$$\frac{P(G) \times P(E|G)}{P(I) \times P(E|I)} = \frac{P(E) \times P(G|E)}{P(E) \times P(I|E)}$$

$P(G)$ is in the standard notation for representing a probability, in this case the probability (‘P’) of guilt (‘G’). Where it appears on the left-hand side (‘LHS’) of the equation, it represents the probability of guilt based on the other evidence on the charge (‘the hard evidence’), without knowledge of the propensity evidence — in other words, where the defendant’s record is simply unknown. Similarly, $P(I)$ represents the probability of innocence (‘I’) based on the hard evidence alone. ‘G’ and ‘I’ are in fact complementary terms — ‘I’ could be expressed as ‘not G’, and in mathematical forms it often would be, but I have used the initials of the common terms for purposes of greater clarity.

‘E’ represents the propensity evidence. $P(E|G)$ is in the conventional form for representing a conditional probability, in this case the probability of the propensity evidence (‘E’) given that the defendant is guilty (‘G’). Conditional probabilities are important because they attempt to take into account the fact that many real-world events are interdependent — the occurrence of one makes the

²⁰ For ease of discussion, the term ‘record’ is used to refer to past criminal conduct, whether or not it has resulted in a conviction and a criminal record. In the theoretical discussions of both Redmayne and Hamer, they use conviction rates as indicators for rates of offending, so in most contexts the record referred to will be an actual criminal record, and that will be clear from the context. In practice, offending may often go unreported. In the Bayesian analysis, under-reporting of offences would cut both ways (though not necessarily equally). Both guilty and innocent defendants would be more likely to have past offences than statistics would suggest.

other more or less likely. This is fundamental to the theory of comparative propensity because the theory rests on the premise that a record is more likely if a defendant is guilty than if the defendant is innocent.

The multiplication operations are a variation of the product rule, which is used to calculate the overall probability of a conjunction, or coincidence, of events based on their individual probabilities. If events are wholly independent, then the probability of their coincidence is simply the product of their individual probabilities. For example, coin tosses are wholly independent events in that the outcome of one coin toss has no effect on the outcome of the next coin toss. The probability of two coin-tosses being heads is 0.5×0.5 , or 25%. On the other hand, if the outcome of one event has implications for the probability of the other event, then the events are said to be dependent, and conditional probabilities are required when using the product rule. $P(G) \times P(E|G)$, for example, represents the probability of the coincidence of guilt (calculated without knowledge of the propensity evidence) and the propensity evidence.²¹ Because the probability of the propensity evidence is affected by whether the defendant is guilty of a similar crime, 'G' and 'E' are dependent. The probability of the two events occurring together is the probability of one event occurring multiplied by the probability of the other event occurring, assuming that the first event has occurred. This can be expressed in reverse order: $P(G) \times P(E|G)$ expresses the same combination of events as $P(E) \times P(G|E)$. From this observation, one can see that the numerators and denominators on the LHS of the equation equate to the corresponding numerators and denominators on the right-hand side ('RHS'), and so the equation is obviously true. The key point of this analysis is that Bayes' equation is founded on the co-occurrence of events within a given scenario, which has ramifications for how it can be applied. It also provides a link between the probability concept of conjunctions and the legal concept of coincidences.

On the RHS of the equation, the terms $P(E)$ on the top and bottom cancel each other out. By then splitting the expressions on the LHS, one arrives at the odds version of the Bayesian equation that Redmayne and Hamer employ in their Bayesian model, as set out in Equation 2.²²

²¹ This variation of the product rule is known as the General Conjunction Rule.

²² Hamer, 'The Significant Probative Value of Tendency Evidence' (n 15) 530, prefers to break the parts of the equation into separate representations, but I have expressed them together to preserve the concept of the conjunctions underpinning the equation. The representation used here matches that of Redmayne, *Character in the Criminal Trial* (n 15) 35.

Equation 2

$$\frac{P(G)}{P(I)} \times \frac{P(E|G)}{P(E|I)} = \frac{P(G|E)}{P(I|E)}$$

Prior odds of guilt ('the prior') Likelihood ratio ('LR') Posterior odds of guilt ('posterior odds')

The probabilities are expressed in odds form so if, for example, the prior probability of guilt were 0.75 (or 75%), the probability of innocence would be 0.25 (or 25%), and the prior odds of guilt would be a ratio of three to one, or in gambling terms, 'three to one on'. Therefore, odds of three to one on equate to a probability of three out of four ($\frac{3}{4}$ or 75%).

The prior odds represent the odds of guilt based on the hard evidence only. The posterior odds represent the odds of guilt based on both the hard evidence and the record. The LR therefore is supposed to reflect the increase in the overall odds of guilt achieved by adding the propensity evidence to the hard evidence.²³ For that reason, it is seen as a measure of the probative value of that evidence. However, it is expressed as a multiplier rather than an actual increase in probability, so the actual increase in probability by adding the propensity evidence depends on what the LR is multiplying.

In both Redmayne's and Hamer's application of the model, the numerator of LR, $P(E|G)$, is based on recidivism statistics, which reflect the likelihood of a guilty person being a past offender. The denominator, $P(E|I)$, is based on the crime rate in the general population, which is said to reflect the likelihood of an innocent person having a record. They both regard the LR derived from this statistical methodology as a metric for determining the probative value of propensity evidence; Redmayne refers to it as a 'rough approximation'.²⁴ Hamer has recently reiterated his belief in this approach.²⁵ For any serious offence, since the likelihood of a person randomly selected from the general population having a record is extremely small, the LR calculated by this means will always be large (so large that Hamer concludes that propensity evidence 'will generally not struggle to achieve the ... threshold' of significant probative value under s 97 of the *Uniform Evidence Acts*).²⁶

By focusing on the LR only and deriving it by reference to general crime and population statistics, the effect is to derive a generic probative value of propensity

²³ It is pertinent here to clarify what the LR is in odds terms. It is not the odds of finding the evidence — that would be $P(E)/P(\text{not } E)$, which is the odds of finding the evidence relative to not finding it. The LR assumes the existence of the evidence and is the relative likelihood of finding the evidence if the defendant is guilty, compared to if he is innocent.

²⁴ Hamer, 'The Significant Probative Value of Tendency Evidence' (n 15) 544–5; Redmayne, 'The Relevance of Bad Character' (n 15) 693.

²⁵ Hamer, 'Myths, Misconceptions and Mixed Messages' (n 6) 238.

²⁶ Hamer, 'The Significant Probative Value of Tendency Evidence' (n 15) 549–50.

evidence without regard to the other evidence in the case or to the actual increment to the probability of guilt contributed by the propensity evidence. Hamer confirms his belief in this approach by rejecting a ‘strongly contextual model’ and asserting that, in assessing the probative value of propensity evidence, the other evidence should only be considered for identifying the issue to which the evidence is to be applied or to corroborate the tendency.²⁷

B Applying the Formula

Despite expressing reliance on the theories of Redmayne and Hamer, the Royal Commission declined to adopt the calculations derived from those theories. The calculations referred to by the Commission appear to be those set out by Hamer as follows:

Someone with a prior conviction is far more likely to offend than someone without a prior conviction. Drawing on conviction statistics for England and Wales, Redmayne suggests that ‘violent offenders in the 2009 cohort were 98 times more likely to commit an offence of violence than a member of the general population’. The comparative propensity figure for sexual offences is 2,353.²⁸

If one applied the figure for sexual offences to the Bayesian model, it would mean that a prosecution case with only a 2% probability of guilt based on the hard evidence would be catapulted to a near mathematical certainty by learning that the defendant had previously done something similar on some remote occasion:

Equation 3

$$\frac{0.02}{0.98} \times 2,353 = \frac{47.06}{0.98}$$

Odds of 47.06 to 0.98 approximate to 48 to one on, which represents a probability of 48 out of 49, or 98%. Although both Redmayne and Hamer promote such calculations as showing that evidence of prior conduct is undervalued,²⁹ the calculations should raise warning flags about the Bayesian model itself and the theories of comparative propensity flowing from it. Quite apart from the fact that the suggested effect of the record evidence is wildly implausible, the idea that any case with such doubtful hard evidence could achieve such a degree of certainty is completely far-fetched.

²⁷ Ibid 526–8.

²⁸ Hamer, ‘Proof of Serial Child Sexual Abuse’ (n 16) 253, cited in *Royal Commission*, Parts III–VI (n 16) 606. See Redmayne, *Character in the Criminal Trial* (n 15) 24, Table 2.7.

²⁹ Hamer, ‘Proof of Serial Child Sexual Abuse’ (n 16) 253–4; Hamer, ‘The Significant Probative Value of Tendency Evidence’ (n 15) 549; Redmayne, *Character in the Criminal Trial* (n 15) 23.

Lest it be thought that this result is an artefact of an unusually high LR, it should be noted that if the LR were halved, the posterior probability of guilt would be 24 to one or a probability of 24 out of 25 or 96%. Thus, halving the LR equates to only a 2% decrease in probability — a demonstration of why using the LR as a numerical gauge of probative value is, at best, deceptive. This arises from the fact that, not only is the LR a multiplier rather than a quantity, but what it is multiplying is an odds ratio, which is also not a quantity. Hence, the LR's relationship to any actual probability is obscure, especially if it is promoted as a free-standing measure of probative value to be used without regard to the prior, as Hamer suggests.

Redmayne himself gives an example of a burglary case with a prior probability of guilt of 50%, which was elevated (by an LR of 125) to a 99% certainty by the revelation of a similar record.³⁰ In the following discussion, I will explain why these calculations fail to reflect the real-life scenario being modelled.

III THE IMPORTANCE OF THE EVIDENCE ON WHICH THE PRIOR IS BASED

Hamer asserts that the LR is the Bayesian measure of the probative value of the propensity evidence,³¹ but that is somewhat misleading. The LR is a multiplier, not a probability. The actual probability it indirectly reflects depends on what it is multiplying, namely, the prior. The problem can be illustrated by elaborating Hamer's own model. One of the insights that Equation 1 gives us is that the Bayesian equation is actually evaluating a coincidence of events. It is only valid if the events are based on the same scenario, which means that the 'givens' of the model must be the same for both the prior and the LR (except that the LR adds the propensity evidence). In Hamer's model, there are no givens specified for the prior. In fact, the prior is simply ignored. However, the denominator of the LR is based on a randomly selected person from the general population. If the defendant really were a randomly selected individual from the general population, then the prior would also be based on the same assumption. It would be assessing the odds of a random individual being guilty of a specific alleged crime on a particular occasion in the absence of any evidence against him or her at all. It would calculate to an infinitesimally small number approaching zero. (Since legal verdicts must be based on evidence, in a trial context it would actually be zero). The LR would then tell us how much that infinitesimal number would increase if one learns that the random person has a record. Apart from the fact that an assumption of no hard evidence bears no resemblance to the real-life scenario being modelled, the exercise would in any case be futile because the

³⁰ Redmayne, 'The Relevance of Bad Character' (n 15) 695-6.

³¹ Hamer, 'The Significant Probative Value of Tendency Evidence' (n 15) 531.

increase in actual probability by adding the record would be tiny. Thus, to have any authenticity and utility, the Bayesian model needs some plausible hard evidence so that the LR has something meaningful to multiply.

Although Hamer does not explain why he treats the defendant as a random person, it would be wrong to suggest that, in applying random selection to the denominator of the LR, he implies that defendants are randomly selected. There is another way that random selection could be justified. It could be argued that, although the defendant is not randomly selected in terms of the likelihood of having a record, an innocent defendant is no different from a member of the general population, and therefore general population statistics can be used as a proxy to estimate the denominator. This is called using the general population as a reference class. In order to justify such an approach, it would be necessary to identify the characteristics of the defendant to justify applying such a reference class. Despite much literature emphasising the reference class problem in this context, Hamer makes no attempt to do this.

To correctly assign a reference class, it would be necessary to take into account the givens of the problem, which include the hard evidence. Alternatively, if one were attempting to create a generalised model, as Hamer appears to be doing, it would be necessary to construct a typical defendant, including any adverse conclusions drawn from the typical sort of evidence against him or her. When one considers the nature of the hard evidence, there is no basis for suggesting that a typical defendant who has been wrongly charged is similar to a person randomly selected from the general population. The hard evidence is likely to contain much evidence that is adverse to the defendant's character even if he or she is innocent. That may well be why he or she was wrongly charged. As Redmayne said in this context, the trial is likely to be 'awash with character inferences'.³² The facts of the *res gestae* might suggest adverse character inferences, and the defendant may appear to fall into a number of socioeconomic categories with a higher than normal incidence of crime. The general population is not an appropriate reference class for a typical defendant, as a number of commentators have noted.

A leading American author, Ronald Allen, when criticising the application of Bayes' theorem to legal cases in general, argued that the problem of identifying an appropriate reference class in real-life cases is inscrutable, because an event will fall into an infinite number of real-life classes, which would provide different reference rates:

Here is the critical point. The event under consideration ... is a member of an infinite number of reference classes, the boundary conditions of which can be gerrymandered in countless ways, some of which lead to the inference that the agent is reliable and some to the inference that he is unreliable, given that particular class. And — outside

³² Redmayne, *Character in the Criminal Trial* (n 15) 61.

of the reference class consisting only of the event itself — nothing in the natural world privileges or picks out one of the classes as the right one ...³³

Richard O Lempert, an American evidence scholar who is an advocate for the use of Bayesian analysis in law, rejects its use with respect to propensity evidence for the very reason that a typical defendant is not like a random person plucked from the general population. He points to a number of factors that would suggest that an innocent defendant, wrongly charged, is much more likely to have a prior record than a randomly selected person:

- the tendency of police to pursue ‘the usual suspects’ and to use mug shots of prior offenders in identifying a perpetrator;
- the tendency of prosecutors to try a weaker case if it is supported by a prior record, which is more likely if the defendant is innocent because innocent defendants tend to generate weaker cases; and
- the tendency of guilty defendants with prior records to accept a plea bargain, thus increasing the proportion of innocent defendants with records going to trial.³⁴

Lempert’s arguments are based on a systemic bias in the justice system towards selecting defendants with records, and he specifically rejects the approach adopted by Hamer for those reasons.³⁵ This issue has also been raised by Allen³⁶ and Mosteller.³⁷ While the first of Lempert’s criticisms would apply more strongly to identification cases, in which a crime is known to have been committed and the question is who did it, selectivity is also present in commission cases, where the defendant is identified but the question is whether he or she did it. Where there is doubt about the strength of the prosecution’s case — eg, where the complainant is a minor or testifying many years later, or where, on a sexual offence charge, it is one person’s word against another — a prosecution is more likely to be commenced and pursued against someone with a prior record than, for example, a Catholic priest.

Dahlman went further. He incorporated selectivity into a Bayesian model to argue that learning of a prior record actually increased the likelihood of innocence rather than guilt, a conclusion which has also been canvassed by Lempert.³⁸ He argued that the correct reference class is the population of defendants, not the

³³ Ronald J Allen and Michael S Pardo, ‘The Problematic Value of Mathematical Models of Evidence’ (2007) 36(1) *Journal of Legal Studies* 107, 112. See also Ronald J Allen et al, *An Analytical Approach to Evidence* (Wolters Kluwer, 6th ed, 2016) 181–4.

³⁴ Richard O Lempert et al, *A Modern Approach to Evidence: Text, Problems, Transcripts and Cases* (West Academic Publishing, 5th ed, 2014) 353–5.

³⁵ *Ibid* 353.

³⁶ Ronald J Allen et al, *Evidence: Text, Cases, and Problems* (Aspen Publishers, 2nd ed, 1997) 303.

³⁷ Robert P Mosteller, ‘Pernicious Inferences: Double Counting and Perception and Evaluation Biases in Criminal Cases’ (2015) 58(2) *Howard Law Journal* 365; Lempert (n 34) 354.

³⁸ Christian Dahlman, ‘The Felony Fallacy’ (2015) 14(3) *Law, Probability and Risk* 229; Lempert (n 34) 354.

general population, and that wrongly charged defendants were more likely to have a record than guilty ones. While Dahlman's precise modelling can be questioned,³⁹ it is undoubtedly the case that the class of defendants is quite different from the general population with respect to prior offending. Indeed, this was acknowledged by Park,⁴⁰ from whom Redmayne derived the theory of comparative propensity.⁴¹ In a footnote, Hamer acknowledges that the Bayesian odds should be conditioned on the other evidence, but failed to recognise the implications for his model: '[s]trictly speaking, all of the probabilities and odds in Bayes' theorem should be conditioned on background knowledge and other previously considered evidence. ... However, for brevity and simplicity, this condition has been omitted from the equations'.⁴²

IV THE RELEVANCE OF THE SUSPECT POOL

The problem of the prior has led some advocates of Bayesian methods to attempt a generic model by framing the prior in terms of the size of the potential suspect pool. Their reasoning only applies to identification cases. If there is no hard evidence at all, the prior likelihood of guilt approximates zero, but if one knows that a crime has definitely been committed, the probability of any particular person committing the offence rises to one divided by the size of the entire population. If hard evidence can be added that restricts the suspect pool to something much less than the general population, the prior might attain a level whereby the value of the LR becomes significant.

This scenario has been the subject of a classic debate around a hypothetical scenario known as the 'island problem', in which it is postulated that a crime is known to have been committed on an island with a limited population.⁴³ Walsh, Buckleton and Triggs proposed a solution to the problem by weighting a suspect's probability of committing the crime by reference to statistics on the geographical location of known offenders relative to the location of the crimes.⁴⁴ In effect, suspects who resided in the locality of the crime received a greater weighting than more distant potential suspects. Fenton et al more recently developed the concept

³⁹ Peter M Robinson, 'Incorporating Implicit Knowledge into the Bayesian Model of Prior Conviction Evidence: Some Reality Checks for the Theory of Comparative Propensity' (2020) 19(2) *Law, Probability and Risk* 119.

⁴⁰ Roger C Park, 'Character at the Crossroads' (1998) 49(3) *Hastings Law Journal* 717, 742.

⁴¹ Redmayne, 'The Relevance of Bad Character' (n 15) 684, 693.

⁴² Hamer, 'The Significant Probative Value of Tendency Evidence' (n 15) 530 n 133.

⁴³ For early examples, see David J Balding and Peter Donnelly, 'Inference in Forensic Identification' (1995) 158(1) *Journal of the Royal Statistical Society. Series A, (Statistics in Society)* 21; AP Dawid and J Mortera, 'Coherent Analysis of Forensic Identification Evidence' (1996) 58(2) *Journal of the Royal Statistical Society. Series B, (Methodological)* 425.

⁴⁴ KAJ Walsh, JS Buckleton and CM Triggs, 'Assessing Prior Probabilities Considering Geography' (1994) 34(1) *Journal of the Forensic Science Society* 47.

of the ‘opportunity prior’ to address this issue.⁴⁵ Their idea is that in identification cases the suspect pool can be restricted to the estimated number of people who had the opportunity to do it, taking into account their proximity to the crime scene within the relevant time span of the crime.

Hamer is dismissive of the concept of a restricted suspect pool, asserting that the size of the suspect pool ‘has nothing to do with the probative value of tendency evidence’.⁴⁶ This is symptomatic of his approach that the probative value of the record should be assessed without regard to the other evidence.

In practice, it is difficult to see how the restriction of the suspect pool would assist in a decision at trial unless the pool were limited to a rather small number of suspects whose backgrounds and other involvement could be investigated and laid before the court. The tendency of a criminal trial to focus on the defendant alone makes this kind of case rare, although there are some examples in the case law.⁴⁷

V DEPENDENCIES BETWEEN THE PRIOR AND THE LIKELIHOOD RATIO

Analysis to this point has centred on the failure of the Bayesian model to address the evidence on which the prior is based in estimating the LR. That failure in itself undermines the utility of the model because correcting it places the calculation outside the realms of statistical analysis. However, the model has a further related problem, which undermines its mathematical soundness. One of the preconditions for application of the product rule, on which Bayes’ equation is founded, is that the probabilities of the coinciding events must be wholly independent. The test for this is to ask, ‘does the probability of one event imply anything about the probability of the other event?’ If the answer is ‘yes’, then the assumption of independence fails.

The conditional probabilities of the LR, $P(E|G)$, and $P(E|I)$, are designed to take into account the dependency between guilt/innocence and the existence of a record. The trouble is that guilt/innocence is not an indivisible variable representing only a single finding of fact. It is a composite consisting of multiple factual findings on a range of variables that could influence the likelihood of a record, eg, all the intermediate facts which may affect the conclusion on the defendant’s general character or disposition. It also includes findings on intermediate facts that may affect the significance of general character to the overall finding of guilt, such as findings on the defendant’s immediate mental state at the time of the alleged crime. The conditionals represented in the LR are inadequate to account for such multiple dependencies. Although the fact finders

⁴⁵ Norman Fenton et al, ‘The Opportunity Prior: A Proof-Based Prior for Criminal Cases’ (2019) 18(4) *Law, Probability and Risk* 237.

⁴⁶ Hamer, ‘The Significant Probative Value of Tendency Evidence’ (n 15) 527.

⁴⁷ *Lowery v The Queen* [1974] AC 85; *R v Randall* [2003] UKHL 69.

may have doubts about guilt, they may have no doubt at all that the defendant had the character necessary to commit such an offence. Propensity evidence will only ever be corroborative of, and incremental to, the evaluation of character already incorporated into the prior. In certain circumstances, the adverse conclusions about character made in the prior may be sufficient to completely exhaust the utility of the propensity evidence.⁴⁸

The adverse effect of redundancy on Bayesian models of evidence has been previously noted by Lempert.⁴⁹ Redmayne also recognised that the propensity evidence must be contributing something completely new,⁵⁰ but failed to recognise that, since the intermediate fact sought to be proven by the propensity evidence is not the record itself but the character or disposition attributed to it, the contribution of the propensity evidence will never be entirely new, at least not in the sense required to justify the mathematics of his model. In particular, the multiplication operation of the product rule, which is what supposedly demonstrates the power of propensity evidence,⁵¹ is unwarranted.

Hamer acknowledges (in a cursory manner) the possibility of redundancy, but fails to recognise its impact on the validity of his mathematical model,⁵² instead providing an analysis that is misleading. He says that as corroborative evidence is added, its probative value decreases until, at a certain point, it 'falls off a cliff'.⁵³ Here, Hamer is assuming that there is a single stream of proof towards guilt, so that probative value only expires when the case as a whole is proven beyond reasonable doubt.⁵⁴ However, because propensity evidence is only probative of the intermediate fact of character (or perhaps, disposition), the ceiling of probative value may fall well short of the point where overall guilt is established.

Within Hamer's model, the point at which the LR falls off a cliff is as soon as the first piece of evidence adverse to character is presented as part of the hard evidence (assuming it is given some credit). This first piece of credible evidence transforms that defendant from a random citizen to a person with a probability of having the character necessary to commit that type of crime. This step will invariably be achieved by the hard evidence (unless the hard evidence is totally disbelieved).

⁴⁸ As Lempert has observed, if the evidence were merely corroborating the ultimate fact of guilt, no harm would be done by admitting redundant evidence because it would only become wholly redundant when guilt is established. However, because the evidence is used to prove an intermediate or constituent fact (namely character), the redundant evidence can prejudice the assessment of the ultimate fact of guilt, which may still be in substantial doubt despite the adverse character inferences: Richard O Lempert, 'Modeling Relevance' (1977) 75(5-6) *Michigan Law Review* 1021, 1048 n 63.

⁴⁹ *Ibid* 1041-2, 1051-2.

⁵⁰ Redmayne, *Character in the Criminal Trial* (n 15) 37.

⁵¹ *Ibid* 15.

⁵² This criticism could also be made of Redmayne, *Character in the Criminal Trial* (n 15).

⁵³ Hamer, 'The Significant Probative Value of Tendency Evidence' (n 15) 527.

⁵⁴ *Ibid* 521-2.

VI THEORETICAL PROBLEMS WITH THE BAYESIAN APPROACH

At a more theoretical level, there are fundamental problems with a Bayesian model in this context. It is a model specifically based on the coincidence of evidence, but Hamer specifies no such coincidence. In his model, the evidence on which the prior is based is unspecified, yet the model assumes that the defendant has been individually selected at random. This random selection of a specific individual is what generates high figures for the LR. In practice, it is the hard evidence that selects a particular defendant for prosecution and creates the coincidence to which Hamer would apply the Bayesian equation.

The significance of this can be seen by assuming that an offence is known to have been committed, but that there is no evidence pointing to a particular person — a so-called ‘identification case’. This is actually more hard evidence than Hamer specifies in his model. If a crime is known to have been committed, but there is no evidence pointing to a particular defendant, how likely is it that the offender had a prior record? The following table provides some insight. It sets out historical data from New South Wales on the prevalence of prior records among persons convicted of the more common types of offences against other persons.⁵⁵

Table 1: Prior Offences of Persons Convicted of Offences Against Other Persons

Jurisdiction/Type of offence	No prior record (%) (of any type)	Prior record (%) Same offence type	Prior record (%) Different offence type ⁶
Homicide-related (murder, attempted murder and manslaughter)	45.3	1.2	53.4
Assault (non-sexual)	45.7	27.2	27.1
Sexual assault (not involving a child)	54.2	6.3	39.5
Child sexual assault	63.8	8.3	27.9

Note: prior record is based on record for previous five years.⁵⁷

⁵⁵ These are sources which Hamer has cited, along with figures from other jurisdictions: Hamer, ‘The Significant Probative Value of Tendency Evidence’ (n 15) 545.

⁵⁶ In the sources, similarity of offence type was based on the offence type described. For example, a prior offence of assault would be regarded as different from a homicide-related offence.

⁵⁷ These figures are derived from a series of crime and justice statistics published by the NSW Bureau of Crime Statistics and Research: Isabel Taussig, ‘Sentencing Snapshot: Homicide and Related Offences’ (Issue Paper No 76, February 2012); Isabel Taussig, ‘Sentencing Snapshot for Assault’ (Issue Paper No 66, February 2011); Clare Ringland, ‘Sentencing Snapshot: Sexual Assault, 2009–2010’ (Issue Paper No 72, January 2012); Jessie Holmes, ‘Sentencing Snapshot: Child Sexual Assault, 2009–2010’ (Issue Paper No 68, May 2013). For homicide-related offences, I have not included the reported figures for driving offences causing death, and for assault, I did not include the reported figures for stalking.

Table 1 shows that for sexual offences, the party convicted was more likely to have no record of any kind than to be a previous offender, whereas for non-sexual offences, they were only slightly more likely to have a record. If the offender did have a record, dissimilar records were much more prevalent than similar ones, except with respect to non-sexual assault, where similar and dissimilar records were virtually tied. Despite the known biases in the system towards pursuing suspects with prior records, such figures suggest that, if a serious crime is known to have been committed, it is no better than a toss-up whether the offender had a record or not.

So, how can these figures be reconciled with the idea that previous criminal records are probative of guilt? The answer lies in base rates. A specific individual with a prior record may be more likely to commit a crime than a randomly selected individual without a record, but the number of people without a record is much higher than the number of previous offenders; so, overall, the commission of serious crimes tends to be balanced fairly evenly between prior offenders and first offenders (at least based on conviction records). Accordingly, if one were to attempt a generic model of the probative value of record evidence on the assumption that there is no hard evidence pointing to a particular defendant, or as Hamer has done, on the basis that such evidence should be disregarded, then the record would have no probative value at all. However, if there is hard evidence pointing to a particular defendant, or at least to a limited suspect pool, then the coincidence of the hard evidence and the record may provide a valid basis for reasoning about this unlikely coincidence.

The analysis above assumes that a crime has been committed. If no such assumption is made, one is left with a bare allegation of a crime accompanied by no evidence apart from the record. This is the scenario that would flow from Hamer's rejection of a 'strongly contextual model'.⁵⁸ If no other evidence is taken into account, a coincidence model is inappropriate, and the problem becomes one of simple prediction. How well does a record predict a specific crime on a particular occasion in the absence of any other evidence? Hamer recognises that 'past offending ... provides a poor basis for predicting future offending'.⁵⁹ It is even worse for predicting re-offending on a specific occasion. The reason again relates to base rates. Recidivism statistics only predict repetition of behaviours over large time spans — one year, three years, perhaps longer. If an offender has multiple opportunities to re-offend over a lengthy period of time, the fact that they have some probability of re-offending on a couple of those occasions tells you only that they are *unlikely* to re-offend on most occasions when given the opportunity. The sort of behavioural evidence that is soundly based on prediction

⁵⁸ Hamer, 'The Significant Probative Value of Tendency Evidence' (n 15) 526–7.

⁵⁹ Hamer, 'Myths, Misconceptions and Mixed Messages' (n 6) 238; See also David Hamer, 'Before the High Court: Tendency Evidence in *Hughes v The Queen*: Similarity, Probative Value and Admissibility' (2016) 38(4) *Sydney Law Review* 491, 495–6 ('Tendency Evidence in *Hughes v The Queen*'); Hamer, 'Proof of Serial Child Sexual Abuse' (n 16) 252–3.

is evidence of habit or regular business practice, where behaviours are repeated like clockwork in particular circumstances.

Park, the originator of the idea of comparative propensity, recognised that the value of record evidence depends on its co-occurrence with evidence specifically pointing to the defendant.⁶⁰ Similarly, Redmayne seems to recognise that the predictive power of propensity evidence is weak and that its value depends on other evidence pointing to the guilt of the specific defendant.⁶¹ Hamer has acknowledged on many occasions that coincidence plays a role in propensity reasoning but nevertheless argues that his approach to comparative propensity provides a distinct form of propensity reasoning, even though it disregards the very evidence that creates the coincidence.⁶²

VII THE DISTINCTION BETWEEN TENDENCY EVIDENCE AND COINCIDENCE EVIDENCE

Much of the confusion in this area, I would argue, derives from the entrenchment in Australia of two notionally distinct forms of reasoning: (1) tendency or propensity reasoning, and (2) coincidence or probability/improbability reasoning, with tendency or propensity reasoning apparently holding sway when the coincidence involves similarities in conduct or disposition. Hamer distinguishes the holistic nature of coincidence reasoning, based on the unlikely coincidence of events, from the sequential nature of propensity reasoning, which proceeds as follows:⁶³

- 1 The defendant committed other similar misconduct.
- 2 This demonstrates that the defendant has a propensity to commit this kind of misconduct.
- 3 This increases the probability that the defendant committed the charged offence.

This structure, by its very nature, focuses on the predictive effect of a known propensity, and it is the form of reasoning that is generally impermissible under the common law.

Hamer maintains that the probative value of tendency evidence is not based on its predictive power but rather on comparative propensity.⁶⁴ However, as will

⁶⁰ Park (n 40) 723–4.

⁶¹ Redmayne, 'The Relevance of Bad Character' (n 15) 692.

⁶² Hamer, 'Tendency Evidence in *Hughes v The Queen*' (n 59) 496, 499; Hamer, 'The Significant Probative Value of Tendency Evidence' (n 15) 526–8.

⁶³ David Hamer, 'The Legal Structure of Propensity Evidence' (2016) 20(2) *International Journal of Evidence & Proof* 136, 145.

⁶⁴ Hamer, 'The Significant Probative Value of Tendency Evidence' (n 15) 544; Hamer, 'Tendency Evidence in *Hughes v The Queen*' (n 59) 499.

appear, reasoning based on his Bayesian model bears no real resemblance to the sequential structure set out above. Furthermore, Bayes' equation does not provide any distinct form of propensity reasoning. It is simply a mathematical means of calculating a coincidence of events. To the extent that propensity is relevant, it is not due to reasoning from propensity; it is due to the fact that a surmised propensity may be a more likely explanation of an unlikely coincidence of offences (or allegations of offences) than the alternative that an innocent person, wrongly charged, happened to have such a record. If any proper distinction were to be drawn between the two forms of reasoning, it would be that when coincidence reasoning is applied to tendency evidence, there will inevitably be substantial dependencies between the hard evidence and the tendency evidence rendering the mathematics of Bayes' equation unsound.

Recognising that tendency evidence is generally not predictive, and that its probative value depends on coincidence reasoning, has the potential to simplify reasoning in this area. Coincidence reasoning is in fact at the heart of all reasoning about evidence. Typically, when we are considering the value of evidence, we weigh up competing theories of the case. These theories represent alternative narratives connecting events suggested by the evidence. In other words, they are elaborate conjunctions of events, and we weigh them up by assessing which conjunction is more plausible. Additional evidence will only contribute probative value if it distinguishes between one theory of the case (the theory leading to guilt) and another theory (the theory suggesting innocence).⁶⁵

This idea can be seen in an analogy previously adopted by Hamer — motive evidence.⁶⁶ He asserts the power of motive evidence as follows:

Motive evidence ... is valued very highly. This does not reflect a view that a person with a motive to murder is highly likely to murder, or that motive evidence by itself would constitute proof. The point is that someone with a motive is far more likely to commit murder than someone without a motive.⁶⁷

Using this as an analogy with propensity evidence involves a subtle misconception. Within the Bayesian model, the comparison is not between a defendant with a record and a defendant without a record. It is a comparison between a defendant with a record and a defendant whose record is unknown. Juries draw adverse conclusions about the defendant based on the hard evidence alone, including the possibility that he has done the same sort of thing before. Propensity evidence is merely supplementary to, and corroborative of, those conclusions.

Turning to the motive analogy, let us take an example of a husband accused of the murder of his wife. We learn that he is the beneficiary of a life policy taken

⁶⁵ Of course, this is simplified. There can be multiple theories of the case leading to guilt or innocence.

⁶⁶ Hamer, 'Proof of Serial Child Sexual Abuse' (n 16) 253; Hamer, 'Tendency Evidence in *Hughes v The Queen*' (n 59) 496.

⁶⁷ Hamer, 'Proof of Serial Child Sexual Abuse' (n 16) 253.

out on his former wife's life. One should note at the outset that, unlike propensity evidence where a single act may be assumed to show an enduring character disposition, motive evidence does not of itself bespeak any mental state. No one would suggest that the presence of a life insurance policy on a spouse's life evidences, even weakly, any actual mental disposition in a person to murder their spouse. In assessing the probative value of the evidence of the life policy, these are the two alternatives posed for comparison:

1. How likely is it that a husband guilty of murdering his wife would be the beneficiary of such a policy?
2. How likely is it that an innocent husband wrongly charged with murdering his wife would be the beneficiary of such a policy?

This comparison bears little resemblance to Hamer's argument that someone with a motive is far more likely to commit a murder than someone without a motive. It is a good example of why the value of new evidence should be assessed in the context of the hard evidence already accounted for. Once one knows from the hard evidence that the defendant is the husband of the deceased, it becomes clear that an innocent defendant is very likely to be the beneficiary of such a policy, because spouses often make financial provisions of this type. Such motive evidence has little or no probative value because there is no unlikely coincidence in a husband potentially gaining financially from the death of their spouse, whether he is guilty or not.

Additional hard evidence may cast a different light on the evidence of a life policy — say, for example, the husband had actively negotiated the policy shortly before the wife's death, when there was no such insurance before. That would be a considerable coincidence. The two questions would then become:

1. How likely is it that a guilty husband would negotiate life insurance on his wife for the first time shortly before he killed her?
2. How likely is it that an innocent husband would happen to negotiate new life insurance on his wife's life shortly before she was killed by someone else?

The value of the evidence in this revised example is not based on any change in the probative value of the policy itself. It is based on the unlikely coincidence of the contemporaneous negotiation of the policy and the killing. One further caution could be added. If there is already evidence of a mental state sufficient to commit the murder, then evidence of motive may be substantially redundant (though unlike tendency evidence, its admission may not be regarded as prejudicial).

In earlier work, Hamer used the case of *R v Baden-Clay*⁶⁸ to demonstrate the power of motive evidence. In that case, a husband, under pressure to end his marriage due to an extramarital relationship and unable to afford a divorce, was convicted of murdering her on circumstantial evidence. Unlike the example of the life policy, there was clear evidence of motivational state because the defendant had promised his lover that he would end the marriage, but had not fulfilled that promise.⁶⁹ Hamer supported his argument by a quote from the High Court that reflected pure coincidence reasoning:

‘it tested credulity too far to suggest that his evident desire to be rid of his wife was fortuitously fulfilled by her unintended death’.⁷⁰ In the following discussion, I will compare how Hamer’s approach to the probative value of propensity evidence differs from the coincidence approach.

VIII ASSESSING THE PROBATIVE VALUE OF THE RECORD

A Hamer’s Approach

Hamer’s application of the LR involves comparing the consistency of the record with guilt (the numerator $P(E|G)$) and the consistency of the record with innocence (the denominator $P(E|I)$).⁷¹ I will refer to these as the guilt hypothesis and the innocence hypothesis, respectively.

1 The Guilt Hypothesis

Hamer asserts that, for determining consistency with guilt, ‘the predictive characterisation broadly captures the strength of the consistency element’, and he describes it in terms of ‘the predictability of an offender reoffending’.⁷² This appears to underpin his assertion that his approach represents a distinct form of reasoning based on propensity. Elsewhere, he recognises that this approach involves a reversal of the proper logic. The correct approach to the numerator of the LR is not to reason whether the record predicts guilt, but rather whether guilt predicts the record, but Hamer does not regard the reversal of the ‘prediction’ as problematic.⁷³ However, the two forms are conceptually very different and would attract different statistics. If one wishes to assess statistically whether generic

⁶⁸ (2016) 258 CLR 308.

⁶⁹ Ibid [22]. Despite the fact that financial gain was also canvassed as a motive, the prosecution specifically declined to argue that a life policy on the wife’s life was evidence of a motive: at [29].

⁷⁰ Ibid [69]; Hamer, ‘Tendency Evidence in *Hughes v The Queen*’ (n 59) 496. Hamer followed this case with a reference to *Pfennig* (n 9), which was also clearly based on coincidence reasoning: Hamer, ‘Tendency Evidence in *Hughes v The Queen*’ (n 59) 497.

⁷¹ In fact, Hamer more often refers to inconsistency with innocence, which is the complement of the denominator.

⁷² Hamer, ‘The Significant Probative Value of Tendency Evidence’ (n 15) 535.

⁷³ Ibid 535 n 150.

guilt predicts a record, Table 1 provides an appropriate reference class, namely a sample of guilty defendants. The figures reflect the likelihood of a guilty defendant having a record for particular types of offence, and they do not support the view advocated by Hamer that similarity of features, or even similarity of offences, increases consistency with guilt, since guilty defendants were more likely to have dissimilar past offences than similar ones.⁷⁴ Again, one sees the importance of base rates. Serious offences are much less common than lesser offences, so guilt may be more likely to coincide with lesser, dissimilar offences than similar ones. To the extent that distinctive or unusual similarity is relevant, it is not because it strengthens any supposed propensity, but because it makes an innocent coincidence more unlikely.

2 The Innocence Hypothesis

With respect to the innocence hypothesis, Hamer does not suggest the same reversal of logic, which is just as well because a criminal record predicts far more innocent behaviour than guilty behaviour. The LR would be less than one, and the record would predict innocence rather than guilt. Instead, he asks whether generic innocence (as represented by a randomly selected citizen) predicts the record.⁷⁵ I have already addressed two factors which undermine this approach:

- a) Unlike randomly selected individuals, people who are wrongly charged with serious offences may well have a record; and
- b) if the hard evidence is at all credible, it will already incorporate adverse conclusions about the defendant's character, rendering the record to a significant extent redundant.

However, there is another factor that must be taken into account. While coincidences of unusual events may be rare as isolated combinations, when they have manifold opportunities to occur, they may be quite common. This is what Murphy J was driving at in the following oft-quoted passage:

Common assumptions about improbability of sequences are often wrong. A suggested sequence, series or pattern of events is often incorrectly regarded as so extremely improbable as to be incredible. However highly improbable, as well as merely improbable, sequences and combinations are constantly occurring. In random tossing the occurrence of a run of ten consecutive heads or tails is generally regarded as highly improbable. But this will occur on the average once in every 512 tosses, and the lesser sequences more frequently (2 runs of 9; 4 runs of 8; 8 runs of 7). If one randomly tosses a coin 257 times, more likely than not there will be a sequence of ten heads or tails.

⁷⁴ Ibid 532.

⁷⁵ See Hamer, 'The Significant Probative Value of Tendency Evidence' (n 15) 533, 544–5.

Although it is extremely improbable that any particular ticket will win a large lottery, it is certain that one will.⁷⁶

Murphy J was highlighting another issue of base rates. The conventional, simplistic concept of probability is an average probability based on a single random selection, but the true likelihood of something occurring in the real world depends not only on that average but also on the number of opportunities the event (or combination of events) has to occur. This principle is particularly important when the system is not based on random selection but on selecting particular types of unusual case, eg, serious crimes. Given sufficient opportunities, rare combinations do occur on a regular basis. The case with which Murphy J was dealing involved a female defendant who had a surprising number of male relatives who died from arsenic poisoning. Apart from that coincidence, there was no hard evidence inculcating the defendant in their demise.⁷⁷ The coincidence of those deaths in a randomly selected family would be highly unlikely, but when one considers the infinite variety of life (and death) and the infinite opportunities for such a coincidence to occur, the fact that one such case arose and arrived at the courts is perhaps not as surprising as it may seem.

3 Issues

The divergence between the Bayesian model and coincidence reasoning comes into focus when Hamer considers how probative value varies depending on the issue in the case. In two places, he considers the example of a sexual assault case in which the complainant describes an unusual or peculiar predilection of the perpetrator and there is evidence that the defendant has displayed that predilection on previous occasions.⁷⁸ He expresses the argument in terms of coincidence. If identity is the issue, 'it would be quite a coincidence for the complainant to report on the defendant's predilection if it were someone other than the defendant who committed the assault'.⁷⁹ However, if the defendant admitted the sexual acts and simply put consent in issue, then evidence of the predilection would no longer be 'an incriminating coincidence'.⁸⁰

This is very different from the Bayesian model in which the LR is conditioned on a generic proposition of guilt or innocence. Hamer provides no explanation of how the Bayesian model is to be adjusted to account for these varying probative values. In fact, the propositions of generic guilt or innocence in the model are effectively replaced by much more specific propositions derived from the hard evidence. For the guilt hypothesis, Hamer says that '[t]he consistency element

⁷⁶ *Perry v The Queen* (1982) 150 CLR 580, 594 [11] ('Perry').

⁷⁷ *Ibid* 591 [1] (Murphy J).

⁷⁸ Hamer, 'The Significant Probative Value of Tendency Evidence' (n 15) 523, 533–4.

⁷⁹ *Ibid* 524.

⁸⁰ *Ibid*.

would be at a moderate level in line with recidivism data',⁸¹ but what needs to be assessed is the likelihood of a person, guilty of the acts with the predilection described in the hard evidence, having a history displaying that same predilection. That assessment has nothing to do with recidivism statistics. Assuming that Hamer adopted a consistent approach for the innocence hypothesis, its likelihood would be informed by crime statistics for the general population, but the likelihood of an innocent person having a matching predilection also has nothing to do with such statistics.

By defining the case as one in which identification is the only issue, the coincidence of the particular alleged conduct and the past predilection is conflated with the general concept of guilt versus innocence. The logic of coincidence is directed to specific factual propositions involving a coincidence between the hard evidence and the propensity evidence, not to generic guilt or innocence. The fact that the answer to those propositions may lead to an inference of guilt does not alter the more specific nature of the logic. In cases where multiple issues are outstanding, a proper focus on the precise coincidence would be essential.

In the identification example, the hard evidence is important not only to define the precise act itself but also to point to the defendant as a potential perpetrator. Without this selection of the defendant, the innocence hypothesis poses the following question: how likely is it that *somebody* from the general population other than the guilty party had a history of the same predilection? Unless the predilection were extraordinarily unique, the probability of the innocence hypothesis would likely be high, possibly higher than the guilt hypothesis. Thus, the detail of the hard evidence, rather than the issue itself, frames the LR.

The need to assess the propensity evidence by reference to its precise coincidence with the hard evidence is not something that arises from a refinement of the issue. The same analysis would apply if all the issues were at large. In the variation where the defendant admits the act but argues consent, Hamer discounts the predilection evidence as it does not give rise to an incriminating coincidence. That is true, but the same logical inference is available as in the identification example — it is just that the admission of the act makes evidence of the specific predilection redundant. However, even in a consent case, the defendant's more general character or disposition to commit non-consensual sex is also in issue, and propensity evidence may be relevant to that issue. I consider the relevance of a more general propensity to the question of consent when I discuss the case of *Phillips v The Queen* below.

To adjust the Bayesian model to accommodate the particular issue, it would be necessary to abandon the generic model and reframe the prior as a more specific factual proposition defined by the hard evidence, rather than simple guilt

⁸¹ Ibid 533.

or innocence. In effect, this is what Hamer partially did when he expressed his conclusions in terms of a more specific coincidence of predilections. In the identification case, the prior would represent the odds of the defendant being the person who committed the act absent the predilection evidence. If the hard evidence did not in some way identify the defendant as a potential perpetrator, at least as a member of a limited suspect pool, then the prior would be close to zero and the LR would be largely irrelevant. If the hard evidence did provide some identification evidence pointing to the defendant, then there would be a potential coincidence to assess. The LR would compare the likelihood of the defendant having displayed the same predilection if they had or had not committed the act. The result would be an assessment of the odds of the defendant having committed the act, given the predilection evidence. If identification were the only issue, commission of the act would conflate with guilt, but otherwise it would simply be a factor to be assessed along with the evidence on the other issues in determining overall guilt.

This approach would more closely align with that of forensic scientists who advocate for a Bayesian evaluation of forensic evidence. In their approach, the forensic scientist is presented with specific alternative propositions for the prosecution and the defence and also the background of the other evidence ('a framework of circumstances') to enable an assessment.⁸² For example, evidence may show a coincidental match between a partial DNA trace at the crime scene and the defendant's DNA. The prosecution proposition is that it was the defendant's DNA at the crime scene. The defence proposition is that it was somebody else's. Even with these very narrowly defined propositions, one must be mindful of the other hard evidence. For example, if the innocent defendant were modelled as a random person, the probability of the defence hypothesis may be, say, one in one million. But if the hard evidence shows that a relative of the defendant (with a similar DNA profile) was also a suspect, the defence hypothesis might jump to a 50/50 proposition. The result of the assessment would be evaluated by the fact finder along with the other evidence to determine guilt. For example, there may be possible scenarios in which an innocent defendant's DNA could have found its way to the crime scene (or to the crime lab by contamination) that the fact finder would have to assess.

If that approach were applied to propensity evidence, in addition to the usual problems of dependencies and redundancy (which are *less* if the factual propositions are expressed more narrowly), there would be the added problem of incorporating a Bayesian LR on one factual issue into non-Bayesian findings on other issues. Hamer's assertion that probative value depends on the issue impliedly concedes that propensity evidence is only probative of certain issues. This means that one cannot automatically convert a Bayesian finding on a specific

⁸² Charles EH Berger et al, 'Evidence Evaluation: A Response to the Court of Appeal Judgment in *R v T*' (2011) 51(2) *Science & Justice* 43, 44.

fact into a corroborative finding on overall guilt. The specific fact may not advance the proof of other issues. Given that (unlike DNA evidence) the results of such an analysis do not seem to be calculable in any event, it is difficult to see what the Bayesian paradigm contributes over and above ordinary coincidence reasoning.

In the following explanation of coincidence reasoning, the question of the issue is dealt with by framing the factual propositions specifically and by taking account of redundancy. If there is no issue on which the propensity evidence would be relevant, then it is redundant.

B Assessing Probative Value Based on Coincidence

I would argue that, to reason soundly about propensity in this way, the numerator and denominator of the likelihood ratio should be replaced by a comparison of the following probabilities representing coincidences of the hard evidence with the record:

The guilt hypothesis

How likely is it that a person who is guilty in the circumstances defined by the hard evidence would have the record alleged?

The innocence hypothesis

How likely is it that an innocent person, wrongly charged in the circumstances described by the hard evidence, would have the record alleged?

In weighing up these alternative hypotheses, one must also take account of the following factors:

- any adverse assessments of character drawn from the hard evidence that would make a record more likely, and possibly wholly or partially redundant;
- any conclusions about the immediate *mens rea* that would make general character wholly or partially redundant;
- the fact that innocent defendants with records tend to be much more common than random individuals with records; and
- the fact that unusual combinations of circumstances have multiple opportunities to occur in real life, rendering them more common than is often thought.

Selection bias should be treated with caution. Selection of a case for prosecution does not in itself alter the probative value of the evidence, but it can have practical implications for assessing that value. Selection bias increases the probability of a record for both guilty and innocent defendants. Figures like those in Table 1,

which incorporate systemic selection bias, therefore overstate the association of guilt with a record, while subjective intuitions that an innocent coincidence would be remarkable may be exaggerated because selection bias systematically picks out unusual cases from the multiplicity of human activity. If the evaluation of the guilt hypothesis uses biased statistics like those in Table 1, then the evaluation of the innocence hypothesis must also account for the same bias.

Although coincidence reasoning is well known in similar fact cases, it is useful to consider how it applies in practice to propensity. The two hypotheses simply represent competing explanations for the coincidence of the record and hard evidence tending to inculpate the defendant (if only by placing him or her in a limited suspect pool in an identification case). They are not considered separately from the hard evidence and the ratio of their probabilities is not treated as a multiplier to be applied to a discrete assessment of the hard evidence on its own. They are simply weighed against each other, taking into account the factors outlined above, to form part of the overall assessment of the case. The initial premise is that the coincidence of the record and the hard evidence would be unlikely if the defendant were innocent. If the hard evidence does not tend to inculpate the defendant (similar to the situation in Equation 3), then there is no incriminating coincidence to consider, and it would look very much as if the defendant had been charged simply because of his or her record. However, contrary to cases that distinguish coincidence reasoning from propensity reasoning on the basis of the need for similarities,⁸³ similarity of the events is not required to give rise to an unlikely coincidence. In the motive example, the unlikely coincidence arose from the negotiation of a life insurance policy shortly before the wife was killed. The two events bear no resemblance to each other, apart from proximity in time.

The basis for asserting probative value toward guilt is that a criminal propensity is a more plausible explanation of the coincidence than an innocent interpretation. Unlike reasoning that focuses on the strength of the propensity, often by reference to similarities that may or may not affect its predictive power, the focus is on explaining the coincidence, and the propensity may be wholly inferred as the most likely explanation. This avoids arguments about circularity of reasoning that arise in the serial form of propensity reasoning, where the propensity has to be established by contested evidence before it can be used to predict the offence under charge.⁸⁴

The hard evidence is important in generating hypothetical scenarios supporting guilt or innocence. For example, the plausibility of the propensity explanation may be undermined by evidence showing that the charge relates to a different victim or class of victims to the record or that the charged events (if they

⁸³ *R v PWD* (2010) 205 A Crim R 75, 91 [79], approved in *Saoud v The Queen* (2014) 87 NSWLR 481, 491 [46].

⁸⁴ See, eg, *Sutton v The Queen* (1984) 152 CLR 528, [5] (Gibbs CJ), [20] (Brennan J); *Thompson v The Queen* (1989) 169 CLR 1, 5; Annie Cossins, 'Similar Facts and Consent' (n 13) 60.

occurred) were the result of some more immediate mental state unrelated to general character. The record itself may also weaken the inference of propensity if it does not show repetitive behaviour over a length of time — for example, if it is only one instance or if it relates to the distant past. The tendency towards recidivism is known to subside as criminals age, and a lengthy period of time without re-offending is somewhat inconsistent with a persistent propensity. (As mentioned earlier, under the recent reforms of the *UEA*, some of these arguments are severely curtailed with respect to child sexual assault cases).

In considering the innocence hypothesis, as Hamer has pointed out,⁸⁵ the rarity of the offence is important. *Prima facie*, it would be a considerable coincidence if hard evidence wrongly inculpated an innocent defendant in a serious crime and he or she happened to have a known record for similar offences (unless the witnesses' evidence were tainted by collusion or knowledge of the record). Unusual similarities between the record offences and the charged offence may heighten the coincidence, but features that are commonplace for innocent behaviours, such as an institutional setting or geographical location close to home, do not.

The unlikelihood of innocent coincidence may be diluted in a number of ways. If the hard evidence detracts from character — for example, if the circumstances relate to people who frequent criminal classes, or if they are members of socio-economic groups with a higher than usual crime rates — the record might not be so unexpected compared to an average person selected from the general population. Innocent people who have been wrongly charged are much more likely than ordinary citizens to have a record.

One must also consider the fact that the record will only ever be corroborative of other findings on character or disposition. This does not render record evidence irrelevant, but it means that exaggerated calculations of unlikely coincidence generated mathematically by the product rule or by an intuitive assumption that the record evidence provides something entirely new are unwarranted. When considering the admission of record evidence, it should be remembered that, if the record is not admitted, the jury will be presented with a defendant whose record is simply unknown. They will not automatically assume that he or she has no record. This is particularly so if there is hard evidence detracting from character, but even without that, juries approaching their task with an open mind will be alive to the possibility that the defendant may have a chequered past. As such, they will already be accounting for the possibility that the defendant has the character or disposition capable of committing the crime. The record may be relevant to confirm their suspicions, but when the effect of the evidence is merely confirmatory rather than wholly new, it will to some extent be redundant. It will be less redundant if the defendant appears to fall into a class among whom criminal records would be highly unlikely, such as priests.

⁸⁵ Hamer, 'The Significant Probative Value of Tendency Evidence' (n 15) 532.

It is in this context that one can also consider the actual issues in the case. As seen in Hamer's predilection example, admissions by the defendant may also render the propensity evidence either wholly or partially redundant.

Finally, any assessment of the plausibility of the innocence explanation must take into account the fact that combinations of events that may be unlikely on a randomly selected occasion are regularly occurring in everyday life. In this respect, the degree of unusualness will again be a practical consideration. For example, if the history involves multiple discrete coincidences and/or the similarities between the charged events and the criminal history are particularly unusual and specific, chance coincidence may still seem implausible. On the other hand, if the similarities between the past and charged events are only generic, the possibility that the coincidence occurred by chance is much more real when one considers the multiplicity of human affairs.

Case law has long acknowledged an overlap between propensity and coincidence reasoning.⁸⁶ Recognition that the probative value of propensity evidence lies in coincidence reasoning would provide a more transparent evaluative process than the current preoccupation with the distinctiveness of similarities. While Hamer's mathematical methodology exaggerates the probative value of propensity evidence, it is nevertheless true that in many case contexts, a serious criminal record would be substantially more likely in a guilty defendant than an innocent one. However, in order to adopt this approach, the courts would have to accept that a generalised character tendency, or 'rank propensity', may have substantial probative value.

The preoccupation with a distinct form of propensity reasoning has led (I would argue) to an exaggerated focus on the modus operandi of criminals, in particular sexual offenders. Hamer presages a detailed analysis of this issue with the following comment:

[T]he higher admissibility threshold appears to reflect an assumption that child sexual offenders are relatively unlikely to reoffend, but if they do, the offences will all share distinctive similarities. As legal commentators have recognised, this assumption can be tested against empirical data.⁸⁷

That may be a logical surmise from a preoccupation with the distinctiveness of similarities, but one would be hard-pressed to find in the historical reasoning of judges much evidence that they are making either of these assumptions. In focusing on distinctiveness, judges seem to be attempting to analyse, however misguidedly, the capacity of past behaviour to predict the alleged offence, by finding a precisely matching propensity. The theory, as I perceive it, is that in order to predict specific behaviour on a particular occasion, it is necessary to find a propensity sufficiently specific to match it. A rank propensity is insufficient.

⁸⁶ See, eg, *KJR v The Queen* (2007) 173 A Crim R 226, 236 [43], [46]; *Saoud v The Queen* (n 83) 491 [43]; *Hoyle v The Queen* (2018) 339 FLR 11, 39 [165]–[169]; *R v WBN* (2020) 5 QR 566, 604 [112].

⁸⁷ Hamer, 'Proof of Serial Child Sexual Abuse' (n 16) 251.

If one recognised that a rank propensity can have substantial probative value by reason of coincidence reasoning, it would become clear that similar features would only add to probative value if they heightened the coincidence by rendering it even more unlikely in an innocent defendant than the coincidence of the bare offences (or allegations of offences) themselves. This would provide some support for judicial arguments that one must examine whether the behaviour is unusual *for the type of offence*, but only if one simultaneously conceded that the rank similarity of offences supplies most of the unlikelihood.

I would argue that the problem with adopting the coincidence approach is not logical or mathematical, but philosophical. The controversial common law case of *Phillips v The Queen* ('*Phillips*') provides some insight.⁸⁸ Phillips was charged with a series of seven offences involving rape and indecent assault of five complainants between August 2000 and November 2001, when he was 16–17 years of age. While he was on bail for those charges, he was accused of another assault with intent to rape in May 2003, which became an eighth count in the indictment. The trial judge declined several applications to sever the trials and all counts were heard together on the basis that the evidence of all complainants was cross-admissible on the other complaints, leading to convictions on six of the eight counts, with lesser verdicts of unlawful carnal knowledge on two of those six. These convictions were upheld by the Queensland Court of Appeal,⁸⁹ but overturned by the High Court on the basis that the evidence was cross-admitted on the limited issue of consent, but evidence of one complainant's failure to consent could not be relevant to the issue of whether another complainant consented.⁹⁰ In other words, although there was a remarkable coincidence of allegations, since the stated issue was the complainants' consent rather than the defendant's behaviour, it was difficult to frame a propensity to explain it. The fruitless search for an applicable propensity, rather than an unlikely coincidence, was therefore the stumbling block to cross-admissibility.

The trial judge's decision was clearly founded on coincidence reasoning:

So you ask yourselves this, what are the probabilities that all six girls have lied when they say they did not consent ... If you think it could possibly be just an unlucky coincidence then you consider each incident and the evidence of each girl's completely separately and you reach your verdicts in light of your view of the evidence relating to each incident completely separately. But if you are satisfied that the only reasonable conclusion to be drawn is that they are all telling the truth when they say they did not consent ... then you may use that conclusion in your thinking along the path to deciding whether [the defendant] is guilty or not guilty of each of the offences.⁹¹

The trial judge was not reasoning *from* propensity. His Honour was not inferring the girls' consent from the defendant's inferred propensity. If anything, his

⁸⁸ *Phillips* (n 13).

⁸⁹ *R v PS* [2004] QCA 347.

⁹⁰ *Phillips* (n 13).

⁹¹ *Ibid* [67].

Honour was doing the reverse. The multiplicity of similar allegations made the innocence hypothesis (that they were all lying) unlikely. This led to consideration of the alternative guilt hypothesis, that the girls had not consented and that the defendant had acted on an illicit propensity. This balancing of alternative hypotheses relates to a specific coincidence, not to overall guilt. The plausibility of the hypotheses would have to be weighed up alongside assessments of the other evidence in each particular case, which may support or detract from the inferences of guilt or innocence derived from the multiplicity of similar allegations.

The High Court's reasoning indicates why courts that are minded to admit tendency evidence often engage in a gymnastic search for similarities:⁹²

Criminal trials in this country are ordinarily focused with high particularity upon specified offences. They are not, as such, a trial of the accused's character or propensity towards criminal conduct. That is why, in order to permit the admission of evidence relevant to different offences, the common law requires a high threshold to be passed. The evidence must possess particular probative qualities; a really material bearing on the issues to be decided. That threshold was not met in this case.⁹³

Elsewhere, the Court emphasised that this threshold could only be achieved by evidence that had a 'sufficient nexus' or 'specific connexion' with the issues in the subject case.⁹⁴

The first two sentences of the extract from *Phillips* seem to suggest a philosophical argument rather than a logical or mathematical one, namely, that a defendant should be tried on evidence specific to the offence, not on evidence of his or her character or past. This argument is closely aligned with the presumption of innocence. However, the subsequent sentences imply that the problem with such evidence is that it lacks probative value, which is true if one reasons from propensity, but not if one argues from coincidence. The coincidence of rapes or rape allegations may be unlikely even if the modus of the rapes differs.

The decision in *Phillips* was not well-received in the academic world. Hamer described it as 'artificial, disjointed and pernicious'.⁹⁵ Gans said the Court's reasoning was 'at odds with reality'⁹⁶ and 'a poor, and possibly counterproductive response to the significant risks of miscarriage of justice arising from joint trials'.⁹⁷ Perhaps the unkindest challenge to the Court's objectivity came only slightly more subtly from Cossins, who presaged one criticism with the words: '[t]he High Court, comprised of five male judges', etc.⁹⁸

⁹² See, eg, the review of authorities in Robinson, 'Reasoning About Tendency' (n 14).

⁹³ *Phillips* (n 13) 327--8 [79].

⁹⁴ *Ibid* 321 [55], citing *Pfennig* (n 9) 483 (Mason CJ, Deane and Dawson JJ).

⁹⁵ Hamer, 'Similar Fact Reasoning in *Phillips*' (n 13).

⁹⁶ Gans, 'Similar Facts after *Phillips*' (n 13), 230.

⁹⁷ *Ibid* 233.

⁹⁸ Cossins, 'Similar Facts and Consent' (n 13) 72.

It is open to the law to privilege past conduct and convictions from admission if they merely represent a generalised, ‘rank’ tendency, but the only basis for doing so is that they are either too prejudicial or that their admission would offend a sacrosanct right to the presumption of innocence. Suggestions that a rank tendency cannot have significant probative value are, I submit, unsupportable.

IX COMPARATIVE PROPENSITY AND THE PRESUMPTION OF INNOCENCE

Coincidence reasoning about rank propensity is at least based on the conduct of the defendant himself or herself, so one could argue that adverse conclusions about the defendant drawn from such conduct do not offend the presumption of innocence. On the other hand, the theory of comparative propensity presents a greater challenge to that presumption. Within that theory, propensity evidence is evaluated without reference to the hard evidence, by assuming that an innocent defendant must be a randomly selected individual from ‘the general law-abiding population’.⁹⁹ This means that the defendant is fixed with a predetermined, generic LR — a multiplier — which attaches to him or her statistically before he or she enters the courtroom. Whatever hard evidence is led of the crime, its probative value will automatically be escalated by that multiplier, and as we have seen, the multiplier calculated in this manner will always be very substantial — so substantial that the flimsiest case can be promoted to a near certainty by admitting the defendant’s record.

The approach in which the probative value of the record is predetermined by some generic calculation could hardly be more at odds with the presumption of innocence. It goes further than simply raising a ‘highly suspicious, prejudicial atmosphere’ as feared by Murphy J,¹⁰⁰ but rather puts the defendant with a record in a position where he or she would be lost at the outset. The only way that such an outcome could be averted is by rejecting not only the calculations derived from the theory of comparative propensity, but also any ‘rough’ approximations derived from it.¹⁰¹

X CONCLUSIONS

This article analysed the Bayesian model, which is the foundation for the theory of comparative propensity. That theory has held some sway in Australian jurisprudence in recent times. The analysis demonstrates that both the model’s statistical assumptions and mathematical foundations fail to reflect the real-world scenario it purports to depict. In addition, the Bayesian model does not

⁹⁹ Hamer, ‘The Significant Probative Value of Tendency Evidence’ (n 15) 513, 528, 545, 547, 549.

¹⁰⁰ Perry (n 76) 594 [11] (Murphy J).

¹⁰¹ See, eg, Redmayne, ‘The Relevance of Bad Character’ (n 15) 693.

represent a distinct form of propensity reasoning but rather a mathematical representation of coincidence reasoning. The analysis provides some insight into the issues that arise when applying such coincidence reasoning to real world cases, and indicates that the logic of reasoning about propensity would be better served by recognising that the probative value of propensity evidence derives from the coincidence between the hard evidence inculcating the particular defendant and his or her record, than by reasoning *from* propensity. In doing so, however, one would also have to acknowledge that the 'rank' coincidence of uncommon offences (or alleged offences) is inherently unlikely in an innocent defendant even if those offences have no distinctive features. If that were acknowledged, courts could address the real question of whether rank propensity should continue to be privileged from admission despite its coincidental probative value.

I am conscious that the criticisms made in this article might seem strident, and that those criticisms are primarily directed at the theories of one of Australia's leading evidence scholars — I would say, *the* leading evidence scholar. These criticisms are not intended to diminish Professor Hamer's contributions in this area. Nevertheless, I believe that it would be dangerous to allow exaggerated 'approximations' of probative value, generated by the Bayesian model, to pervade legal thinking about prior conduct evidence. In many ways, the problems of dependencies and composite variables described in this article are symptomatic of a wider problem with applying mathematical formulae, which tend to be based on independent, one-dimensional variables or simple dependencies, to the complex combinations of interacting factors involved in human behaviour. There seems to be a trend in modern society for numbers generated by such mathematical approaches to be given much greater credence than their rudimentary nature warrants.

