

CHRISTIE, SECTION 137 AND FORENSIC SCIENCE EVIDENCE (AFTER *DUPAS v THE QUEEN* AND *R v XY*)

GARY EDMOND,* DAVID HAMER,** ANDREW LIGERTWOOD*** AND MEHERA SAN ROQUE****

I INTRODUCTION

This article addresses the implications of *Dupas v The Queen*¹ and *R v XY*² for the handling of expert evidence under s 137 of the Uniform Evidence Law (UEL).³ Focused on forensic science and medicine evidence, and in anticipation of High Court consideration, we offer an alternative lens through which to approach the current dialogue between the New South Wales and Victorian appellate courts.⁴ Whatever their differences, both *Dupas* and *XY* are concerned with the obligation of the court to limit the admissibility of evidence that has the potential to mislead or be misused by the trier of fact and thereby undermine rational decision making. This article directs attention to an enhanced role for s 137 in the regulation and, where appropriate, exclusion of incriminating ‘expert’ evidence.⁵

In this article we offer an account — influenced by emerging evidence on the weaknesses of the contemporary forensic sciences — of how expert evidence led by prosecutors has the potential to mislead and be misused by the trier of fact (in particular the jury) in ways that cannot be easily or adequately addressed by conventional trial safeguards. These risks are accentuated where opinions are derived from forensic science techniques that have not been formally evaluated,

* Professor, Australian Research Council Future Fellow, and Director, Expertise, Evidence & Law Program, School of Law, The University of New South Wales, Sydney 2052, Australia. Email: g.edmond@unsw.edu.au.

** Associate Professor, School of Law, University of Sydney.

*** Emeritus Fellow, School of Law, University of Adelaide.

**** Senior Lecturer, School of Law, The University of New South Wales.

1 (2012) 218 A Crim R 507 (*‘Dupas’*).

2 (2013) 84 NSWLR 363 (*‘XY’*). For a more detailed description, see Michael W R Adams and Christopher K Wareham, ‘Is Judicial Consideration of Credibility and Reliability under Section 137 of the Uniform Evidence Law a Guarantee of Fairness or “Moral Treason”?’ (2014) 40 *Monash University Law Review* 243.

3 The uniform evidence legislation (UEL) now covers all major jurisdictions with the exception of Queensland, South Australia and Western Australia. See *Evidence Act 1995* (Cth); *Evidence Act 2011* (ACT); *Evidence Act 1995* (NSW); *Evidence Act 2004* (Norfolk Island); *Evidence (National Uniform Legislation) Act 2011* (NT); *Evidence Act 2001* (Tas); *Evidence Act 2008* (Vic).

4 It is not our intention to resolve the tensions in these cases, but rather to discuss the decisions in the context of an attempt to reiterate the purpose of s 137 and its common law progenitor *R v Christie* [1914] AC 545 (*‘Christie’*), and the particular risks introduced by incriminating expert evidence.

5 ‘Expert’ is placed in quotation marks because in a surprising number of cases we do not know whether the individuals allowed to proffer their incriminating opinions actually possess any relevant expertise.

and where admissibility depends primarily on legal recognition of the witness's experience or some putatively relevant formal qualifications.⁶

While historically courts have expressed concerns about the potential for expert opinion to sway or mislead juries, in general they have not translated these concerns into an effective regulatory jurisprudence.⁷ In relation to ss 79, 135 and 137 of the UEL, Australian courts have tended to rely on outmoded or inadequate heuristics to regulate admission and inform the evaluation of incriminating expert evidence.⁸ Our approach to expert evidence requires lawyers and judges to attend carefully under s 137 to the probative value (and therefore reliability) of forensic and medical evidence when considering whether there is a risk that the jury will attribute to that evidence more probative weight than it warrants. Simultaneously, it supplements s 79(1) — the primary regulator of expert opinion evidence. Requiring expert opinions to be based on 'specialised knowledge' demands evidence of 'knowledge' that exists beyond the individual 'expert'. If expert opinion is to assist the jury in reaching a more accurate decision, logic would appear to demand that the underlying 'knowledge' be reliable (and identifiable). If, as is currently the case, 'specialised knowledge' is not interpreted in a way that demands reliability, the trier of fact must be placed in a position where it is able to rationally assess the probative value of the 'expert' evidence. If not, the risk of the evidence being overvalued (or otherwise misused) arises under s 137.⁹

At present, s 137 is usually read in a way that does not attend to the actual probative value of 'opinions based on specialised knowledge'. In most cases there

- 6 For a comprehensive critique of the contemporary forensic sciences, see National Research Council, *Strengthening the Forensic Sciences in the United States: A Path Forward* (National Academies Press, 2009) ('NAS Report'); Expert Working Group on Human Factors in Latent Print Analysis, *Latent Print Examination and Human Factors: Improving the Practice through a Systems Approach* (National Institute of Standards and Technology and National Institute of Justice, 2012) ('NIST Report'); Sir Anthony Campbell, *The Fingerprint Inquiry Report* (APS Group Scotland, 2011); Stephen T Goudge, *Inquiry into Pediatric Forensic Pathology in Ontario* (Queen's Printer for Ontario, 2008). For a detailed review of these reports, see Gary Edmond, 'What Lawyers Should Know about the Forensic "Sciences"' (2015) 36 *Adelaide Law Review* (forthcoming).
- 7 Gary Edmond, 'Specialised Knowledge, the Exclusionary Discretions and Reliability: Reassessing Incriminating Expert Opinion Evidence' (2008) 31 *University of New South Wales Law Journal* 1. Section 79 of the UEL has not been read or applied in a way that requires lawyers and judges to attend to the existence of 'specialised knowledge' or, in relation to forensic science and medicine, to demand evidence that the opinions are demonstrably reliable. Section 79 has not been read to require those with qualifications or ostensible experience to provide evidence that their techniques work (and how well) and, if so, that they are proficient in using them. In the absence of attention to evidence of reliability under s 79, there is a persistent risk that opinions — expressed by witnesses called by the prosecution and recognised by courts as experts — will be overvalued, and that any limitations, even those identified or explained by the defence, will not be understood or appropriately valued by the trier of fact.
- 8 Gary Edmond, 'The Admissibility of Forensic Science and Medicine Evidence under the Uniform Evidence Law' (2014) 38 *Criminal Law Journal* 136. The recent decision of the High Court in *Honeysett v The Queen* (2014) 88 ALJR 786 ('*Honeysett*'), held that the 'expert' comparison evidence had been wrongly admitted, but declined to address the larger question of whether 'specialised knowledge' required indicia of reliability such as independent validation of techniques or proficiency: at 794 [42].
- 9 In addition, there may be occasions where even demonstrably reliable techniques might nonetheless produce opinions that have the potential to mislead or create specific difficulties for the trier of fact. An example might be the failure to attend to cognitive bias, contextual factors or cross-contamination: see Part III C of this article.

is little informed scrutiny of incriminating expert opinion evidence under either s 79 or s 137 (or s 135). In many cases the need to persuade the trier of fact that the opinions are weak, speculative or unreliable, and that the state's routine practices are neither rigorous nor methodologically sound, falls to an impecunious defendant at considerable personal risk. We argue that s 137, explicitly oriented to evidence adduced by the prosecutor, has a substantial (and perhaps expansive) role to play in circumstances where s 79 has been construed, and applied, in a manner that disregards 'specialised knowledge' and reliability. In this regard, ss 79(1) and 137 are intimately connected and partially overlap. Where s 79(1) is rigorously applied, s 137 will generally have less of a role to play in regulating the admissibility of expert evidence.

Drawing upon *Dupas* and *XY*, particularly the explicit recognition in *Dupas* that judges can, and should, consider the actual probative value of evidence when applying s 137,¹⁰ we argue that, in considering the probative value of incriminating expert evidence, attention should be directed to the results of (or the lack of) formal evaluation of its reliability. In advancing this position we take as axiomatic the fact that most forensic science and medicine techniques are susceptible to formal evaluation that provides vital information about the weight that might legitimately be accorded to evidence derived from them.¹¹ Such an approach to expert evidence is consistent with the purpose of s 137 within the context of an accusatorial trial. Regardless of how the tensions between *Dupas* and *XY* come to be resolved in relation to, for example, eyewitness identification or disputed admissions, we contend that incriminating expert evidence will remain a special class of evidence, and that its admissibility within the terms of s 137 must necessarily include some consideration, by the trial judge, of its validity and reliability.

This article begins with a principled restatement of the fundamental role of s 137 (and the corresponding common law *Christie* discretion). It outlines some of the key differences between *Dupas* and *XY* and, more critically for our purposes, the degrees of convergence between the two cases. Noting the points made in *Dupas* in relation to the evaluation of expert evidence, we then turn to consider some of the critical issues pertaining to expert evidence that have yet to be adequately addressed through conventional approaches to (incriminating) expert opinion. In advancing the need to attend to the probative value (and therefore the reliability) of expert evidence, we accept that many of our arguments could (and should) be employed at the earlier admissibility stages. Our focus, however, is primarily upon s 137 and is not constrained by what may or may not have been done in

10 (2012) 218 A Crim R 507, 524–5 [63].

11 It is not our intention to suggest that expert evidence possesses a proper probative value. Rather, testing usually provides some indication of the range of appropriate values (which might be moderated by other relevant considerations). We accept that some kinds of expert evidence are not readily susceptible to meaningful independent evaluation, but all have meaningful alternatives. The comparison techniques in routine use — which form the primary focus of this article — can and should be tested.

relation to relevance (s 55) or the exception to the opinion rule (s 76) for opinions based on specialised knowledge (s 79).¹²

II THE PURPOSE OF SECTION 137 AND *CHRISTIE*

A *Back to Basics*

Both at common law and under the UEL, any tendered evidence that is relevant — that is, capable of rationally affecting, directly or indirectly, alone or in combination with other evidence, the probability of a material fact in issue (s 55) — is admissible in proof of that fact unless subject to an exclusionary rule (s 56).¹³

Specific exclusionary rules, such as the proscription against opinion evidence, also cover similar ground at common law and under the UEL, although there are some differences. The general exclusionary rule found in s 137 reflects almost precisely the common law principle derived from *Christie*.¹⁴ While acknowledging the potential importance of the precise statutory language, both the common law and the UEL rules permit or demand the exclusion of prosecution evidence in a criminal case where the probative value of relevant evidence is outweighed by the danger of unfair prejudice to the defendant.

The purpose of the admissibility rules now contained in ss 135 and 137 should not be controversial. They seek to avoid inaccurate verdicts by preventing certain evidence going to the jury for its consideration.¹⁵ They provide a safety valve to exclude evidence which does not fall under a more definitive exclusionary rule where, although relevant (that is, to some degree probative), the tender of the

12 When it comes to expert evidence, both ss 55 and 137 require attention to be given to probative value and, therefore, the reliability of expert evidence. In s 55, the phrase ‘if it were accepted’ must be subservient to ‘could rationally affect’. For most types of forensic science and medicine evidence the relevance and weight should be informed by evaluation. Formal evaluation determines whether techniques work and therefore whether (and to what degree) they can be considered to ‘rationally effect’. Validation studies help answer the question of whether the analyst is better than chance and able to accomplish more than laypersons without assistance. It does not matter if the trier of fact accepts something if it has no logical ability to influence. For some discussion of probative value in relation to relevance and s 137, see *Papakosmas v The Queen* (1999) 196 CLR 297, 323 [86]; *Adam v The Queen* (2001) 207 CLR 96, 115 [59]–[60]; *R v Rahme* [2004] NSWCCA 233 (14 July 2004) [197]–[205].

13 At common law, *legal* relevance was slightly more circumscribed than the *logical* approach imposed in ss 55, 56 and the UEL Dictionary. To some extent, earlier notions of legal relevance were incorporated in s 135 of the UEL which, in addition to a concern about the potential of unfair prejudice to any party, also provides for discretionary exclusion on the basis that the probative value of the evidence is ‘substantially outweighed by the danger that the evidence might ... (b) be misleading or confusing; or (c) cause or result in undue waste of time’.

14 In *Christie*, the rule is regarded as permitting an exercise of discretion, while in s 137 the rule is expressed in mandatory terms. The difference is of limited practical consequence. The *Christie* discretion may also have some application in civil cases whereas s 137 applies only to the tender of prosecution evidence in a criminal case; cf s 135 which applies to all parties and forms of evidence. See also Australian Law Reform Commission, *Uniform Evidence Law*, Report No 102 (2005) 554–70; Tim Smith and Stephen Odgers, ‘Determining “Probative Value” for the Purposes of Section 137 in the Uniform Evidence Law’ (2010) 34 *Criminal Law Journal* 292.

15 Section 135 is also directed toward confusion, cost and delay.

evidence runs an unacceptable risk of producing a wrongful verdict because of the likelihood that the jury will accord it more weight than it rationally deserves, or because it will compromise the high criminal standard of proof.

The onus lies upon the accused to persuade the court that the evidence should be excluded.¹⁶ To secure exclusion, what must be established is the risk that the evidence in question will be inaccurately or irrationally assigned undue probative value in determining guilt. It is this risk that makes the evidence unfairly prejudicial. It must also be emphasised that the purpose of the safety valve is not to withdraw evidence from the jury merely because the judge regards it as unreliable (having little probative value).¹⁷ The general rule, expressed in s 56 remains — if the evidence is relevant it is admissible unless an exclusionary rule applies. Once evidence is admitted its probative value is a matter for the jury. For s 137 to apply, the judge must determine that there is an unacceptable risk that the jury, even with appropriate directions and warnings, will be incapable of adjudging accurately the probative value of the evidence in question (or will otherwise be misled by the evidence or the witness) and, as a consequence, give the evidence more weight than it deserves; reaching a verdict that does not meet the high criminal standard of proof. Unless the risk of misuse can be sufficiently mitigated by the probative nature of the evidence or by direction from the trial judge, the evidence must be excluded. This decision is expressed in the conclusion that the probative value of the evidence is outweighed by the unfair prejudice that would flow from its admission.

Working from the fundamental principles underpinning s 137, it would seem obvious that in determining exclusion via the safety valve, the judge may have to analyse the probative value of the evidence in deciding the risk of misuse. If the evidence has probative value and the jury is capable of assessing that probative value then the evidence may be *described* as having probative value outweighing any unfair prejudice.¹⁸ Though it is one thing to state the purpose of the safety valve and another to apply it in practice.¹⁹ As many judges have

16 See, eg, *R v Blick* (2000) 111 A Crim R 326, 332–3 [19]–[20].

17 Undoubtedly, *Christie* and s 137 raise dilemmas. If the evidence is admitted it may be misused, leading to a wrongful conviction. Alternatively, where the evidence is excluded the jury may be denied the ability to consider potentially probative evidence. This may lead to a mistaken acquittal. However, the issue is whether the danger of misuse can be rectified or managed. If not, then otherwise relevant evidence should be excluded.

18 For example, where it is argued that identification evidence should be excluded, the only issue is whether the jury is capable of accurately assessing its probative value. But, in some cases, where the probative value of the evidence may be accurately determined by the jury, exclusion might yet be sought on the basis that the evidence incidentally reveals information, or may have an emotional effect, which might prevent the jury from objectively applying the standard of proof; for example, where it is argued that photographs of a victim's injuries run the risk of biasing the jury against the accused so that it might either interpret the evidence irrationally and inaccurately, or fail to rigorously apply the criminal standard of proof. Such a case might more easily be described as involving a balance of probative value against unfair prejudice.

19 In particular, how does one assess the capabilities of juries so as to assess the risk of the evidence being given inappropriate weight? Critically (as discussed further below), can one assume that directions given to a jury will be understood and enable it to undertake its task rationally and accurately? A lawyer can look to past decisions where information about the capabilities of juries is assumed by judges to lie. But an empiricist might be more skeptical as to what these capabilities really are.

recognised, this exercise is not just a matter of weighing probative value and unfair prejudice; the exercise is not advanced by attempting a balancing act based on literal interpretations of the words ‘probative value’ and ‘unfair prejudice’.²⁰ As the recent judgments illustrate, such attempts simply confuse the nature of the exercise demanded by *Christie* and s 137, by placing the emphasis upon arcane legal interpretation — for example, asserting that the definition of ‘probative value’ in the UEL dictionary compels the judge to take the probative value of the evidence in dispute ‘at its highest’ in order to conduct some sort of balancing exercise — rather than attempting to determine whether the reception of the evidence runs the risk of being misunderstood and/or overvalued if left to the jury.²¹ Regrettably, the UEL seems to have encouraged judges to take this arcane approach and, in decisions following *XY* (in New South Wales), there are signs that this posture may be hardening.²²

B Section 137, Dupas and XY

In *Dupas* and *XY*, significant decisions of five-judge benches of the appellate courts of Victoria and New South Wales respectively, the question of the extent to which s 137 requires trial judges to exclude evidence on the basis of its potential unreliability was revisited. As discussed in detail by Adams and Wareham, an apparent difference of opinion has emerged as to the correct approach to the assessment of probative value by the trial judge, when determining whether contested incriminating evidence should go to the jury.²³ With some exceptions, the New South Wales cases have generally been interpreted as taking a ‘restrictive’ or ‘narrow’ approach to the role of the trial judge, limiting their capacity to evaluate the reliability of evidence prior to admission.²⁴

Against this trend, in *Dupas* the Victorian court held that the ‘restrictive approach’ — conventionally associated with *Shamouil*²⁵ — was ‘manifestly wrong’,²⁶ and recognised in the trial judge a greater power to intervene to exclude unfairly prejudicial evidence. In summarising its position, a unanimous bench insisted that:

- (a) The common law did require the trial judge, in assessing probative value, to evaluate the weight that the jury could rationally attach to the evidence. The contrary conclusion was inconsistent with a continuous line of High Court authority.

20 See, eg, *Pfennig v The Queen* (1995) 182 CLR 461, 514–15 (McHugh J); *Bendix Autolite Corp v Midwesco Enterprises Inc*, 486 US 888, 897 (1988) (Scalia J).

21 This is the approach associated with *R v Shamouil* (2006) 66 NSWLR 228 (‘*Shamouil*'). See discussion below.

22 See, eg, *R v Burton* [2013] NSWCCA 335 (20 December 2013) [158]–[170] (Simpson J, RA Hulme J and Barr AJ agreeing).

23 Adams and Wareham, above n 2.

24 *XY* (2013) 84 NSWLR 363, 381 [64] (Basten JA).

25 (2006) 66 NSWLR 228, 237 [60].

26 *Dupas* (2012) 218 A Crim R 507, 524 [63].

- (b) The legislative intention, as disclosed by the language of s 137 and its context, is that the task under s 137 is the same as that at common law.²⁷

Less than six months later, the New South Wales Court of Criminal Appeal in *XY* declined to adopt the reasoning in *Dupas*, affirming *Shamouil* as the leading authority in New South Wales. However, it is difficult to draw a clear ratio from the five separate judgments in *XY*.²⁸ As indicated in the judgment of Basten JA, some of the differences between *Dupas*, *Shamouil* and *XY* may in fact be more apparent than real. To some extent, it could be said that the Court of Criminal Appeal reinterpreted, rather than offered a straightforward reaffirmation of *Shamouil* and earlier New South Wales authorities.²⁹ Perhaps paradoxically — while the analysis and statement of principle in *Dupas* appears to favour a more interventionist and exclusionary approach, the evidence was held to have been correctly admitted; whereas the evidence in *XY* was held to have been correctly excluded. Nevertheless, in holding that *Shamouil* remains the law in New South Wales, the decision in *XY* is generally understood to have reasserted the more restrictive position, and in this respect continues to act as a barrier to judicial consideration of the probative value of contested forensic science evidence.³⁰

One way of approaching the convoluted s 137 jurisprudence is to identify the various ways in which courts have sought to balance jury deference with the trial judge's responsibility to weigh probative value. Clearly s 137 requires the trial judge to engage in *some* kind of assessment of the probative value of evidence. However, in delineating the scope of the trial judge's role, the courts have been wary of usurping the jury's 'constitutional'³¹ and 'traditional'³² role as fact-finder. This is captured in the insistence that the trial judge should take the evidence 'at its highest'³³ and leave it to the jury to weigh (that is, determine) its 'actual probative value'.³⁴

The court in *Dupas* suggested that the responsibilities of the jury and the trial judge are balanced, both at common law and under the UEL, by assigning each to a separate domain. Witness honesty is the exclusive domain of the jury.³⁵ In this respect at least, evidence is taken at its highest; doubts about honesty provide no basis for the trial judge to exclude incriminating evidence under s 137. The jury

27 Ibid.

28 See Adams and Wareham, above n 2.

29 *XY* (2013) 84 NSWLR 363, 377–81 [50]–[65]. In *XY*, as noted below, Basten JA goes to some lengths to point out that the Victorian Court has misunderstood, and thus overstated, the restrictive effect of *Shamouil*: *ibid*.

30 See, eg, *R v Jacobs [No 2]* [2013] NSWSC 943 (12 June 2013) [26]–[31]; *R v Burton* [2013] NSWCCA 335 (20 December 2013); *LP v The Queen* [2013] NSWCCA 330 (23 December 2013) [82]–[84]; *Lau v The Queen* [2014] NSWCCA 179 (12 September 2014) [106]. See discussion of the interpretation of s 137 in Judicial College of Victoria, *Uniform Evidence Manual* (2 June 2014) <<http://www.judicialcollege.vic.edu.au/eManuals/UEM/index.htm#28737.htm>>.

31 *XY* (2013) 84 NSWLR 363, 371 [25] (Basten JA).

32 *Shamouil* (2006) 66 NSWLR 228, 238 [64].

33 See, eg, *XY* (2013) 84 NSWLR 363, 376 [46] (Basten JA).

34 See, eg, *ibid* 400 [167] (Simpson J).

35 See, eg, *Dupas* (2012) 218 A Crim R 507, 524–5 [63], 552 [162], 559 [184], 562 [191].

is regarded as perfectly capable of determining a witness' honesty. According to the Court in *Dupas*, the trial judge has freer rein to assess other aspects of witness credibility, and also to determine the jury's capacity to assess the strength of the inference of guilt that the evidence, if credible, would support.³⁶ Though similar statements can be found elsewhere in the case law,³⁷ such sharp, apparently clear distinctions are not entirely persuasive. In some circumstances, grave doubts about witness honesty deplete probative value to such a degree that to even admit the evidence would be misleading.³⁸ Furthermore, in relation to much incriminating expert evidence: honesty, credibility and reliability are often inextricably interwoven.

Without endorsing the sharp distinction between witness honesty and other aspects of evidentiary reliability and weight in *Dupas*, the various judgments in *XY* balanced trial judge and jury responsibilities in different ways. Price J was most favourably disposed toward trial judge intervention. His Honour expressed broad support for the interventionist principles in *Dupas*, indicating that taking evidence 'at its highest in favour of the Crown' may allow 'evidence being before a jury which in reality ... has little probative value and is outweighed by the danger of unfair prejudice to the accused'.³⁹ According to Price J, 'enabling the trial judge to consider questions of credibility, reliability or weight ... is likely to enhance the fundamental principle that an accused is to receive a fair trial'.⁴⁰

At the other end of the spectrum, Simpson J was most deferential to the jury. Her Honour suggested that 'where a court is considering ... s 137, questions of credibility, reliability or the weight to be attributed to the evidence in question have no part to play'.⁴¹ This appears to endorse a rather extreme interpretation of *Shamouil* under which the trial judge retains a highly constrained, limited capacity to exclude evidence.⁴² If the trial judge must take the evidence 'at its highest' in all these respects, then the evidence would be considered to have maximal probative value, leaving very little scope for probative value to be outweighed by the danger of unfair prejudice. As we shall see, much unfair

36 Ibid 524–5 [63].

37 See, eg, Stephen Odgers, *Uniform Evidence Law* (Thomson Reuters, 11th ed, 2014) 857; *R v Best* [1998] 4 VR 603, 610–11. See also the statements of the High Court in *Doney v The Queen* (1990) 171 CLR 207: 'the purpose and the genius of the jury system is that it allows for the ordinary experience of ordinary people to be brought to bear in the determination of factual matters. It is fundamental to that purpose that the jury be allowed to determine, by inference from its collective experience of ordinary affairs, whether and, in the case of conflict, what evidence is truthful': at 214.

38 Evidence of an accomplice — see, eg, *Rozenes v Beljajev* [1995] 1 VR 533, cited in *Dupas* (2012) 218 A Crim R 507, 542 [124]; unsigned records of interview containing a purported admission — see, eg, *Driscoll v The Queen* (1977) 137 CLR 517, 542, cited in *Dupas* (2012) 218 A Crim R 507, 540 [116]–[117]; coincidence evidence of multiple alleged victims which may be the product of concoction — see, eg, *PNJ v DPP* (2010) 27 VR 146, of which the court in *Dupas* (2012) 218 A Crim R 507 observes '[w]hether those circumstances were viewed as going to credibility or to reliability, they have been consistently treated as matters for the judge to assess in the evaluative process, notwithstanding *Shamouil*': at 553 [165].

39 *XY* (2013) 84 NSWLR 363, 409 [225].

40 Ibid 408 [224].

41 Ibid 386 [97]. See also at 381 [64] (Basten JA).

42 But see, eg, *R v Cook* [2004] NSWCCA 52 (12 March 2004) [43]; *R v Smith [No 3]* [2014] NSWSC 771 (3 June 2014).

prejudice arises from misunderstanding, misusing and particularly overvaluing expert evidence. Taking the evidence ‘at its highest’ may mean that real dangers of unfair prejudice that flow from an overinflated acceptance of the evidence’s probative value are not able to be explored or considered. Allowing the trier of fact to rely upon weak, speculative or unreliable expert evidence often renders the section and its protection against a range of dangers nugatory. This point is made strongly in *Dupas*, and is a critical factor in the Court’s rejection of the principle of taking the evidence ‘at its highest’ as an interpretive gloss on s 137.⁴³

Basten JA also confirmed that *Shamouil*’s ‘general approach’ remains the law in New South Wales, though as noted he understands *Shamouil* to be less restrictive in effect.⁴⁴ Consequently, his Honour accepts that the trial judge may, in some circumstances, consider probative value. For Basten JA, the trial judge does not attempt to assess ‘actual weight’, but rather ‘the capacity of the evidence to support the prosecution case’.⁴⁵ This approach is consistent with the legislative definition of ‘probative value’ — ‘the extent to which the evidence could rationally affect the assessment of the probability of ... a fact in issue’.⁴⁶ Basten JA contends that it was ‘erroneous’ for the Victorian Court to suggest that ‘*Shamouil* ... conclud[ed], inflexibly and without qualification, that the weight of the evidence was irrelevant’.⁴⁷

The difference between a trial judge’s assessment of actual weight and their assessment of the weight that could rationally be assigned to a piece of evidence depends upon the extent to which it is recognised that there may be a ‘great diversity of reasonable opinion ranging between widely set limits’.⁴⁸ The same issue arises in the context of an appeal, where an appellate court is asked to determine whether a conviction should stand. In some cases appellate courts have deferred to the jury on the basis that ‘there may be reasonable views of facts which do not commend themselves to the court’.⁴⁹ On such a view, an appellate court may uphold a conviction notwithstanding its own doubts as to whether the evidence is sufficiently strong. Other appellate courts have been more prepared to intervene, insisting that if the appellate court doubts the strength of the evidence, ‘a reasonable jury should be of a like mind’.⁵⁰ As in the context of s 137, the former approach pays greater respect to the constitutional role and competence invested in the jury, while the latter approach provides greater protection against

43 (2012) 218 A Crim R 507, 563 [194], 565 [206]–[211]. See also *DSJ v DPP (Cth)* (2012) 259 FLR 262 (*‘DSJ’*); *R v Burton* [2013] NSWCCA 335 (20 December 2013).

44 *XY* (2013) 84 NSWLR 363, 381 [65].

45 *Ibid* 371 [25].

46 UEL Dictionary.

47 *XY* (2013) 84 NSWLR 363, 380 [62].

48 *Naxakis v Western General Hospital* (1999) 197 CLR 269, 292 [64] (Kirby J), quoting Harold H Glass, Michael H McHugh and Francis M Douglas, *The Liability of Employers in Damages for Personal Injury* (Law Book, 2nd ed, 1979) 217.

49 *Ratten v The Queen* (1974) 131 CLR 510, 519 (Barwick CJ) (*‘Ratten’*). The related question here is the extent to which the criminal appeal court should defer to the jury verdict in determining whether the conviction is unreasonable or unsupportable: see, eg, *Criminal Appeal Act 1912* (NSW) s 6.

50 *Ratten* (1974) 131 CLR 510, 516 (Barwick CJ). Again, the context of this observation was a criminal appeal on the basis that the verdict was unreasonable or unsupported.

the risk of an unfair trial and a wrongful conviction. Both approaches may lie on the same ‘rational capacity’ spectrum.⁵¹

A further dimension to the question of the trial judge’s assessment of probative value concerns the breadth of information to which the trial judge may refer. There is a procedural dimension to this question. Typically, s 137 is considered prior to trial, at which point the trial judge does not yet have full access to the other evidence (or to the defence’s response). Simpson J emphasised the impracticality of a trial judge attempting a broad contextual assessment at this preliminary stage.⁵² However, as Price J explains, these practical difficulties are not necessarily insuperable.⁵³

In *XY* the ‘other information’ issue was characterised in terms of whether, and the extent to which, the trial judge may have regard to inferences from, or explanations for, the disputed evidence, other than those relied upon by the Crown. Again, taking a strongly deferential line, Simpson J suggested that assessing probative value does not ‘involv[e] the trial judge in weighing the relative merits of competing explanations’ for the evidence.⁵⁴ Basten JA suggested that s 137 may require the trial judge to ‘consider the weight to be given to each possible set of inferences’,⁵⁵ although the competing inferences may be discounted quite swiftly on the basis that ‘[d]etermining which inferences should be accepted was quintessentially a function for the jury’.⁵⁶ The majority judgments in *XY* recognised greater scope for ‘the possibility of such an alternative explanation [to] substantially alter [the trial judge’s] view as to the significant capacity of the Crown evidence’.⁵⁷

The danger that the evidence is not susceptible to rational evaluation or will be misused by the trier of fact is the starting point. As mentioned above, low probative value is not itself sufficient for exclusion. If there is ‘no real risk of unfair prejudice ... s 137 ... [is] not engaged’.⁵⁸ Where the defence can provide a persuasive competing interpretation of the disputed evidence, this may both reduce the evidence’s probative value, and reveal the risk that it will be overvalued by a jury that accepts the prosecution version without due consideration of alternative interpretations.

Apart from Parts IIA and IIB, this article is not concerned with the application of s 137 at large. Rather, our attention is focused on expert evidence where, as is often the case, the reliability of the evidence is not interrogated at the point of admission such that its probative value is unknown. It will be argued that,

51 See, eg, John T McNaughton, ‘Burden of Production of Evidence: A Function of a Burden of Persuasion’ (1955) 68 *Harvard Law Review* 1382; David Hamer, ‘The Privilege of Silence and the Persistent Risk of Self-Incrimination: Part I’ (2004) 28 *Criminal Law Journal* 160, 172.

52 *XY* (2013) 84 NSWLR 363, 400 [170].

53 *Ibid* 408 [224].

54 *Ibid* 400 [166].

55 *Ibid* 383 [73].

56 *Ibid* 382 [68].

57 *Ibid* 405–6 [202] (Blanch J), quoting *DSJ* (2012) 259 FLR 262, 278 [78] (Whealy J). See also *XY* (2013) 84 NSWLR 363, 385 [88] (Hoeben CJ at CL), 408–9 [223]–[225] (Price J).

58 *XY* (2013) 84 NSWLR 363, 383 [73] (Basten JA).

having regard to the particular qualities of such evidence, the conventional trial safeguards of cross-examination, jury directions and warnings do not necessarily offer a meaningful corrective.⁵⁹ Before turning to explain our position in more detail, it is useful to summarise the Victorian Court of Appeal's statements about expert evidence in *Dupas*.

C Expert Evidence in Dupas (and XY)

The Court in *Dupas* recognised a general ability to attend to reliability when considering the weight that could be 'rationally attach[ed]' to the evidence, and considered this to be incorporated in the terms of s 137.⁶⁰ In undertaking a brief survey of potentially problematic evidence (including expert evidence), the Court was open to the possibility of directly attending to the weight of the evidence, and that consequently reliability may be an appropriate question for the trial judge to consider when determining the risk of misuse or misunderstanding. In doing so, the Court accepted that overlooking probative value and disregarding the reliability of expert evidence has the potential to eviscerate what is intended to be a substantial protection against unfair prejudice to the accused.⁶¹ The Court indicated that, '[e]xpert evidence has often been excluded in the exercise of the discretion where the trial judge was dissatisfied with its reliability and probative value'.⁶² Additionally, in discussing the importance of threshold questions of relevance, as well as the role of reliability in the context of s 79, the Court emphasised that the risks associated with the admission of expert evidence of little (or no) weight are 'important consideration[s]' in the context of discretionary exclusion.⁶³ While recognising that 'modern attitudes towards [the admission of] expert evidence may be less exclusionary than in the past',⁶⁴ the *Dupas* Court is

59 It is possible to entertain significant reservations about the actual effectiveness of jury directions with respect to the contested evidence in *Dupas* and *XY*. There are real questions, especially with respect to the identification evidence in *Dupas*, as to the extent that conventional directions are capable of offering useful guidance to a jury. For a discussion of directions, see, eg, Gary Edmond, Kristy Martire and Mehera San Roque, 'Unsound Law: Issues with ("Expert") Voice Comparison Evidence' (2011) 35 *Melbourne University Law Review* 52. See also the new *Jury Directions Act 2013* (Vic), drafted in response to concerns about the length, effectiveness and complexity of jury directions; Victorian Law Reform Commission, *Jury Directions*, Final Report No 17 (2009); Justice Mark Weinberg, 'Simplification of Jury Directions Project: A Report to the Jury Directions Advisory Group' (Report, Judicial College of Victoria and Department of Justice, August 2012). The New South Wales Law Reform Commission has similarly reported on concerns about the effectiveness of conventional jury directions, although its recommendations fall short of the substantial procedural reform undertaken in Victoria: New South Wales Law Reform Commission, *Jury Directions*, Report No 136 (2012).

60 *Dupas* (2012) 218 A Crim R 507, 524 [63].

61 *Ibid* 544 [132]. The Court stated '[a]cademic writers have argued strongly that, as a consequence of *Shamouil*, exclusionary discretions have been "emasculated" in the area of expert evidence'.

62 *Ibid* 542 [125], citing *R v Elliott* (Unreported, Supreme Court of New South Wales, Hunt J, 6 April 1990) ('*Elliott*'); *R v Lucas* [1992] 2 VR 109 (Hampel J); *R v Green* (Unreported, New South Wales Court of Criminal Appeal, 26 March 1993); *R v Pantoja* (1996) 88 A Crim R 554 (Hunt CJ at CL). While judges do on occasion exclude expert evidence using s 137, exclusion appears to have been less frequent in recent years.

63 *Dupas* (2012) 218 A Crim R 507, 543 [128].

64 *Ibid* [125].

not entirely sanguine about the implications of this development with respect to s 79 jurisprudence, nor with exclusionary powers inherent in ss 135 and 137.⁶⁵

On the subject of image comparison (so-called face and body mapping evidence) and voice comparison evidence for the purpose of identification, the Court accepted that the failure to consider the probative value of incriminating expert opinions meant that:

expert comparison evidence and cross-lingual voice comparisons were expected to be undertaken by lay jurors where the probative value of the evidence was unknown, the risk of error high and the dangers associated with a range of biases and exaggerated confidence, significant and unlikely to be conveyed to the jury. Yet exclusionary powers designed to protect the accused from the danger of unfair prejudice lay idle.⁶⁶

Clearly, neither the contested evidence, nor the approach of the court in *XY* lent itself to such a broad survey of the jurisprudence. Nonetheless, in turning to consider expert evidence in the aftermath of *Dupas* and *XY*, the most important point to make is that both decisions, to a greater and lesser extent, accept that there may be a need to attend to the probative value, and therefore the reliability, of incriminating expert evidence in determining the risk that the jury might give the evidence more probative value than it deserves. This, as the previous sections explain, was explicit in *Dupas* and its reading of both the common law and its substantial embodiment in s 137. In *XY*, Price J clearly expressed a preference for the approach in *Dupas*, and Basten JA recognised that, in some limited circumstances, it may be appropriate to consider reliability in the context of assessing the danger of unfair prejudice.⁶⁷

The strong commitment of the *Dupas* Court to intervention is also reflected in its invocation of *Elliott*.⁶⁸ There, Hunt J (who also delivered judgment in *R v Carusi*)⁶⁹ characterised the common law trial judge's duty in the following, rather expansive, terms:

If scientific testing in the particular case is unreliable or if it has a tendency to produce a misleading or confusing impression to the jury or if the weight to be afforded to the result is so minimal as to preclude the jury being satisfied beyond reasonable doubt that the Crown has established the fact which it seeks to prove, then clearly I have a duty to exclude it from

65 Both courts adverted to evolving jury abilities and modern judicial attitudes without providing any evidence that modern juries are able to cope with increasingly sophisticated techniques and statistical expressions.

66 *Dupas* (2012) 218 A Crim R 507, 544 [131], quoting Gary Edmond, Kristy Martire and Mehera San Roque, "'Mere Guesswork": Cross-Lingual Voice Comparisons and the Jury' (2011) 33 *Sydney Law Review* 395, 420.

67 (2013) 84 NSWLR 363, 380 [62], 381–2 [67], 383 [73] (Basten JA), 408 [224] (Price J). We note that Basten JA makes clear that this will only occur exceptionally in the sense that there may only be limited classes of evidence that warrant such evaluation prior to admission. We explain below why expert evidence is such a class of evidence.

68 (Unreported, Supreme Court of New South Wales, Hunt J, 6 April 1990).

69 (1997) 92 A Crim R 52.

the jury — whether it is a result of ruling that the evidence is inadmissible or whether it is excluded in the exercise of my discretion.⁷⁰

The Court in *Dupas* emphasised the need to ensure that the jury is equipped to rationally assess weight, and concluded that assessing reliability will often be a necessary corollary of that assessment.

There are, as indicated, many unresolved tensions between *Dupas* and XY. It is not our intention, and we are in no position, to attempt to resolve them. Rather, having reiterated the historical rationale for *Christie* and s 137 and aspects of the judgments, our purpose is to explain why the basic approach in *Dupas* — and its willingness to attend to the actual probative value (and reliability) of the evidence — is the most appropriate way of managing the very real dangers attending the increasing reliance placed upon forensic science and medicine evidence in criminal proceedings.

III THE PROBATIVE VALUE OF EXPERT EVIDENCE AFTER DUPAS (AND XY)

At the outset, it is important to make clear that our focus on probative value and reliability is not for the purposes of determining admissibility per se. This, as some of us have argued elsewhere, is the appropriate role for s 79(1).⁷¹ Rather, our concern with probative value and reliability in relation to s 137 is to enable the trial judge to determine whether incriminating expert evidence is susceptible to rational evaluation within the confines of criminal proceedings. In the absence of information about the reliability of forensic science and medicine techniques, there is, as we now explain, a heightened risk of unfair prejudice created by the trier of fact misunderstanding, misusing or overvaluing the evidence and improperly deferring to ‘experts’.

In explaining the importance of attending to the probative value of incriminating expert evidence we intend to build on *Dupas* and XY. We suggest that forensic science and medicine evidence should, in general, be treated as a special class of evidence. Because of the danger of unfair prejudice arising from the admission of incriminating ‘expert’ opinions that are mistaken, misleading or exaggerated, prosecutors and judges should direct their attention to the probative value of the evidence. We contend that there is space within the current jurisprudence to treat forensic science and medicine evidence as a special class, and that to do so is

70 *Elliott* (Unreported, Supreme Court of New South Wales, Hunt J, 6 April 1990), quoted in *Dupas* (2012) 218 A Crim R 507, 543 [128]. While noting the decision in *R v Jung* [2006] NSWSC 658 (29 June 2006), the Court in *Dupas* also drew attention to the fact that Spigelman CJ (who delivered judgment in *Shamouil* (2006) 66 NSWLR 228) had, in another case, ‘found that it was a fundamental error for the trial judge not to have undertaken a systematic analysis of the probative value of the expert’s evidence for the purpose of considering discretionary exclusion under s 135’ in a civil trial: *Dupas* (2012) 218 A Crim R 507, 544 [130], citing *Australian Securities and Investments Commission v Rich* (2005) 218 ALR 764, 799 [163].

71 Edmond, ‘Specialised Knowledge, the Exclusionary Discretions and Reliability’, above n 7; Edmond, ‘The Admissibility of Forensic Science and Medicine Evidence’, above n 8.

consistent with the broad purpose of s 137 (as well as s 135, and the limits of the trial). In addition, such an approach is preferable to one that attempts to lay down a prescriptive rule for all evidence, especially if such an approach tends towards the narrow or restrictive approach associated with Simpson J's reading of *Shamouil*. There are good reasons to think that many types of expert opinion evidence, especially those that have not been evaluated or which carry serious methodological limitations, might justifiably be brought within the scope of s 137 on the basis of diminished probative value — recognised by Basten JA in determining the risk of the evidence being given undue weight by the jury.

The inclusionary trends referred to by the Court in *Dupas* are not, in the absence of widespread and systematic evaluation, an appropriate response to increasingly technical forms of expert evidence. There is instead a need for some kind of indication, ideally through formal evaluation (for example, validation), that forensic science techniques are demonstrably reliable and that analysts are proficient with demonstrably reliable techniques.⁷² The need for demonstrable reliability is fundamental because in many cases it is not appropriate, and may not even be rational, to attempt to evaluate opinions derived from techniques that can and should have been tested but have not been.⁷³ This means that in cases where there has been no formal evaluation, we cannot be confident that the trier of fact will not misuse or misunderstand the evidence. Indeed, in some circumstances the lack of information about the validity of techniques and the proficiency of analysts will mean that we cannot be confident that this has not occurred. The adversarial criminal trial does not usually provide an effective (or an appropriate) forum to explore subtle, and frequently complex, methodological, technical and psychological issues associated with expert evidence, let alone contested expertise. The result is a very real risk that forensic science evidence will be given more weight than it deserves.

A Excluding Forensic Science Evidence under Section 137: A Special Case?

In thinking about the application of s 137 to incriminating expert evidence, a vital consideration is whether forensic science (and forensic medicine) evidence should be treated like other forms of incriminating evidence. Without seeking to descend

72 *NAS Report*, above n 6; *NIST Report*, above n 6.

73 This applies equally to judicial attempts to make a rational assessment of what 'taken at its highest' might actually mean.

into complex and abstruse epistemological terrain,⁷⁴ it is our contention that there are compelling practical reasons to direct attention toward probative value and reliability when s 137 is raised in response to expert evidence.⁷⁵

Forensic science and medicine evidence are different from the normal course of sensory descriptions and the categories of lay opinion tolerated by the common law and the UEL. Unlike ordinary witnesses, those proffering forensic science evidence — really ‘opinions based on specialised knowledge’⁷⁶ — are not simply sensory (or direct) witnesses. They are allowed to testify on the basis that they offer something beyond what the trier of fact already knows or can do for itself.⁷⁷ This is why they are allowed to proffer opinions derived from identified facts (whether proved or assumed or observed by them) based on ‘knowledge’ linked to their ‘training, study or experience’.⁷⁸ Even if admitted with a proper foundation, forensic science and medicine evidence is often complex and not always readily understood by judges and jurors.⁷⁹ Lay people do not usually engage in the assessment of technical, methodological and statistical issues.⁸⁰ In contrast to most forms of witness testimony, laypersons are not usually in a good position to assess the opinions of ‘experts’, or the methods and reasoning behind them. Notwithstanding the adversarial nature of proceedings, jurors are likely to assume, understandably, that the admission of ‘expert’ opinion reflects the court’s imprimatur. Expert evidence, especially opinions proffered by state-employed forensic analysts, may exert a disproportional influence on judges and jurors. There is a manifest danger that trial safeguards will not place the jury in a position to appreciate the magnitude of methodological shortcomings and the weakness of evidence, even if these are explained.

A further, related reason why forensic science evidence ought to be treated with great care is because most techniques, along with the abilities of the analysts

74 Cf Frederick Schauer and Barbara A Spellman, ‘Is Expert Evidence Really Different?’ (2013) 89 *Notre Dame Law Review* 1. We are not epistemic essentialists or reductionists. Rather, in adopting a pragmatic course informed by legal principle, we contend that many techniques can and should be tested in order to provide the decision-maker with the kinds of information that enable them to make use of interpretive opinions. In general, philosophical attempts to demarcate science from non-scientific activities have not proven effective. See Larry Laudan, ‘The Demise of the Demarcation Problem’ in Michael Ruse (ed), *But Is It Science? The Philosophical Question in the Creation/Evolution Controversy* (Prometheus Books, 1996) 337; John A Schuster and Richard R Yeo (eds), *The Politics and Rhetoric of Scientific Method* (Reidel Publishing, 1986); Thomas F Gieryn, *Cultural Boundaries of Science: Credibility on the Line* (University of Chicago Press, 1999).

75 There may be grounds for treating other ‘classes’ of evidence with similar caution, but these will need to be made on a case-by-case (or class) basis.

76 UEL s 79.

77 But see UEL s 80.

78 UEL s 79(1); *Makita (Australia) Pty Ltd v Sprowles* (2001) 52 NSWLR 705; *Davie v Magistrates of Edinburgh* [1953] SC 34; *Dasreef Pty Ltd v Hawchar* (2011) 243 CLR 588.

79 This is especially the case with statistical forms of evidence: see Kristy A Martire, Richard I Kemp and Ben R Newell, ‘The Psychology of Interpreting Expert Evaluative Opinions’ (2013) 45 *Australian Journal of Forensic Sciences* 305; Dawn McQuiston-Surrett and Michael J Saks, ‘The Testimony of Forensic Identification Science: What Expert Witnesses Say and What Factfinders Hear’ (2009) 33 *Law and Human Behavior* 436.

80 Cf Alan Irwin and Brian Wynne (eds), *Misunderstanding Science? The Public Reconstruction of Science and Technology* (Cambridge University Press, 1996); Steven Epstein, *Impure Science: AIDS, Activism, and the Politics of Knowledge* (University of California Press, 1996).

who use them, can (and should) be formally assessed through validation and/or rigorous proficiency studies.⁸¹ Such studies are vital in accurately gauging the probative value of this kind of evidence. Ideally such assessment should be carried out before the technique and derivative opinion are relied upon. The ability to evaluate underlying techniques, particularly those in regular use, distinguishes forensic science and medicine evidence from the descriptive (that is, factual) reports provided by other sensory witnesses. Formal evaluation and evidence of proficiency are particularly important because, as we explain below, it is difficult to rationally evaluate most techniques and derivative opinions unless information about validity, indicative error rates and proficiency is provided.

Before moving to explain the need to attend to formal evaluation, we recognise that courts have not been particularly interested in reliability. Historically, concerns have been mainly directed to partisanship and the threats posed to fact-finding by exaggerated influence and deference.⁸² In *HG v The Queen*,⁸³ Gleeson CJ reiterated these concerns:

Experts who venture ‘opinions’ (sometimes merely their own inference of fact), outside their field of specialised knowledge may invest those opinions with a spurious appearance of authority, and legitimate processes of fact-finding may be subverted.⁸⁴

We agree that experts expressing opinions ‘outside their field of specialised knowledge’ introduce the danger of the jury being overwhelmed or misled. However, we contend that this risk is not restricted to experts who trespass beyond the legitimate reach of their ‘field’ or actual expertise. Rather, and more troublingly, when it comes to forensic science and medicine evidence, real risks arise whenever a witness, recognised by a court as an expert — however qualified and however experienced — expresses opinions based on techniques and abilities (that is, proficiency) that have not been formally evaluated.⁸⁵ Perhaps the most curious feature of longstanding judicial anxiety is the putative concern with boundary transgression but strangely incongruous insensitivity to the (validity and) reliability of techniques and proficiency. In most cases professional boundaries and claims about ‘fields’ provide less information, and less valuable insights, than attention to evidence demonstrating that techniques are valid and reliable, and that analysts are proficient in their use. Indeed, it is the absence of

81 *NAS Report*, above n 6.

82 See, eg, *Wood v The Queen* (2012) 84 NSWLR 581.

83 (1999) 197 CLR 414.

84 *Ibid* 429 [44]; See also *Murphy v The Queen* (1989) 167 CLR 94, 131 (Dawson J); *Wood v The Queen* (2012) 84 NSWLR 581; *Gilham v The Queen* (2012) 224 A Crim R 22; *Morgan v The Queen* (2011) 215 A Crim R 33, discussed in Gary Edmond, David Hamer and Andrew Ligertwood, ‘Expert Evidence after *Morgan*, *Wood* and *Gilham*’ (2012) 112 *Precedent* 28.

85 Studies suggest that witnesses characterised (or understood) as experts exert disproportional influence when they are no more capable than ordinary persons and even when they are mistaken. Confidence, persuasiveness and ability to withstand quotidian cross-examination may reveal (or convey) little about ability, reliability or probative value. See Richard Kemp, Stephanie Heidecker and Nicola Johnston, ‘Identification of Suspects from Video: Facial Mapping Experts and the Impact of Their Evidence’ (Paper presented at 18th Conference of the European Association of Psychology and Law, Maastricht, 2–5 July 2008).

information about validity, error rates and proficiency that introduces the most serious and pernicious threat to fact-finding posed by expert evidence as a result of juries investing evidence of unknown probative value with inappropriate (that is, exaggerated) weight.

This article is primarily concerned with forensic science and medicine techniques in regular use. We are most interested in the comparison, pattern recognition and identification ‘sciences’. These are typically used to identify persons or to link some trace to a person or object. They include DNA profiling, latent fingerprint comparison, ballistics, bite marks, shoe, tyre and foot impressions, handwriting and documents, paint fragments, arson investigation, vehicle collision, drug and fire chemistry, comparing CCTV images and sound recordings, and so on. All of these techniques and practices can be assessed to provide an indication of their probative value and limits (through validation studies and indicative error rates), the kinds of protocols and standards that are appropriate and should be in place, as well as to inform the expression of opinions. The proficiency of individuals in the use of validated techniques can also be assessed.⁸⁶ Most of the techniques that are in routine use and regularly relied upon in criminal proceedings are susceptible to formal evaluation and should be formally evaluated. Surprisingly few have been assessed.⁸⁷ Until formally assessed, the risk of juries giving such evidence inappropriate weight persists. In principle, admission should be regulated by s 79. Unless s 79 jurisprudence and practice is reformed, the main weapon against weak, speculative and unreliable opinion evidence causing miscarriages of justice is s 137.

B Can Forensic Science Evidence Be Rationally Assessed without Formal Evaluation?

Perhaps our strongest argument in favour of courts giving careful consideration to the exclusion of forensic science and medicine evidence under s 137 is the difficulty of rationally assessing the probative value — and therefore the attendant danger to the defendant — raised by many, perhaps most, techniques without the insights provided through formal evaluation. That is, there is a need to provide information about whether the technique works, how well, in what conditions, and whether the analyst is proficient using the technique, based on some kind of validation study (or rigorous proficiency test) conducted in conditions where the correct answers (that is, ground truth) are known to those conducting the evaluation.⁸⁸

86 Jonathan J Koehler, ‘Fingerprint Error Rates and Proficiency Tests: What They Are and Why They Matter’ (2008) 59 *Hastings Law Journal* 1077; Jonathan J Koehler, ‘Proficiency Tests to Estimate Error Rates in the Forensic Sciences’ (2013) 12 *Law, Probability and Risk* 89.

87 *NAS Report*, above n 6, 7–8, 87. The report goes on to state: ‘With the exception of nuclear DNA analysis, however, no forensic method has been rigorously shown to have the capacity to consistently, and with a high degree of certainty, demonstrate a connection between evidence and a specific individual or source’: at 7.

88 See, eg, Jason M Tangen, ‘Identification Personified’ (2013) 45 *Australian Journal of Forensic Sciences* 315.

Legal institutions have performed poorly in response to forensic science and medicine evidence and contemporary practice is conspicuously inconsistent with the best scientific advice.⁸⁹ Uncomfortable as it might be, for those steeped in the adversarial tradition, the legal heuristics conventionally used to gauge the admissibility and probative value of expert evidence are not especially informative. Common law and UEL categories such as formal qualifications, experience, legal recognition of a ‘field’, apparent impartiality, and previous admission reveal little, if anything, about the probative value of particular techniques. They are inferior to knowing about the conditions in which the technique works, how well it works and the analyst’s proficiency using the technique. These later considerations are the factors that afford primary insight into reliability and, therefore, probative value. On reflection, it must be more useful to *know* that a technique works and that the analyst is proficient in its use than to be told that the analyst has a formal qualification and has done this sort of thing before.

Many techniques routinely used in criminal investigations and prosecutions have not been formally evaluated.⁹⁰ We do not know whether they are reliable. Analysts nevertheless use them to produce opinions, but opinions of unknown probative value. Such opinions are often expressed in strong terms such as ‘one and the same’ or ‘a high level of similarity’, without empirical justification.⁹¹ In many cases they are ipse dixit — impressionistic and/or speculative assertions.⁹² The fact that such opinions do not have an empirical predicate must mean that very often there is a nontrivial risk of unfair prejudice.⁹³ In some cases the opinion will be mistaken and in some cases exaggerated opinions will be accepted and, therefore, overvalued by the trier of fact.⁹⁴ That is, mistaken or exaggerated opinions will be taken at face value or insufficiently discounted.

89 See *NAS Report*, above n 6, 85; *NIST Report*, above n 6; Law Commission, *Expert Evidence in Criminal Proceedings in England and Wales* (Law Com No 325, 2009) 5 [1.20], 6 [1.24].

90 *NAS Report*, above n 6, 7–8.

91 See, eg, *R v Tang* (2006) 65 NSWLR 681, and the debate around the uniqueness and expression of latent fingerprint evidence; Simon A Cole, ‘Forensics without Uniqueness, Conclusions Without Individualization: The New Epistemology of Forensic Identification’ (2009) 8 *Law, Probability and Risk* 233; Michael J Saks and Jonathan J Koehler, ‘The Individualization Fallacy in Forensic Science Evidence’ (2008) 61 *Vanderbilt Law Review* 199; cf David H Kaye, ‘Probability, Individualization, and Uniqueness in Forensic Science Evidence: Listening to the Academics’ (2010) 75 *Brooklyn Law Review* 1163, 1176–7.

92 Ipse dixit is bare assertion by an analyst portrayed by the prosecutor (and accepted by the court) as an expert. In such cases the analyst tends to claim some special ability and implicitly asks for the trust (or indulgence) of the judge and the jury. The analyst’s procedure may or may not be explained. We might understand the process (or even the basis) but the interpretation itself is inscrutable. Moreover, understanding a process does not necessarily provide means of assessing it. There is no straightforward means of determining if the procedure or the analyst’s opinion is of any value. Ipse dixit reinforces the need for evidence, independent of the analyst, that supports the claim to expertise. Evidence of relevant expertise, doing precisely the thing that was done in the instant case, is necessary. In the absence of such information we cannot know how probative is the analyst’s opinion and even whether it is relevant.

93 In most cases we cannot be sure if the evidence will be (or was) overvalued. Such uncertainty requires special care in admission of weak, speculative and unreliable expert opinion evidence.

94 In some cases the opinion may actually be under-valued but this is not an issue that concerns *Christie* and s 137.

The fact that a small group of laypeople (for example, a jury) might ‘accept’ incriminating opinions of unknown probative value is not particularly meaningful. It reveals nothing about the probative value of the opinions. In the absence of information about the reliability of a technique (and the scope of its validity) there will always be significant (that is, nontrivial) risks that the trier will overvalue the evidence. There is always a risk that real weaknesses, including unreliability and major uncertainties, will be marginalised or not accepted. These risks are extreme where limitations and oversights are not raised and explained, and persist when attempts are made to address them, particularly where such interventions are not informed or effective.⁹⁵

In practice, those endeavouring to evaluate forensic science and medicine evidence have not been provided with information about the value of techniques. Very rarely are they placed in a position conducive to the rational assessment of incriminating expert opinions. Rather, and in the context of an adversarial proceeding, they tend to be peppered with selective insights, criticisms, claims and counter-claims that, in general, do not enable a judge or jury to rationally evaluate the expert’s incriminating testimony.

C Understanding the Dangers of Contamination and Pseudo-Corroboratorion

The danger that forensic science evidence will be admitted in circumstances that impede the rational assessment of its probative value is often accentuated by the combination of different pieces of evidence. This can operate in a number of different ways. One aspect of this may be a tendency to admit weak expert evidence because it appears to corroborate an otherwise strong Crown case.⁹⁶ A very real but largely unrecognised threat to the rational evaluation of evidence is the suggestion at trial that forensic science and medicine evidence provide *independent corroboration* to other parts of the case when they do not. Historically, there has been insufficient attention directed to the question of whether forensic science and medicine evidence is independent of the other evidence. In many cases, expert evidence is contaminated by the analyst’s gratuitous exposure to domain irrelevant information — that is, information about the case or the accused that is not relevant to the analysis.⁹⁷

95 Uncertainty and limitations should be proactively identified and explained by the prosecutor. See Gary Edmond, ‘(Ad)ministering Justice: Expert Evidence and the Professional Responsibilities of Prosecutors’ (2013) 36 *University of New South Wales Law Journal* 921.

96 Notwithstanding the strength of the Crown case, the probative value of expert evidence should generally be assessed independently of both the strength of the case overall and of other, apparently corroborating, evidence. Generally, the probative value of expert opinion evidence should be informed by the results of formal evaluation of the underlying technique or process.

97 Itiel E Dror, David Charlton and Ailsa E Péron, ‘Contextual Information Renders Experts Vulnerable to Making Erroneous Identifications’ (2006) 156 *Forensic Science International* 74. See generally D Michael Risinger et al, ‘The *Daubert/Kumho* Implications of Observer Effects in Forensic Science: Hidden Problems of Expectation and Suggestion’ (2002) 90 *California Law Review* 1.

Forensic analysts are routinely exposed to domain irrelevant information, such as the suspect's criminal record or the fact that stolen goods were found in the accused's possession. Such information has been shown to have the potential to influence, and indeed change, interpretations by very experienced forensic analysts, including those who know about the potential dangers. Generally, analysts cannot avoid or circumvent contextual biases and unconscious influences by cognitive effort. These influences are particularly insidious, often operating below the level of consciousness.⁹⁸ Significantly, such influences, and evidence derived through processes infected by them, have the potential to completely destroy the probative value of derivative opinion — even opinions obtained using techniques that are otherwise demonstrably reliable. They also have the ability to contaminate other expert and non-expert evidence.⁹⁹

Forensic science and medicine techniques only provide independent corroboration where the analysts using them are blind to other incriminating/prejudicial information (and where the underlying techniques are reliable). Indeed, the major value of forensic science and medicine would seem to lie in the provision of independent and reliable insights or feedback (whether confirmatory or disruptive) on a case, particularly on the suspicions of the investigators. Analysts should not be exposed to gratuitous information, or should be exposed sequentially through a procedure that is documented.¹⁰⁰ Such procedures are rarely used.

Contextual bias and pseudo-corroboration threaten both the probative value and the rational evaluation of a large proportion of forensic science and medicine evidence. They introduce real risks of misuse and misunderstanding that are unlikely to be addressed, let alone corrected, at trial.

IV TRIAL SAFEGUARDS AS A CORRECTIVE TO THE DANGER OF UNFAIR PREJUDICE

In determining whether to admit incriminating expert evidence, it is necessary as part of the s 137 obligation for the trial judge to consider whether any risks associated with misunderstanding and misuse can and will be corrected through the course of proceedings, particularly via judicial directions and warnings (but also cross-examination and rebuttal witnesses).

In the absence of formal evaluation (against correct answers, that is, ground truth), it is not possible to determine how well a technique works.¹⁰¹ In such circumstances we do not know whether the analyst is better than a juror or chance. It might be possible for a trial judge to explain the need for validation studies, the

98 This means that cross-examination may be restricted to methodological abstractions.

99 Gary Edmond et al, 'Contextual Bias and Cross-Contamination in the Forensic Sciences: The Corrosive Implications for Investigations, Plea Bargains, Trials and Appeals' (2014) *Law Probability and Risk* (forthcoming).

100 Dan E Krane et al, 'Sequential Unmasking: A Means of Minimizing Observer Effects in Forensic DNA Interpretation' (2008) 53 *Journal of Forensic Science* 1006.

101 In some cases it may not be possible to determine whether the technique works at all.

significance of error rates and proficiency and even some of the implications of oversights. It might also be possible for a trial judge to explain why experience and the kinds of other heuristics that they and the lawyers have traditionally relied upon — such as formal qualifications, the existence of a field and prior admission — are not particularly salient. However, these issues and limitations are not always exposed during the expert's testimony or the judge's summing up. Moreover, judges are not usually referred to relevant scientific literature and are not in a position to invoke the literature unilaterally.¹⁰² Consequently, judicial warnings and directions rarely refer to the significance of scientific methods (and oversights), they rarely draw upon relevant scientific literature, rarely if ever address the threats posed by contextual bias and pseudo-corroboration, and therefore almost never capture the weakness and vulnerability of untested experience.¹⁰³

Even if judges did explain the main problems, it must remain doubtful whether such explanations would overcome the apparent force of even unreliable expert evidence. The difficulty conveying complex methodological issues, along with the lack of positive (that is, demonstrable) evidence supporting a technique and the proficiency of an analyst using that technique according to standards (or an established protocol), would seem to make trial safeguards, particularly judicial instruction and warnings, ineffective means of exposing, let alone overcoming, deficiencies with incriminating expert evidence. In most cases trial safeguards will not address (and will not explain) the methodological, procedural, logical, psychological and statistical problems with incriminating expert evidence.¹⁰⁴ Even when they do, there are still very real risks that in the context of the trial they will not exert the appropriate influence or that the jury may use other incriminating evidence — whether forensic science or some other evidence — irrationally as a makeweight (even when the expert evidence is not genuinely independent).

Furthermore, generally, cross-examination, rebuttal experts and judicial instructions and directions provide only criticisms and caveats. At best they draw attention to limitations and weaknesses. They do not provide the kinds of information that enable the rational assessment of opinions derived from

102 See *Aytugrul v The Queen* (2012) 247 CLR 17 ('*Aytugrul*'); UEL pt 4.2. In the 2014 Paul Byrne SC Memorial Lecture, Dyson Heydon suggested that a trial judge might nonetheless refer to scientific literature in the attempt to gauge how 'high' evidence could legitimately be taken 'at its highest'. In his defence of the New South Wales approach, Heydon suggested that a trial judge might refer to scientific support for the assessment of the weakness of the evidence. His example referred to the use of scientific research, by experimental psychologists and cognitive scientists, on problems with eyewitness identification evidence. Unfortunately, the example from the talk is not included in the published essay: Dyson Heydon, 'Is the Weight of Evidence Material to its Admissibility?' (Speech delivered at the Paul Byrne SC Memorial Lecture, University of Sydney, 15 October 2014). This approach seems to be inconsistent with the approach adopted by the High Court in *Aytugrul*, where McClellan CJ at CL was criticised for his independent researches in the New South Wales Court of Criminal Appeal (at 199–200 [66]–[67]), and difficult to reconcile with the antipathy toward trial and appellate judges undertaking their own research or taking unilateral notice of anything but uncontroversial scientific findings.

103 Gary Edmond and Mehera San Roque, 'The Cool Crucible: Forensic Science and the Frailty of the Criminal Trial' (2012) 24 *Current Issues in Criminal Justice* 51.

104 The use of defence experts sometimes facilitates or eases the admission of otherwise untested incriminating expert evidence. See, eg, *Honeysett v The Queen* [2013] NSWCCA 135 (5 June 2013).

many techniques in routine use. Issues of training and experience, impartiality, confidence, and methodological scruples may help us to interpret opinions derived from validated techniques. Unfortunately, they tend to provide limited insight into the probative value of techniques (and derivative opinions) that have yet to be formally evaluated.¹⁰⁵

Trial safeguards might have untapped potential for exposing and explaining the problems with expert evidence, including evidence that has not been formally evaluated, to those responsible for its evaluation. But so far it seems that they have been used in ways that generally do not identify or convey the serious problems to those obliged to assess the probative value of forensic science evidence. The exclusion available through s 137 provides one of the few avenues for effectively addressing serious threats to reasoning and proof.

V CONCLUSION: RELIABILITY, RATIONALITY AND RISK

The *Christie* discretion and its modern incarnation in s 137 (and s 135) have not been credibly or consistently applied to expert evidence. Apart from a few cases, perhaps most conspicuous in the early appearance of DNA profiling (for example, *R v Lucas*,¹⁰⁶ *R v Green*¹⁰⁷ and *R v Pantoja*),¹⁰⁸ Australian judges have generally been reluctant to exclude incriminating expert opinions notwithstanding the fact that in many cases techniques are untested and so the information required to make sense of them is not available.¹⁰⁹ Reluctant to trespass on the constitutional prerogatives of the trier of fact, Australian judges have maintained a highly conventional and ideological commitment to the efficacy of the trial and its many ‘safeguards’, along with the capabilities and performance of juries (in less than ideal conditions). Courts have been inclined to treat expert evidence as unexceptional and not especially problematic. In consequence, *Christie* and s 137 have been infrequently applied to exclude expert evidence.

In the absence of independent evidence that a technique works (that is, that it is sufficiently reliable) and that the analyst is proficient in using the reliable technique, should we assume that the evidence is probative? The fact that a jury might *accept* an expert’s untested opinion should not stand in the place of evidence of actual ability. If some technique or purported ability to ‘rationally

105 This might lead to some kind of discounting, but in the absence of information about validity and reliability, such discounting will itself be speculative. Juries might, for example, not discount sufficiently on the basis that the particular witness is highly experienced or performed well during cross-examination.

106 [1992] 2 VR 109.

107 (Unreported, New South Wales Court of Criminal Appeal, 26 March 1993).

108 (1996) 88 A Crim R 554.

109 *R v Lucas* [1992] 2 VR 109 (Hampel J); *R v Green* [1993] NSWCCA 10 (26 March 1993) (Gleeson CJ, Cripps JA and Abadee J); *R v Pantoja* (1996) 88 A Crim R 554 (Hunt CJ at CL, with whom Hidden J agreed). See also Edmond, ‘Specialised Knowledge, the Exclusionary Discretions and Reliability’, above n 7, 48. As already noted in above n 8, recent cases such as *Morgan v The Queen* (2011) 215 A Crim R 33 and *Honeysett* (2014) 88 ALJR 786 have avoided addressing questions about the reliability of ‘techniques’ relied on by the prosecution’s expert.

affect' the assessment of facts in issue is doubtful (or even speculative or unknown) then there are real dangers in leaving it to laypersons to assign weight. That risk is accentuated where the trier is not provided with effective assistance. Risks are also accentuated where the opinion is presented as that of an 'expert' with considerable, though untested, experience. The results of validation and proficiency testing enable the relevance and probative value of opinions to be determined rationally. Ideally the evaluation of new and emerging techniques and technologies should be independently undertaken prior to use in the courtroom.

In circumstances where the probative value of an incriminating opinion is unknown, because appropriate evaluation has not been undertaken, courts should be cautious about admitting incriminating expert testimony. There will always be a nontrivial risk — especially where it is presented as expert and reliance placed on the analyst's experience — that such incriminating opinions will be overvalued or otherwise misused. There is a real risk that criminal proof will be subverted. The liberal admission of speculative and unreliable expert evidence threatens both the goals of the trial (namely truth and fairness), while simultaneously rewarding forensic analysts for disregarding mainstream scientific methods such as validation studies.

As things stand in New South Wales, no section of the *Evidence Act 1995* (NSW) has been interpreted to require prosecutors or trial judges to direct their attention to the reliability of expert evidence. At no stage is there a formal expectation that forensic science techniques, including those in regular use, will have been formally evaluated. Rather, there is a belief that our criminal courts and lay jurors provide an appropriate forum to explore the epistemic value of scientific and medical techniques that have not been formally evaluated. We have serious doubts whether this is always so.

The Victorian Court of Appeal in *Dupas* takes what is, in our view, a more appropriate approach to the role and scope of s 137. It is our contention that this approach, if followed in other courts, has repercussions for the admission of forensic and medical science evidence, encouraging counsel and courts to carefully consider whether the trier of fact is in a position to assess accurately the probative value of such evidence. *Christie* and s 137 are designed to prevent the misuse of evidence and the consequent risk of undermining the high standard of criminal proof. Considering that the likelihood of misuse of forensic science and medical evidence is high, and that current interpretations of s 79 do not engage with reliability, the lead in *Dupas* should be followed. Counsel and courts should be encouraged in every case to consider whether otherwise relevant and admissible forensic science evidence should be excluded under s 137 on the basis that it is not capable of rational consideration by the trier of fact.

Our emphasis has been on the traditional purpose of *Christie* and its incarnation in s 137 of the UEL. We have insisted on the need to direct attention to the probative value and reliability of incriminating expert evidence as part of the goal of preventing unfair prejudice to the defendant and the subversion of criminal proof. This approach is critical of other approaches, particularly those commonly in use in New South Wales. However, regardless of the approach to *Christie* and

s 137, when it comes to expert evidence, it is not possible to ignore probative value and therefore reliability. Even attempts to circumvent attention to actual probative value by, for example, taking evidence ‘at its highest’, do not overcome the need to attend to the actual probative value of expert evidence when attempting to assess the danger of unfair prejudice to the defendant. Many of the serious risks from unfair prejudice associated with expert evidence will emerge from jurors not understanding, misusing or overvaluing the evidence. In many cases, these risks will be acute where the evidence is less probative (that is, less reliable) than the way it is expressed in reports and testimony, or understood by the trier of fact. These kinds of dangers can only be ascertained and guarded against through an appreciation of the actual probative value of the evidence.¹¹⁰

110 It may be that the danger of unfair prejudice, particularly with respect to expert evidence, is not necessarily (or always) incommensurable with the probative value of the evidence.