

SPECIALISED KNOWLEDGE, THE EXCLUSIONARY DISCRETIONS AND RELIABILITY: REASSESSING INCRIMINATING EXPERT OPINION EVIDENCE

GARY EDMOND*

I INTRODUCTION

This essay documents the failure of the jurisprudence governing the reception of expert opinion evidence in New South Wales.¹ Focused on the admissibility of expert opinion evidence and the operation of the exclusionary discretions, the essay discusses the circumstances in which a trial judge might prevent expert evidence from going before the tribunal of fact. In particular, it examines the apparent reluctance to exclude unreliable expert opinion evidence and expert opinion evidence of unknown reliability adduced by the state.

As things stand, most Australian judges have not exhibited much interest in the reliability of expert opinion evidence. This disinterest, in a system based on truth and justice, and supposedly operating within a rational tradition of evidence and proof, might be considered intriguing, at the very least.² Judicial disinterest in the reliability of expert opinion evidence has meant that the state has been able to secure the admission of incriminating expert opinion evidence of unknown reliability, notwithstanding the ability to ascertain validity and the existence of statutory rules designed to regulate the admissibility of expert opinion evidence.

Once expert opinion evidence adduced by the state is deemed admissible, a judge might still prevent this evidence going before a criminal jury through recourse to the exclusionary discretions. From the early twentieth century, and the decision in *R v Christie*,³ the modern exclusionary discretions empowered trial judges to exclude otherwise admissible evidence in circumstances where the probative value of the evidence was outweighed by the danger of unfair prejudice

* BA(Hons) University of Wollongong, LLB(Hons) University of Sydney, PhD University of Cambridge. Associate Professor, Faculty of Law, The University of New South Wales, Sydney 2052, Australia. I would like to thank Richard Kemp, Mehera San Roque, Kristy Martire, Andrew Ligertwood, Stephen Odgers, Jill Hunter, Maciej Henneberg, Mike Lynch and several referees for comments and suggestions. The research forms part of a study of expert evidence supported by the Australian Research Council (DP0771770).

1 Although this essay is primarily focused on NSW, and by implication the Federal Courts, Tasmania and the ACT, much of the discussion has obvious implications for jurisdictions throughout Australia and beyond.

2 William Twining, 'The Rationalist Tradition of Evidence Scholarship' in *Rethinking Evidence: Exploratory Essays* (1990) 32; Erica Beecher-Monas, *Evaluating Scientific Evidence: An Interdisciplinary Framework for Intellectual Due Process* (2006).

3 [1914] AC 545.

to the accused. In theory, this invested trial judges (and appellate courts) with means of excluding evidence where there was a real risk that its admission would unfairly prejudice the accused. In practice, this is not how the discretions have been applied to expert opinion evidence. Instead, they have been interpreted such that trial and appellate judges have become reluctant to make their own assessment of probative value. Judicial indifference to probative value has been shaped by the primacy attributed to fact-finding by juries, in conjunction with the availability of procedural safeguards such as cross-examination, defence experts and judicial directions.

These developments, as we shall see, have effectively eviscerated the exclusionary potential of the admissibility rules and discretions with respect to expert opinion evidence. That is, the discretionary exclusions, ostensibly concerned with assuring a fair trial and fairness to the accused, have almost no role to play in relation to expert opinion evidence adduced by the state. In response, this essay contends that judges should actually determine the probative value of the state's incriminating expert opinion evidence when applying the exclusionary discretions (ss 135 and 137) derived from *Christie*. It also contends that judges should incorporate *demonstrable reliability* into the admissibility standard for expert opinion evidence (s 79) adduced by the state. This account begins, therefore, with a review of the rules governing the admissibility of expert opinion evidence under the *Evidence Act 1995* (NSW), and an explanation of why the prevailing approach has not prevented unreliable evidence and evidence of unknown reliability from contaminating criminal prosecutions.

II THE ADMISSIBILITY OF EXPERT OPINION EVIDENCE

Because this essay is also focused on the exclusionary discretions, this excursion into the admissibility of expert opinion evidence will, of necessity, be concise. The purpose of this section is to explain why the existing rules of admissibility have not prevented unreliable expert evidence gaining access to criminal trials. As we shall see, the main reason is that judges have expressed disinterest in the reliability of expert opinion evidence and have not assiduously applied the terms of the *Evidence Act* to expert opinion evidence adduced by the prosecution.

The admissibility of opinion evidence is governed by Part 3.3 of the *Evidence Act*. This Part is dominated by the *opinion rule* (s 76) which states that 'evidence of an opinion' is not admissible 'to prove the existence of a fact about the existence of which the opinion was expressed'. There are, however, several exceptions to the exclusionary impact of section 76. Although it does not attempt to codify the common law, section 79 provides the major exception for expert opinion evidence.⁴ It reads:

4 At common law, judges were primarily interested in the existence of a 'field of knowledge' and whether the witness was an 'expert' in that field. See *Bugg v Day* (1949) 79 CLR 442; *Clark v Ryan* (1960) 103 CLR 486; *Commissioner for Government Transport v Adamcik* (1961) 106 CLR 292; *Ramsay v Watson* (1961) 108 CLR 642.

79 Exception: opinions based on specialised knowledge

If a person has specialised knowledge based on the person's training, study or experience, the opinion rule does not apply to evidence of an opinion of that person that is wholly or substantially based on that knowledge.

'An opinion' *based* 'wholly or substantially' on 'specialised knowledge' which is *based* on 'training, study or experience' is not caught by the exclusionary opinion rule.⁵ We can represent this schematically, in the following way:

training, study or experience → specialised knowledge → opinion

Provided these conditions are met, a witness can give relevant opinions subject only to the exclusionary discretions (considered below) and the overarching commitment to a fair trial.⁶

Two of the leading admissibility decisions under the *Evidence Act 1995* (NSW) are *Makita (Australia) Pty Ltd v Sprowles* from the Court of Appeal and *R v Tang*⁷ from the Court of Criminal Appeal ('CCA'). These two decisions help us to understand how section 79 has been interpreted and applied. *Makita* was an appeal relating to an injury on a set of concrete stairs.⁸ Sprowles successfully sued her employer in negligence after a slip and fall at work. For the trial, she called a physics professor who testified that in his opinion, the stairs were too slippery. Sprowles was awarded more than A\$1 million dollars in damages. Makita, the employer and occupier, appealed that verdict. The appeal challenged Sprowles' credibility, particularly her account of the shoes she was wearing at the time of the accident and her description of their subsequent use. The main issue for our purposes, though, is the expert opinion evidence of Associate Professor Morton of the University of New South Wales.

Almost a decade after the accident Associate Professor Morton examined the stairs and the shoes and measured their respective coefficients of friction (ie, slipperiness). Most of the test results appeared to comply with the relevant Australian standard.⁹ Yet, as Associate Professor Morton testified, he believed that a higher standard would have been more appropriate. For Morton, the fact that not all of the test results satisfied even the *low* Australian standard, in conjunction with the failure to roughen the surface of the concrete or fit abrasive strips and edging, made the stairs unnecessarily slippery and was evidence of negligence.

The Court of Appeal recognised Associate Professor Morton's formal qualifications and found him to be a credible witness. Nevertheless, they considered that his reasoning, particularly the interpretation of test results, was

5 See, eg, *HG v The Queen* (1999) 197 CLR 414, [39] (Gleeson CJ) ('HG').

6 *R v Swaffield*; *Pavic v The Queen* (1998) 192 CLR 159, 193 (Toohey, Gaudron and Gummow JJ).

7 *R v Tang* (2006) 65 NSWLR 681, [134] ('Tang').

8 *Makita (Australia) Pty Ltd v Sprowles* (2001) 52 NSWLR 705 ('Makita'). See also *Idoport Pty Ltd v National Australia Bank Ltd* [2001] NSWSC 123 (Einstein J).

9 The Australian standard had not been instituted at the time of the accident.

not always clear. In determining how to respond to this expert opinion evidence, Heydon JA (with Priestly JA agreeing) explained the approach:

In short, if evidence tendered as expert opinion evidence is to be admissible, it must be agreed or demonstrated that there is a field of 'specialised knowledge'; there must be an identified aspect of that field in which the witness demonstrates that by reason of specified training, study or experience, the witness has become an expert; the opinion proffered must be 'wholly or substantially based on the witness's expert knowledge'; so far as the opinion is based on facts 'observed' by the expert, they must be identified and admissibly proved by the expert, and so far as the opinion is based on 'assumed' or 'accepted' facts, they must be identified and proved in some other way; it must be established that the facts on which the opinion is based form a proper foundation for it; and the opinion of an expert requires demonstration or examination of the scientific or other intellectual basis of the conclusions reached: that is, the expert's evidence must explain how the field of 'specialised knowledge' in which the witness is expert by reason of 'training, study or experience', and on which the opinion is 'wholly or substantially based', applies to the facts assumed or observed so as to produce the opinion propounded. If all these matters are not made explicit, it is not possible to be sure whether the opinion is based wholly or substantially on the expert's specialised knowledge. If the court cannot be sure of that, the evidence is strictly speaking not admissible, and, so far as it is admissible, of diminished weight.¹⁰

Justice of Appeal Heydon cited a tremendous volume of predominantly common law authority in support of this interpretation of section 79. In doing so, his Honour characterised the 'prime duty of experts in giving opinion evidence' as the need to '*furnish the trier of fact with criteria enabling evaluation of the validity of the expert's conclusions*'.¹¹

Notwithstanding Associate Professor Morton's expert opinion evidence the judges on the Court of Appeal were persuaded by the 'incident-free' history of the stairs. Sprowles' co-workers testified that the stairs had not presented problems. Even Sprowles conceded that during her years at the site, she had not found the stairs to be slippery. Justices of Appeal Heydon and Powell separately concluded that in the absence of clearer explanations about 'the *validity* of Professor Morton's approach' the apparent safety of the stairs was a very powerful piece of evidence to overcome.¹²

The conclusions in Professor Morton's report ought not to be accepted uncritically. On examination it is difficult to be convinced by them. The lay history of incident-free use of the stairs suggests that they were not slippery.¹³ That inference from that history is preferable to Professor Morton's conclusions.

The *Makita* jurisprudence has been prominent in NSW and was influential on the Court of Criminal Appeal in *Tang*. *Tang* was an appeal concerned with the

10 *Makita* (2001) 52 NSWLR 705, [85]. See also *HG* (1999) 197 CLR 414, [39]–[41] (Gleeson CJ). In subsequent cases, particularly in the Federal Court, the emphasis on the need to identify admissible evidence as the basis for any inferences has been characterised as a 'counsel of perfection'. The Full Federal Court has been more inclined to let these issues go to weight. See, eg, *Sydneywide Distributors Pty Ltd v Red Bull Australia Pty Ltd* [2002] FCAFC 157, [7] (Branson J); Lee Aitken, 'Expert Evidence and *Makita* – "Gold standard" or Counsel of Perfection?' (2006) 28 *Australian Bar Review* 207.

11 *Makita* (2001) 52 NSWLR 705, [59] (emphasis added). Notably, this is not the 'overriding' or 'paramount duty' of an expert witness set out in sch 7 of the Expert Witness Code of Conduct in the *Uniform Civil Procedural Rules 2005* (NSW).

12 *Ibid* [99] (Heydon JA) (emphasis added).

13 *Ibid* [102] (Heydon JA). See also [21] (Powell JA).

admissibility of *facial mapping* and *body mapping* evidence. ‘Facial mapping’ involves the identification of a person of interest using anthropometric and/or morphological analysis of the face. It usually entails, respectively, quantitative and/or qualitative comparisons of security and CCTV images of an unknown person with images of a known person. ‘Body mapping’ involves a similar set of processes focused on the body, posture and movement. On the basis of her facial mapping and body mapping techniques, Dr Sutisno, an anatomist called by the prosecution, opined that the person of interest in security images from a robbery was Hien Puoc Tang. From her comparison of the security images – which were of such poor quality that Spigelman CJ indicated that they ‘could not be left for the jury’¹⁴ – and a set of high quality police images of Tang, Dr Sutisno was convinced that the two persons were the same and testified to that effect. This evidence was admitted over objection and the admissibility of Dr Sutisno’s identification evidence became the principal ground of appeal.

In reviewing the admissibility of this evidence, Spigelman CJ (with Simpson and Adams JJ agreeing) directed his attention to section 79, and explained its operation:

Section 79 has two limbs. Under the first limb, it is necessary to identify ‘specialised knowledge’, derived from one of the three matters identified, ie, ‘training, study or experience’. Under the second limb, it is necessary that the opinion be ‘wholly or substantially based on that knowledge’. Accordingly, it is a requirement of admissibility that the opinion be demonstrated to be based on the specialised knowledge.¹⁵

Applying this approach to the evidence, the Court concluded that facial mapping was not ‘specialised knowledge’ that would enable Dr Sutisno to give her opinion about the identity of the unknown person in the security images.

Facial mapping, let alone body mapping, was not shown, on the evidence in the trial, to constitute ‘specialised knowledge’ of a character which can support an opinion of identity.¹⁶

Dr Sutisno’s opinions about the identity of Tang were not based on ‘specialised knowledge’.¹⁷ Instead, her opinions – including the emphasis on what were described as ‘unique identifiers’ – were characterised by the Court, somewhat pejoratively, as *ipse dixit*.¹⁸ Chief Justice Spigelman reproduced the passage from *Makita* at [85], extracted above. Like Heydon JA before him, he was concerned that the reasoning process – this time the process employed by Dr Sutisno – was inadequately explained. Facial mapping and body mapping were, therefore, incapable of supporting opinions about identity and should not have been admitted.

14 *Tang* (2006) 65 NSWLR 681, [120].

15 *Tang* (2006) 65 NSWLR 681, [134].

16 *Ibid* [146].

17 *Ibid* [140]–[141].

18 *Ibid* [154]. According to the Oxford English Dictionary, *ipse dixit* is a ‘dogmatic statement resting merely on the speaker’s authority’. The Latin translates as ‘he himself said it’. In the context of jurisprudence associated with expert evidence the term has a long and pejorative pedigree. A useful illustration is the US Supreme Court’s decision in *General Electric Co v Joiner*, 522 US 136, 146 (1997).

Things did not end there, however. Dr Sutisno's training in anatomy, combined with the fact that she had repeatedly compared the security images with the police photographs, led Spigelman CJ to qualify her as an ad hoc expert. This common law exception to the general prohibition on opinion evidence was used to enable Dr Sutisno to testify about similarities and (at least in theory) differences between the person(s) in the images. In consequence, Dr Sutisno would be allowed to make de facto identifications incriminating Tang. Even though Dr Sutisno was not giving opinion evidence based on 'specialised knowledge', the Court was willing to allow her to testify about similarities between the two sets of photographs in a future trial. That is, she would be allowed to give expert opinion evidence about similarities, but would be prevented from actually identifying the accused as she had done, in very confident terms, during the first trial. According to Spigelman CJ, any weaknesses or limitations with Dr Sutisno's techniques and opinions were for cross-examination. It would be for a future jury to determine the reliability and weight of her evidence. Interestingly, and perhaps revealingly, the exclusionary discretions played no part in the Court's decision.¹⁹

Perhaps the most intriguing feature of *Tang* is the attitude expressed by the Court toward the *reliability* of Dr Sutisno's opinion evidence. Adopting what might be considered a very narrow approach to the text of section 79, Spigelman CJ explained that the 'focus of attention must be on the words "specialised knowledge", not on the introduction of an extraneous idea such as "reliability"'.²⁰

Drawing on the influential decision in *Daubert v Merrell Dow Pharmaceuticals Inc*, where the Supreme Court of the United States explained its approach to the *Federal Rules of Evidence* (1975) governing the admissibility of opinions derived from 'scientific, technical and other specialized knowledge', Spigelman CJ offered insight into the meaning of 'specialised knowledge'.²¹ Quoting directly from *Daubert*, he accepted that

'knowledge' connotes more than subjective belief or unsupported speculation. The term 'applies to any body of known facts or to any body of ideas inferred from such facts on good grounds'.²²

There is, it might be thought, a serious tension between these two passages. That is, between disinterest in 'an extraneous idea such as "reliability"' and a focus on 'specialised knowledge' that involves 'known facts', inferences from such facts on 'good grounds', and requires more than 'subjective belief' or 'unsupported speculation'.

Nevertheless, in developing the admissibility jurisprudence governing incriminating expert opinion evidence, the most senior judge in NSW explicitly

19 To its credit, the prosecution did not adduce facial mapping evidence at the subsequent re-trial.

20 *Tang* (2006) 65 NSWLR 681, [137].

21 *Daubert v Merrell Dow Pharmaceuticals, Inc*, 509 US 579, 590 (1993) ('*Daubert*'). For some discussion of the US jurisprudence, see Gary Edmond, 'Supersizing *Daubert*: Science for Litigation and its Implications for Legal Practice and Scientific Research' (2007) 52 *Villanova Law Review* 857.

22 *Tang* (2006) 65 NSWLR 681, [138] (emphasis added). In *Daubert*, Blackmun J took this passage from *Webster's Third New International Dictionary* (1986) 1252. See Scott Brewer, 'Scientific Expert testimony and Intellectual Due Process' in Evan Selinger and Robert Crease (eds), *The Philosophy of Expertise* (2006) 111.

dismissed the need for reliability. Instead, emphasis was placed on the more amorphous idea of ‘specialised knowledge’ (in conjunction with the *bases*). If we reflect on the application of these ideas in *Tang*, we can see how easily they can be circumvented. Dr Sutisno’s opinions were admitted even though there was no ‘specialised knowledge’ and no explanation of how her anatomical training (or study of the images) would provide a basis for drawing inferences about identity that were not merely ‘speculative’ or ‘subjective’.

In practice, section 79 is not always applied particularly strictly – at least not to the evidence adduced by the prosecution in criminal proceedings.²³ Rather than focusing on whether opinions are based ‘wholly or substantially’ on ‘specialised knowledge’ that is based on an individual’s ‘training, study or experience’, in the criminal sphere judges have a tendency to privilege *formal* training and *recognisable* expertise. Sometimes formal training, in established fields like medicine, anatomy or biology, enables an *expert* to testify about matters that are not based on ‘knowledge’ (let alone ‘specialised knowledge’), and not based on their actual training, study or experience. There is a tendency to allow trained professionals to testify in areas beyond their actual expertise or beyond the collective ability of any recognisable field or identifiable sub-discipline. There can be, as a subsequent example (see Part IV) will help to illustrate, considerable latitude between an expert’s ‘training, study or experience’ and the bases and knowledge purportedly grounding what becomes admissible *expert* opinion evidence.

These developments might be considered unfortunate. Chief Justice Gleeson certainly thought so when he cautioned that:

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.²⁴

The disinterest in reliability and superficial supervision of the links between opinions, ‘specialised knowledge’ and ‘training, study or experience’ are particularly troubling when we appreciate that there are ways to determine the validity of most forensic scientific techniques.

The opinion rule and its exceptions have not prevented incriminating expert opinion evidence with questionable epistemological provenance from going before criminal juries.

III CHRISTIE AND THE – NOT SO EXCLUSIONARY – STATUTORY DISCRETIONS

Given the failure of Part 3.3 to exclude unreliable expert opinions and expert opinions of unknown reliability adduced by the state, it might be thought that

23 Although beyond the scope of this essay, it is worth noting that in recent years much of the expert evidence jurisprudence (particularly in NSW) has emerged from civil litigation and prosecutions by the Australian Securities and Investments Commission.

24 *HG* (1999) 197 CLR 414, [44]. *HG* was an appeal from the NSWCCA and this passage was cited, with approval, by Heydon JA in *Makita* (2001) 52 NSWLR 705, [84].

judicial discretions based on balancing the probative value of evidence against the danger of unfair prejudice to the accused would afford protection against unreliable expert evidence. That, as we shall now see, is not how things have unfolded.

A Common Law Origins

At common law there has long been a discretion, vested in the trial judge, to exclude otherwise admissible evidence where it might operate unfairly against the accused. The eponymous discretion emerged from the appeal to the House of Lords in *R v Christie*. The extract below is taken from the judgment of Lord Moulton:

The law is so much on its guard against the accused being prejudiced by evidence which, though admissible, would probably have a prejudicial influence on the minds of the jury which would be out of proportion to its true evidential value, that there has grown up a practice of a very salutary nature, under which the judge intimates to the counsel for the prosecution that he should not press for the admission of evidence which would be open to this objection, and such an intimation from the tribunal trying the case is usually sufficient to prevent the evidence being pressed in all cases where the scruples of the tribunal in this respect are reasonable. Under the influence of this practice, which is based on an anxiety to secure for every one a fair trial, there has grown up a custom of not admitting certain kinds of evidence which is so constantly followed that it almost amounts to a rule of procedure.²⁵

In *Christie*, the Law Lords transformed prevailing practice, which had been based on judicial persuasion, into a discretionary rule of exclusion. Significantly, the formalisation of this discretion was ‘based on an anxiety to secure for everyone a fair trial’.²⁶ Appealing to the ‘best traditions of our criminal procedure’, Lord Moulton explained that the discretion (and before that, the ‘influence’) was designed to prevent ‘evidence being given in cases where it would have very little or no evidential value’.²⁷ The Law Lords were concerned that ‘the evidential value’ might have ‘the effect on the minds of the jury ... [that] might seriously prejudice the fairness of his trial’.²⁸

In subsequent decades this exclusionary possibility became known as the *Christie* discretion or the rule from *Christie*. Adopted throughout the Commonwealth, the *Christie* discretion allowed trial judges to weigh the probative value of the evidence adduced by the Crown against any unfair prejudice to the accused and exclude admissible evidence in circumstances where that evidence might unfairly disadvantage a criminal defendant. The discretion was gradually consolidated through appeals in *Noor Mohamed v R*, *Harris v*

25 *R v Christie* [1914] AC 545, 559 (Moulton LJ) (emphasis added) (*‘Christie’*). The year before, in *R v Fletcher* (1913) 9 Cr App R 53, 56, Bankes J explained that a judge could suggest ‘to the prosecution that they should not press it, but he cannot exclude evidence which he holds to be admissible’. The Earl of Halsbury LC embraced that position during the argument in *Christie* (1914) 10 Cr App Rep 141, 149.

26 John Heydon, *Cross on Evidence* (7th ed, 2004) § 21220.

27 The issue before the Lords was whether Christie’s response to a serious accusation was admissible evidence.

28 *Christie* [1914] AC 545, 559–60. This approach was adopted in NSW in *R v Eyles* (1917) 17 SR (NSW) 377. See Rosemary Pattenden, *The Judge, Discretion and the Criminal Trial* (1982) 66–7.

Director of Public Prosecutions and *Kuruma v The Queen*.²⁹ Perhaps the most authoritative of these post-war decisions was delivered by the House of Lords in *R v Sang*. There, Lord Diplock explained the discretion in the following terms:

So I would hold that there has now developed a general rule of practice whereby in a trial by jury the judge has a discretion to exclude evidence which, though technically admissible, would probably have a prejudicial influence on the minds of the jury, which would be out of proportion to its true evidential value.³⁰

These decisions were generally embraced by Australian courts, although, as the ensuing extract from *Driscoll v The Queen* indicates, Australian judges exhibited a tendency to take a slightly more restrained approach than their English counterparts:

It has long been established that the judge presiding at a criminal trial has a discretion to exclude evidence if the strict rules of admissibility would operate unfairly against the accused. The exercise of this discretion is particularly called for if the evidence has *little or no weight*, but may be *gravely prejudicial* to the accused: see, e.g., *R. v. Christie*; *Noor Mohamed v. The King*; *Harris v. Director of Public Prosecutions*; and *Kuruma v. The Queen*.³¹

The precise breadth of the discretion seems to have exhibited some variation over time and between jurisdictions.

Often, the repeated application of the *Christie* discretion in relation to particularly troubling types of evidence – such as identification evidence, the evidence of prison informers and accomplices, similar fact and propensity evidence, confessions, and improperly obtained evidence – led to substantial revision of the common law and occasionally statutory reform. Reforms to the rules relating to confessions (‘admissions’ under the *Evidence Act*) provide a particularly good example of how judicial concerns with probative value and unfair prejudice, particularly the danger of police verballing, changed both police procedures and the substantial law relating to their admissibility.³²

Writing just before the enactment of the *Evidence Act 1995* (NSW), Pattenden summarised the *Christie* discretion in the following terms:

29 *Noor Mohamed v R* [1949] AC 182; *Harris v Director of Public Prosecutions* [1952] AC 694; *Kuruma v The Queen* [1955] AC 197. See David Ormerod and Diane Birch, ‘The Evolution of the Discretionary Exclusion of Evidence’ [2004] *Criminal Law Review* 767.

30 *R v Sang* [1980] AC 402, 434.

31 *Driscoll v The Queen* (1977) 137 CLR 517, 541 (Gibbs J) (emphasis added, citations omitted). See also *R v Sandford* (1994) 72 A Crim R 160, 178 (Hunt CJ at CL) (‘*Sandford*’); *R v Fletcher* (1953) 53 SR (NSW) 70, 76; *Bunning v Cross* (1978) 141 CLR 54; *Alexander v R* (1981) 145 CLR 395; *Cleland v R* (1982) 151 CLR 1; *Harriman v R* (1989) 167 CLR 590; *Doney v R* (1990) 171 CLR 207.

32 See, eg, *McDermott v R* (1948) 76 CLR 501; *Van der Meer v R* (1988) 82 ALR 10; *Duke v The Queen* (1989) 180 CLR 508; *McKinney v R* (1991) 171 CLR 468; *Pollard v The Queen* (1992) 176 CLR 177; *Foster v The Queen* (1993) 113 ALR 1. See The changes introduced through Parts 3.4 and 3.11 of the *Evidence Act 1995* (NSW), especially sections 90 and 138.

The common law discretion requires the trial judge to balance the prejudicial effect of evidence against its probative value. Evidence is prejudicial and hence susceptible to discretionary exclusion if there is a real risk that it will contribute to an *erroneous* verdict, either *because its weight and credibility cannot be effectively tested by the defence* or because it may be *misused* by the jury. *Misuse* of evidence covers *inter alia* putting more weight on evidence than it deserves or drawing false inferences from evidence or use of evidence admitted for one purpose for some other *forbidden* purpose. Evidence which is prejudicial only in the sense that it incriminates the accused is not prejudicial for the purposes of the discretion.³³

The common law discretion emerged from a perceived need to secure a fair trial for the accused. Though, as Pattenden's summary indicates, the rationale for the discretion and its scope gradually expanded. By the early 1990s the *Christie* discretion was used to protect the accused, make sure that the defence could test the weight and credibility of incriminating evidence, prevent the misuse of evidence, as well as help to avoid 'erroneous' verdicts.

B Section 137: Excluding Relevant and Admissible Inculpatory Evidence

When the *Evidence Act 1995* (NSW) came into effect, the common law *Christie* discretion was replaced by sections 135 and 137. Here, the main focus is on section 137. Reproduced below, section 137 is concerned with unfairly prejudicial evidence adduced by the prosecution in criminal proceedings.

137 Exclusion of prejudicial evidence in criminal proceedings

In a criminal proceeding, the court must refuse to admit evidence adduced by the prosecutor if its probative value is outweighed by the danger of unfair prejudice to the defendant.

The formulation leaves little doubt about the common law origins.

Section 137 requires the trial judge to exclude 'evidence adduced by the prosecutor' where 'probative value is outweighed by the danger or unfair prejudice'. Unlike the common law discretion which *enabled* Australian judges to exclude otherwise admissible evidence where the probative value was low and the risk of prejudice grave, section 137 *requires* the judges of NSW to exclude admissible evidence if its probative value is *merely* outweighed by the danger of unfair prejudice:

The onus remains on the accused under s 137 to persuade the trial judge that the danger of unfair prejudice from the evidence outweighs its probative value. Once the judge is persuaded of that fact, there is no further discretion involved and the evidence *must* be excluded.³⁴

33 Rosemary Pattenden, *Judicial Discretion and Criminal Litigation* (1990) 233 (emphasis added). See also Rupert Cross, 'Discretion and the Law of Evidence: When it Comes to the "Forensic Crunch"' (1979) 30 *Northern Ireland Legal Quarterly* 289; Mark Weinberg, 'The Judicial Discretion to Exclude Relevant Evidence' (1975) 21 *McGill Law Journal* 1; Bernard Livesey, 'Judicial Discretion to Exclude Prejudicial Evidence' (1968) 26 *Cambridge Law Journal* 291.

34 *R v Lock* (1997) 91 A Crim R 356, 364 (Hunt CJ at common law) ('*Lock*') (emphasis in original). See also *R v Polkinghorne* [1999] NSWSC 704, [51] (Levine J); *R v Blick* [2000] NSWCCA 61, [13], [20] (Sheller JA) ('*Blick*').

1 'Probative value' and 'Relevance'

To understand the way section 137 (and section 135, which is expressed in analogous terms and considered below) has been interpreted we need to take a step backwards to appreciate how it relates to the most fundamental admissibility criterion – *relevance*. Consider the following rules and definitions from the *Evidence Act*.

55 Relevant evidence

- (1) The evidence that is relevant in a proceeding is evidence that, if it were accepted, could rationally affect (directly or indirectly) the assessment of the probability of the existence of a fact in issue in the proceeding.
- (2) In particular, evidence is not taken to be irrelevant only because it relates only to:
 - (a) the credibility of a witness; or
 - (b) the admissibility of other evidence; or
 - (c) a failure to adduce evidence.

56 Relevant evidence to be admissible

- (1) Except as otherwise provided by this Act, evidence that is relevant in a proceeding is admissible in the proceeding.
- (2) Evidence that is not relevant in the proceeding is not admissible.

Evidence Act Dictionary

probative value of evidence means the extent to which the evidence could rationally affect the assessment of the probability of the existence of a fact in issue.

The intention behind the introduction of the *Evidence Act*, as section 56 makes clear, was to encourage a more straightforward and inclusive approach to admissibility.³⁵ That intention was predicated upon a clear commitment to a rational system of evidence and proof. Accordingly, under the *Evidence Act*, evidence that is not relevant is not admissible. Evidence that is relevant is admissible, subject to a series of exclusionary rules, exceptions and 'discretions'. In order to satisfy the relevance requirement, evidence must possess some 'probative value'. This means that the evidence needs to be capable of *rationally affecting* the 'assessment of the probability of the existence of a fact in issue'. 'The key to this definition ... is the word "rationally". Probative value is about

35 The intention was to replace 'legal relevance' with 'logical relevance'. The common law took a pragmatic, but not always principled, disdain to evidence of low probative value. Legally relevant evidence is a subset of logically relevant evidence. In recent years, several appellate judges have encountered difficulty maintaining this distinction. The majority decision in *Smith v R* (2001) 206 CLR 650 (Gleeson CJ, Gaudron, Gummow and Hayne JJ) is a good example of the persistence of *legal* notions of relevance.

the weight which the tribunal of fact, if acting *rationaly*, could give the evidence.³⁶

The terms of section 55 allow for the fact that not all of the evidence presented during the trial will be *accepted* by the tribunal of fact, and that *acceptance* is not the criterion for admissibility. Rather, all that is needed to satisfy the very low threshold for relevance required by the *Evidence Act* is that the evidence ‘could’ (ie, might), ‘if it were accepted’, ‘rationally affect’ the ‘assessment of the probability of the existence of a fact in issue’. If evidence has the potential to alter our assessment of a fact in issue then it is relevant. Relevant evidence is admissible subject to admissibility rules (eg, sections 76 and 79) and the ‘discretions’ (ss 135 and 137).

2 ‘Unfair prejudice to the defendant’

As its terms and genealogy suggest, section 137 is intended to protect the defendant from *unfair prejudice*. Unlike probative value, ‘unfair prejudice’ is not defined in the *Evidence Act* Dictionary. There is, however, a considerable body of case law on the meaning of the term, much of which predates the *Evidence Act*.

As many judges have recognised, highly probative inculpatory evidence is prejudicial to the accused. Section 137 is not concerned with this kind of prejudice, but rather with prejudice that is somehow unfair or productive of unfairness.³⁷ The meaning of ‘unfair prejudice’ was explained in an early and influential Australian Law Reform Commission (‘ALRC’) Report:

By risk of unfair prejudice is meant the danger than the fact finder may use the evidence to make a decision on an improper, perhaps emotional basis, ie on a basis logically unconnected with the issues in the case. Thus the evidence that appeals to the fact-finder’s sympathies, arouses a sense of horror, provokes an instinct to punish, or triggers some mainsprings of human action may cause the fact-finder to base his decisions on something other than the established proposition of the case. Similarly, on hearing the evidence the fact-finder would be satisfied with a lower degree of probability than would otherwise be required.³⁸

36 Jeremy Gans and Andrew Palmer, *Australian Principles of Evidence* (2nd ed, 2004) 145. See also *Lock* (1997) 91 A Crim R 356, 360 (Hunt CJ at common law).

37 ALRC, *Evidence (Interim)*, Report No 26 (1985) [957]. Other examples include *Papakosmas v The Queen* (1999) 196 CLR 297, [91]–[94] (McHugh J) (‘*Papakosmas*’); *R v Clark* [2001] NSWCCA 494, [164] (Heydon JA, Bell J concurring and Dowd J concurring in part) (‘*Clark*’); *R v GK* (2001) 53 NSWLR 317, [30] (Mason P, Dowd J concurring, and (Sully J); *R v Lisoff* [1999] NSWCCA 364, [52] (Spigelman CJ, Newman and Sully JJ) (‘*Lisoff*’); *R v Toki (No 3)* (2000) 116 A Crim R 536, 548 (Howie J); *R v Ambrosoli* (2002) 55 NSWLR 603, [12], [70] (Mason P, Hulme and Simpson JJ concurring); *R v Chai* [2002] NSWCCA 512, [43] (Mason P, Sperling and Bergin JJ).

38 ALRC, above n 37, [644]. This report formed part of the research background to the drafting of the NSW and Commonwealth *Evidence Acts*. The passage was recently endorsed in ALRC, *Uniform Evidence Law*, Report No 102 (2005) [16.24].

‘Unfair prejudice’ refers to the danger that the jury will misuse the evidence, particularly by treating the evidence in an irrational manner, or assigning greater weight to the evidence than it can rationally sustain.³⁹

It is important to emphasise that many judges have expressed anxiety about the jury misusing evidence or giving evidence undue weight. The following extract from *R v Yates* is, in this way, exemplary:

Prejudice argues for exclusion only if there is a *real risk* of danger of it being unfair ... This may arise in a variety of ways, a typical example being, where it may lead a jury to adopt an illegitimate form of reasoning, or to give the evidence undue weight.⁴⁰

Misuse of evidence would seem to be a practical risk with expert opinion evidence, particularly in circumstances where the reliability and probative value of the evidence are unknown.

3 *The Degree of ‘Danger’*

This brings us to ‘*the danger* of unfair prejudice’. Appellate judges have explained that the risk of misuse, overvaluing or irrationality must not be remote or fanciful. Obviously, the courts are not concerned with all risks or the mere possibility of danger. Rather, any ‘danger’ has to be founded or tangible:

In my view evidence may be unfairly prejudicial to a party if there is a *real risk* that the evidence will be misused by the jury in some unfair way.⁴¹

The risk to which section 137 is directed must be ‘more than a hypothetical risk’; it ‘must be a *real one*’.⁴² There ‘must be a real risk that the evidence will be misused by the jury in some way and that that risk will exist notwithstanding the proper directions’.⁴³

4 *‘Outweighed’: The Balancing Exercise*

Section 137 is, as we have seen, mandatory. There is no discretion for the trial judge to exercise, only a *balancing exercise* to undertake. The trial judge is required to balance the probative value of the proffered and otherwise admissible evidence against any ‘real’ danger(s) of unfair prejudice to the defendant:

Application of s137 requires a balancing, by the trial judge, of the probative value of the evidence against the danger of unfair prejudice to the defendant. If that

39 In some circumstances, procedural unfairness, like the inability to cross-examine a witness, may create the danger of unfair prejudice. See, eg, *Ordukaya v Hicks* [2000] NSWCA 180, [31]–[41] (Sheller JA, Meagher JA concurring); *R v Suteski* (2002) 56 NSWLR 182, [126]–[127] (Wood CJ at CL, Sully and Howie JJ) (‘*Suteski*’); *Galvin v Regina* [2006] NSWCCA 66, [40] (Howie J, McClellan CJ and Latham J concurring).

40 *R v Yates* [2002] NSWCCA 520, [252] (Wood CJ at CL, Hulme and Bidden JJ) (emphasis in original) (‘*Yates*’). See also *R v Suteski (No 4)* [2002] NSWSC 218, [42]–[54] (Kirby J); *R v Lockyer* (1996) 89 A Crim R 457, 460 (Hunt CJ at CL); *Hannes v DPP (Cth) (No 2)* (2006) 205 FLR 217, [315] (Barr and Hall JJ).

41 *R v BD* (1997) 94 A Crim R 131, 151 (Hunt CJ at CL, Bruce J concurring) (emphasis added). This case was endorsed in *Papakosmas* (1999) 196 CLR 297, [29] (Gleeson CJ and Hayne J), [91]–[95] (McHugh J).

42 *Suteski* (2002) 56 NSWLR 182, [117] (Wood CJ at CL, Sully and Howie JJ concurring) (emphasis added). See also *R v GK* (2001) 53 NSWLR 317, [30].

43 *R v Shamouil* (2006) 66 NSWLR 228, [72] (Spigelman CJ) (‘*Shamouil*’).

balancing process results in a finding that the probative value is outweighed by the danger of unfair prejudice, the Court is constrained to refuse to admit the evidence. No element of discretion arises.⁴⁴

In practice, the balancing exercise creates difficulties. For, the trial judge is expected to balance incommensurables. Justice McHugh alluded to this difficulty in a discussion of the common law rules pertaining to the treatment of propensity evidence:

Nevertheless, the proposition that the probative value of the evidence must outweigh its prejudicial effect is one that can be easily misunderstood. The use of the term “outweigh” suggests an almost arithmetical computation. But prejudicial effect and probative value are incommensurables. They have no standard of comparison. The probative value of the evidence goes to proof of an issue, the prejudicial effect to the fairness of the trial. In criminal trials, the prejudicial effect of evidence is not concerned with the cogency of its proof but with risk that the jury will use the evidence or be affected by it in a way that the law does not permit.⁴⁵

Justice Scalia of the Supreme Court of the United States made a similar point, more creatively, when he described the need to determine ‘whether a particular line is longer than a particular rock is heavy.’⁴⁶

C Section 135: Broader Scope, Tougher Scales

135 General discretion to exclude evidence

The court may refuse to admit evidence if its probative value is substantially outweighed by the danger that the evidence might:

- (a) be unfairly prejudicial to a party; or
- (b) be misleading or confusing; or
- (c) cause or result in undue waste of time.

Section 135 overlaps substantially with section 137 and shares the common law heritage linking it to *Christie*.⁴⁷ The terms of section 135, particularly subsection (a), have the same meaning as those in section 137.⁴⁸

The following are the main differences between sections 135 and 137:

- Section 135 is discretionary (ie, ‘the court *may* refuse’) as opposed to section 137, which is mandatory (ie, ‘the court *must* refuse’);

44 *R v Cook* [2004] NSWCCA 52, [27] (Simpson J, Ipp JA and Adams J concurring) (‘Cook’). See also *Blick* [2000] NSWCCA 61, [20] (Sheller JA).

45 *Pfennig v R* (1995) 182 CLR 461, [39]. See also *R v Morris* (1969) 54 Cr App R 69, 83 (Widgery LCJ).

46 *Bendix Autolite Corp v Midwesco Enterprises Inc*, 486 US 888, 897 (1988).

47 The development of the Australian uniform evidence law was also influenced by the US *Federal Rules of Evidence* (1975). Rule 403 is entitled ‘Exclusion of Relevant Evidence on Grounds of Prejudice, Confusion, or Waste of Time’ and the text reads, ‘Although relevant, evidence may be excluded if its probative value is substantially outweighed by the danger of unfair prejudice, confusion of the issues, or misleading the jury, or by considerations of undue delay, waste of time, or needless presentation of cumulative evidence.’

48 *Ainsworth v Burden* [2005] NSWCA 174, [99] (Hunt AJA, Handley and McColl JJ concurring).

- Section 135 applies to evidence that might be ‘misleading or confusing’ or ‘cause or result in undue waste of time’ *as well as* evidence that might be ‘unfairly prejudicial’;
- Section 135 applies in all types of proceedings (ie, civil and criminal) and to all parties (ie, prosecution, plaintiff and defendants) whereas section 137 applies only in ‘criminal proceedings’ and only to ‘evidence adduced by the prosecutor’; and
- Section 135 has a different set of ‘scales’. The balancing exercise in section 135 is biased in favour of admission. According to section 135, ‘the danger’ must ‘substantially outweigh’ the ‘probative value’ to enliven the court’s *discretion* to exclude.

The last difference is significant because the discretion to exclude otherwise admissible evidence only arises where the probative value is ‘*substantially outweighed*’ by one of the enumerated dangers. Here, the judge undertakes a balancing exercise that requires more than just outweighing. The danger must substantially (or ‘well’ or ‘considerably’) outweigh the probative value of the evidence.⁴⁹ Put another way, the enumerated danger must be considerably higher than the probative value. This is a much harder threshold to satisfy than the standard imposed by section 137. It is also a threshold which merely enlivens the trial judge’s discretion. Section 137, therefore, intervenes earlier than section 135(a) and leaves no role for it to play in relation to evidence adduced by the prosecutor.

Subsections (b) and (c) are expressed in different terms to sections 135(a) and 137. There is not much instructive case law in this area.⁵⁰ In practice, judges have been reluctant to exclude expert opinion evidence simply because it might be complicated or confusing or waste the court’s time, particularly in criminal prosecutions. There is, on the contrary, an assumption that juries can (and should) cope with most expert opinion evidence and virtually all expert disagreement.⁵¹

Even though its scope is wider, section 135 has exerted limited impact on criminal proceedings. On its face, section 135 provides a trial judge with discretionary means of excluding evidence in addition to the protections afforded by section 137. In practice, however, section 137 seems to overlap substantially with all of section 135. If recourse to section 137 does not lead to the exclusion of unfairly prejudicial evidence adduced by the prosecutor, it is highly unlikely

49 Clark [2001] NSWCCA 494, [163] (Heydon JA, Dowd and Bell JJ concurring). See also John Wigmore, *The Principles of Judicial Proof* (1913) 52–3.

50 See, eg, *Cadbury Schweppes Pty Ltd v Darrell Lea Chocolate Shops Pty Ltd* (2007) FCR 397, [55], [102]; *Australian Securities Investments Commission v Rich* (2005) ALR 764; *Fina Research SA v Halliburton Energy Services Inc* [2003] FCA 55.

51 *Lisoff* [1999] NSWCCA 364, [60] (Spigelman CJ, Newman and Sully JJ). The appeal in *Lisoff* seems to have been decisive in this regard. The Court explained that ‘[s]ection 137 requires a real risk of unfair prejudice to the defendant by reason of the admission of the [complex scientific] evidence complained of. It is not sufficient to establish that the complexity or nature of the evidence was such that it created the mere possibility that the jury could act in a particular way’. But compare *R v McNeill (Ruling No 2)* [2007] NFSC 3 (Weinberg CJ).

that a judge will use the discretion afforded by section 135 to exclude otherwise admissible evidence on the grounds of confusion or delay.⁵²

D Determining ‘Probative value’: Taking the (Expert) Evidence ‘at its highest’

Now we turn to consider how judges *determine* the ‘probative value’ of evidence when applying sections 135 and 137.

Most of the section 137 jurisprudence is concerned with lay evidence. We can understand why a judge might not want to exclude such evidence on the basis of his or her assessment of its probative value if, notwithstanding his or her own impression, a jury might find the evidence probative or even compelling. Reluctance to trespass on the prerogatives of the jury is probably easiest to understand in relation to the credibility and the reliability of the evidence of lay witnesses. In consequence, judges have tended to approach the determination of the probative value of evidence when applying sections 135 and 137 (or the *Christie* discretion at common law) by attributing to the evidence the highest possible probative value that it can support, and then balancing any real danger of unfair prejudice against that maximum value. That is, they allocate the highest probative value that a jury could (in theory, ‘rationally’) assign, and balance the danger of unfair prejudice against that value. At no stage does the trial judge attempt to determine the actual probative value of the evidence.

This approach to probative value is conspicuous in common law cases such as *R v Edelstein*, *R v McLean and Funk*; *Ex parte Attorney-General*, *Tugaga v R*, *R v Sandford* and *Rozenes v Beljajev*.⁵³ We, however, are primarily concerned with the operation of the *Evidence Act*. In NSW, the case of *R v Carusi* provided an early and influential resource in this area. Hunt CJ at CL (with Newman and Ireland JJ agreeing) applied the Act in terms highly reminiscent of the common law:

The power of the trial judge to exclude evidence in accordance with the *Christie* discretion does not permit the judge, in assessing what its probative value is, to determine whether the jury should or should not accept the evidence of the witness upon which the Crown case depends. The trial judge can only exclude the evidence of such a witness where, *taken at its highest*, its probative value is outweighed by its prejudicial effect ...⁵⁴

These sentiments were re-reiterated by Hunt CJ at CL (with McInerney J and Donovan AJ agreeing) in *R v Singh-Bal* and affirmed by a more recent series of

⁵² See, eg, *R v Rose* (2002) 55 NSWLR 701, [349]–[395] (Smart AJ).

⁵³ *R v Edelsten* (1990) 21 NSWLR 542 (Carruthers, Allen and Badgery-Parker JJ); *R v McLean and Funk*; *Ex parte Attorney-General* (1990) 47 A Crim R 240, 255–260 (Carter J); *Tugaga v R* (1994) 74 A Crim R 190 (Hunt CJ at CL, Gleeson CJ and Abadee J concurring); *R v Sandford* (1994) 72 A Crim R 160, 178 (Hunt CJ at CL, Smart and Studdert JJ concurring); *Rozenes v Beljajev* (1994) 126 ALR 481 (Brooking, McDonald and Hansen JJ).

⁵⁴ *R v Carusi* (1997) 92 A Crim R 52, 65–66 (Hunt CJ at CL, Newman and Ireland JJ concurring) (emphasis added).

decisions, including *R v Yates*, *R v Le*, *R v Rahme* and *R v Nguyen*.⁵⁵ Though, the most detailed and emphatic in this line of authority is the decision of Spigelman CJ in *Shamouil*.⁵⁶

Shamouil was a Crown appeal against the trial judge's exclusion of incriminating lay identification evidence. Writing for the CCA, Spigelman CJ (with Simpson and Adams JJ agreeing) offered a version of the historical development of the 'taken at its highest' doctrine.⁵⁷

Before the Evidence Act the *Christie* discretion to exclude evidence at common law for which s137 is a replacement, did not involve considerations of reliability of the evidence. ... After the enactment of s137, the same approach was taken in *R v Singh-Bal* (1997) 92 A Crim R 39 at 403 and *R v Yates* [2002] NSWCCA 520 at [255]-[256], in both of which the formulation from *R v Carusi* was expressly adopted, i.e. the evidence must be "taken at its highest" in order to determine its probative value.⁵⁸

He continued, referring to the definition of 'probative value' in the *Evidence Act Dictionary*:

In my opinion, the critical word in this regard is the word *could* in the definition of probative value ... namely, 'the extent to which the evidence could rationally affect the assessment ...'. The focus on *capability* draws attention to what it is *open* for the tribunal of fact to conclude. It does not direct attention to what a tribunal of fact is *likely* to conclude. Evidence has 'probative value', as defined, if it is capable of supporting a verdict of guilty.⁵⁹

The decision in *Shamouil* illustrates how at common law and (perhaps even more uncompromisingly) under the *Evidence Act*, judges have been reluctant to determine the actual probative value of evidence when undertaking the balancing exercise between probative value and the danger of unfair prejudice. *Shamouil* stands for the proposition that judges should take the evidence *at its highest* when engaged in the balancing exercises mandated by sections 135 and 137.

Of especial interest, this approach is not restricted to the evidence of lay witnesses but also seems to apply to expert opinion evidence.

E The Irrelevance of 'Reliability' and 'Credibility'

We have already observed how the Criminal Court of Appeal explicitly, though perhaps unwisely, distinguished 'specialised knowledge' from 'extraneous' ideas like 'reliability' in *Tang*. The exclusionary discretions, insofar as they involve undertaking the balancing exercise on the basis of determining

55 *R v Singh-Bal* (1997) 92 A Crim R 397, 403-4 (Hunt CJ at CL, McInerney J and Donovan AJ concurring); *R v Yates* [2002] NSWCCA 520, [255] (Wood CJ at CL, Hulme and Bidden JJ); *R v Le* [2002] NSWCCA 186, [50] (Heydon JA, Dunford and Buddin JJ concurring); *R v Rahme* [2004] NSWCCA 233, [198]-[204] (James J) ('*Rahme*'); *R v Nguyen* [2006] NSWSC 834, [33], [42] (Hulme J). This last decision represents Hulme J's acquiescence after *R v Shamouil* (2006) 66 NSWLR 228 ('*Shamouil*').

56 See also *R v Mundine* [2008] NSWCCA 55 (Simpson J, McClelland CJ at CL and Grove J concurring).

57 This was the same panel that heard the appeal in *Tang*.

58 *Shamouil* (2006) 66 NSWLR 228, [49]. The expression 'taken at its highest' has also been used in relation to removing cases from the jury. See, eg, *R v Galbraith* [1981] 1 WLR 1039, 1042; *Doney v The Queen* (1990) 171 CLR 207, 215 (Deane, Dawson, Toohey, Gaudron and McHugh JJ).

59 *Shamouil* (2006) 66 NSWLR 228, [61] (emphasis in original). Omission of 'rationally' shifts the meaning from an objective to a subjective standard.

probative value by taking the evidence at its highest, perpetuate general disinterest in the reliability of evidence and the credibility of witnesses. We can see this clearly in *Shamouil*:

The preponderant body of authority in this Court is in favour of a restrictive approach to the circumstances in which issues of reliability and credibility are to be taken into account in determining the probative value of evidence for purposes of determining questions of admissibility. There is no reason to change that approach.⁶⁰

This approach to ‘probative value’ resembles its treatment in other parts of the *Evidence Act*. Considering the admissibility of tendency evidence (under sections 97 and 101), also concerned with balancing probative value against prejudicial effect, Adam J (with whom Spigelman CJ and Sully J concurred) followed the ‘preponderant body of authority’:

In my view, the probative value of that evidence was high in the circumstances of this case, *upon the assumption, of course, that it was true*, but that must be necessarily the assumption with which s101 is concerned.⁶¹

Here, taking the evidence at its highest is equated with assuming ‘that it was true’.

Earlier, in *Adam v The Queen*, Gaudron J had offered an interpretation of ‘probative value’ that seems to have reinforced this general outlook. Her Honour incorporated the phrase ‘if it were accepted’ from section 55 into the definition of ‘probative value’.

The dictionary to the Act defines “probative value” to mean “the extent to which the evidence could rationally affect the assessment of the probability of the existence of a fact in issue”. That definition echoes the substance of s 55(1) of the Act which provides that “evidence is relevant in a proceeding is evidence that, if it were accepted, could rationally affect (directly or indirectly) the assessment of the probability of the existence of a fact in issue in the proceeding”. It is to be noted that the dictionary definition differs from s 55 in that it is not predicated on the assumption that the evidence will be accepted.

The omission from the dictionary definition of “probative value” of the assumption that the evidence will be accepted is, in my opinion, of no significance. As a practical matter, evidence can rationally affect the assessment of the probability of a fact in issue only if it is accepted. Accordingly,⁶² the assumption that it will be accepted must be read into the dictionary definition.

Justice Gaudron’s approach was explicitly endorsed by the majority in *R v Rahme*⁶³ as well as in *Shamouil*.⁶⁴

60 Ibid [60], [64]–[65].

61 *R v AB* [2001] NSWCCA 496, [17] (Adams J, Spigelman CJ and Sully J concurring) (emphasis added). Section 101 governs the prosecution use of tendency and coincidence evidence. More onerous than section 137, section 101(2) states that tendency and coincidence evidence ‘adduced by the prosecution cannot be used against the defendant unless the probative value of the evidence substantially outweighs any prejudicial effect it may have on the defendant’. Where section 101 applies, sections 137 and 135(a) have no application. See also *Lock* (1997) 91 A Crim R 356, 360; *R v Fletcher* [2005] NSWCCA 338, [112]–[113] (Simpson J, McClellan CJ at CL concurring) (*‘Fletcher’*).

62 *Adam v The Queen* (2001) 207 CLR 96, [59]–[60] (*‘Adam’*). This evades questions of acceptance, credibility and reliability.

63 *Rahme* [2004] NSWCCA 233, [201]–[202] (James J).

64 *Shamouil* (2006) 66 NSWLR 228, [53], [55], [62].

Even if we accept that the position championed by Spigelman CJ, Adam and Gaudron JJ represents the dominant approach to ‘probative value’, we should not overlook the fact that several senior judges have, at least, questioned this interpretation. A number of judges have suggested that issues of reliability and credibility should not be automatically abandoned to the jury. There may yet, according to this lesser line of authority, be a place for judicial assessments of reliability and probative value. We can identify these tensions in cases like *Papakosmas* and the dissent in *Rahme*.

Papakosmas was decided a couple of years before *Adam*. There, McHugh J intimated that when assessing ‘probative value’ the ‘reliability’ of the evidence could not be disregarded.⁶⁵

To the extent that other policies of evidence law, such as procedural fairness and reliability, required the strict logic of the relevance rule to be modified, that could best be done by the exclusionary rules—such as the hearsay rule and the credibility rule—and by conferring discretions on the court as in ss 135-137. ...

The distinction which the Act makes between relevance and probative value also supports the view that relevance is not concerned with reliability. Probative value is defined in the Dictionary of the Act as being “the extent to which the evidence could rationally affect the assessment of the probability of the existence of a fact in issue”. *That assessment, of course, would necessarily involve considerations of reliability.* “Probative value” is an important consideration in the exercise of the powers conferred by ss 135 and 137.⁶⁶

Justice McHugh’s decision provides a stark contrast to the position advanced by Gaudron J in *Adam* and Spigelman CJ in *Shamouil*. Justice McHugh distinguished between relevance (s 55) which is not concerned with reliability – remember ‘if it were accepted’ – and the definition of probative value – which requires that the evidence ‘could *rationally* affect’. There is, notwithstanding *Adam*, no qualification in the definition of ‘probative value’. For McHugh J ‘reliability’ is a necessary consideration.

The second example emerges from Hulme J’s dissent in *Rahme*. This statement provides one of the clearest expressions of this alternative line of reasoning:

... given the history and care that went into the drafting of the Evidence Act, I am unable to accept, consistently with general canons of construction, that the omission in the definition of “probative value” of any reference along the lines “if it were accepted” should be treated as a matter of no significance, when these words do not appear in the exposition of relevance in s55. ...

By virtue of the words used in the definition, any consideration under the Evidence Act of the probative value of evidence requires an assessment of “the extent to which the evidence could rationally affect the assessment of the probability of the existence of a fact in issue. ...

The need to consider the “extent” in the context of “rationally affect” to my mind argues for an assessment of the credibility of the author and the likelihood of the evidence being accepted. This is not to deny that operation must also be given to the word “could” in the expression “could rationally affect”. When a judge is required to consider the probative value of evidence, the test is not simply whether the judge believes it.

65 *Papakosmas* (1999) 196 CLR 297, [86].

66 *Ibid* [81], [86] (emphasis added).

Many of the occasions contemplated by the Evidence Act as to requiring an assessment of probative value also point in the direction of requiring, or at least permitting, as assessment of the credibility or reliability of the evidence under consideration. These include comparison with “any prejudicial effect it (the evidence) may have on the defendant”—s101, “the danger (the evidence) might be unfairly prejudicial ... misleading or confusing, or cause or result in undue waste of time”—s135, and “the danger of unfair prejudice”—s137. It strikes me that a far more useful comparison with these matters can be made if a comprehensive assessment of the value of the evidence under consideration can be made, rather than an assessment circumscribed by a prohibition on considering the credibility or reliability of the author of the evidence.⁶⁷

Questions about the need to assess ‘probative value’ have emerged in many other cases, including *R v Cook* and *R v Zhang*.⁶⁸

Reviewing cases like *Carusi*, *Adam*, *Papakosmas*, *Rahme* and *Cook*, the recent *Uniform Evidence Law Report* (2005) produced by the ALRC, the NSW Law Reform Commission and the Victorian Law Reform Commission suggested that questions about credibility and the reliability of evidence remain ‘open’:

The question is open as to whether probative value is determined solely on the basis of the degree of relevance or whether the court is permitted to consider the credibility and reliability of the evidence. This issue has arisen mostly in the context of jury trials, and hence the relevant question has been whether the judge may consider whether the jury should accept the evidence.⁶⁹

Though relatively conservative in its outlook – and insensitive to issues relating to the validity and accuracy of expert opinion evidence – the Report expressed concern that the ‘inability to test the reliability of evidence may carry with it the danger of ... mis-estimation’.⁷⁰ It continued:

It is therefore consistent with the policy basis for the discretion that the inability to test evidence may constitute a legitimate ground for its exclusion where this will affect the ability of the fact-finder to assess rationally the weight of evidence. ... the Commissions are of the view that questions of credibility and reliability should generally be left to be determined by the tribunal of fact. Factors affecting the reliability or credibility of evidence usually emerge during the course of the trial, particularly in cross-examination. However, where the reliability or credibility of the evidence is such that its weight is likely to be overestimated by the tribunal of fact because of an inability to test the evidence by cross-examination or for some other reason, then these may be considerations relevant to the decision to exclude or limit the use of the evidence.⁷¹

67 *Rahme* [2004] NSWCCA 233, [220]–[223] (Hulme J) (emphasis in original) and see generally [217]–[224].

68 *R v Cook* [2004] NSWCCA 52, [37]–[38], [43] (Simpson J, Ipp JA and Adams J concurring); *R v Zhang* (2005) 227 ALR 311, [46] (Basten JA), [139] (Simpson and Buddin JJ concurring); *Fletcher* [2005] NSWCCA 338, [32]–[35] (Simpson J, McClellan CJ at CL concurring). Interestingly, in *R v Dann* [2000] NSWCCA 185, [37] (Heydon JA, Spigelman CJ and James J concurring), Spigelman CJ appears to endorse Heydon JA’s actual *determination* of the ‘low probative value’ of the medical evidence. See also *Galvin v Regina* [2006] NSWCCA 66, [27] (Howie J, McClellan CJ and Latham J concurring); *Nguyen v R* [2007] NSWCCA 363, [35]–[36] (Smart AJ, Mason P and Adams J concurring); *Nguyen v R* [2007] NSWCCA 249, [58], [65] (James J, Spigelman CJ and Hislop J concurring); *Blick* [2000] NSWCCA 61, [19] (Sheller JA, James and Dowd JJ concurring).

69 ALRC, *UEL Report*, above n 38, [16.16].

70 *Ibid* [16.45]. See also ALRC, *Evidence (Interim)*, above n 37.

71 ALRC, *UEL Report*, above n 38, [16.47] (emphasis added).

Notwithstanding tangible dangers, the only modification to Part 3.11 'Discretions to exclude evidence' recommended in the Report was to change the heading to 'Discretionary and mandatory exclusions' in order to reflect the mandatory nature of section 137.⁷²

So far, almost all of the case law we have considered has been developed in relation to lay witnesses and their evidence. We can understand why judges might not want to unilaterally interfere with juror assessments of the credibility of lay witnesses or the reliability of their evidence. The same arguments do not, however, apply with equal force to opinions based on 'specialised knowledge'. There would seem to be an obvious role for reliability when determining the probative value of inculpatory expert opinion evidence.

What we can say, on the basis of the foregoing review, is that taking expert opinion evidence 'at its highest' encourages the trial judge (and appellate courts) to assume that the techniques used and opinions presented by expert witnesses are reliable in preference to actually requiring the prosecution to persuade the trial judge that such an assumption is sound. It *may* make sense to adopt such a stance in response to the credibility of lay witnesses and the reliability of their evidence. This is because we have few dependable means of gauging credibility or the reliability of lay evidence.⁷³ However, the same cannot be said about the forensic sciences. Almost all forensic scientific techniques can be assessed for validity and reliability. The question then becomes: why do judges not make a principled distinction where the reliability of techniques and the probative value of expert opinion evidence can be meaningfully determined?

F Discussion

The dominant approach to probative value and the balancing exercise seems misconceived when applied to expert opinion evidence. If there are good reasons for doubting the reliability of evidence – and remember that we are primarily interested in forensic scientific evidence adduced by the prosecution – then judges should intervene:

It is difficult to see why a court, aware that there are particular reasons for doubting the reliability of certain evidence, should not be empowered to exclude the evidence if concluding that there is a real danger that the tribunal of fact will, even with the benefit of appropriate directions, misuse or significantly over-value the evidence.⁷⁴

In relation to expert opinion evidence, intervention and exclusion seem particularly warranted where there are formal means, such as testing or validation studies, to determine the reliability of techniques and opinions.

⁷² Ibid Recommendation 16–1.

⁷³ Ken Alder, *The Lie Detectors: The History of an American Obsession* (2007). But see Kent Roach, 'Unreliable Evidence and Wrongful Convictions: The Case for Excluding Tainted Identification Evidence and Jailhouse and Coerced Confessions' (2007) 52 *Criminal Law Quarterly* 210.

⁷⁴ Stephen Odgers, *Uniform Evidence Law* (7th ed, 2006) 638. In their *Australian Principles of Evidence* (2004) 146, Gans and Palmer suggest that if reliability is not used to determine probative value then it might be factored into prejudicial effect.

As things stand, there is a serious flaw in the dominant approach to the exclusionary discretions. The failure to assess the reliability of the evidence means that judges, taking the evidence ‘at its highest’ or assuming that the opinion is ‘true’, are only likely to use the discretions to exclude expert opinion evidence when it has *no probative value*. Odgers captures this limitation:

A difficulty with this analysis is that, if it could not be open to the jury to conclude that the evidence could rationally affect the assessment of the probability of the existence of a fact in issue, the evidence is not relevant. Accordingly, this provision would have no application, since the evidence would be inadmissible under s56(2).⁷⁵

In practice, the only circumstances in which evidence might be excluded using section 137 are when the probative value taken at its highest is low and there is a very conspicuous risk of unfair prejudice.⁷⁶ Where an unreliable opinion purportedly based on ‘specialised knowledge’ has a *high* probative value, ‘if accepted’ or ‘upon the assumption that it is true’, then there seems to be almost no danger of unfair prejudice that could tilt the scales in favour of exclusion.

For expert opinion evidence, the approach presented in *Shamouil* is tautological. Taking the prosecution’s expert evidence at its highest has come to mean assuming that the methods and conclusions *are* reliable. The upshot is that when it comes to balancing this (implicitly) *reliable* evidence against the danger of unfair prejudice – especially the danger that the jury may *overvalue* or *misuse* evidence of unknown reliability – that danger is removed by the initial assumption. There can be no risk of the jury overvaluing unreliable expert opinion evidence if the judge assumes that – notwithstanding obvious limitations, wide-ranging criticisms and a lack of substantiation – the evidence is reliable or true.

What the foregoing overview suggests is that when it comes to assessing opinion evidence based on ‘specialised knowledge’ adduced by the prosecution, judges are generally not interested in probative value. They have little interest in the validity and accuracy of techniques or evidence (or even the credibility of the expert witness). In this way the descriptions ‘Discretions to exclude evidence’ and ‘Discretionary and mandatory exclusions’ are highly misleading. Not only is section 137 not a discretion, but neither sections 135 nor 137 are very exclusionary. In effect, if the prosecution’s expert opinion evidence satisfies the incredibly low criteria for relevance – that is, a jury might accept that it has some bearing on the facts in issue – then it will almost certainly go before the jury. The scope for exclusion, once expert evidence is taken at its highest, is very narrow indeed. Where an expert opinion is unreliable (or of unknown reliability), but if accepted could be damning, under our current jurisprudence section 137 has no substantial role to play.

75 Odgers, above n 74, 640. See also Jill Anderson, Jill Hunter and Neil Williams, *The New Evidence Law* (2002).

76 See, eg, *R v Harvey* [1996] NSWCCA (Unreported, Beazley JA, Smart and James JJ, 11 December 1996); *Zammit v R* (1999) 107 A Crim R 489; *R v Patsalis*; *R v Spathis (No 2)* [1999] NSWSC 714, [6]–[7] (Kirby J); *R v Benecke* [1999] NSWCCA 163; *R v Knight* [2005] NSWCCA 241, [51]–[57] (Howie J, Grove and Rothman JJ concurring).

Now, having reviewed the operation of sections 79, 135 and 137, we can begin to appreciate some of the practical consequences of the failure to determine probative value and take the reliability of expert opinions more seriously. At this juncture it is useful to consider a recent example, from the aftermath of *Tang*, in a little more detail.

IV CASE STUDY: FACIAL MAPPING EVIDENCE AFTER *TANG* AND *SHAMOUIL*

Several issues from *Tang* re-emerged just a few months later in *R v Jung*.⁷⁷ Here we can observe how, notwithstanding the guidance provided by the CCA, a judge of the Supreme Court of NSW found Dr Sutisno's facial mapping evidence to be 'specialised knowledge' and concluded that section 137 would not exclude this otherwise admissible evidence. *Jung* is a good example because it was decided after *Carusi*, *Shamouil*, *Makita* and *Tang* and illustrates how a trial judge, working with the available jurisprudence, applied sections 79 and 137 to controversial opinion evidence in a double murder trial. In *Jung*, Hall J eventually decided to allow the prosecution to adduce and rely upon expert opinion evidence that was untested and of unknown reliability. As the example clearly illustrates, questions about probative value and the reliability of the expert evidence were left for the trial and the jury. The decision in *Jung* exemplifies pervasive disinterest in reliability and exaggerated confidence in cross-examination, rebuttal experts, directions and warnings.

During a police investigation Dr Sutisno was presented with photographs of a person of interest from an automatic telling machine (ATM) which dated back to 1997 and police photographs of Myoung Il Jung obtained in 2004 and 2005. After examining the two sets of images Dr Sutisno identified Jung as the person in the poor quality images taken by the ATM. The admissibility of this opinion evidence was challenged and a *voir dire* was conducted. Here the question was what, if anything, Dr Sutisno would be permitted to say about the identity of the person in the ATM photographs in the aftermath of *Tang*.

During the *voir dire* Dr Sutisno gave evidence that the right ear, in one of the photographs from the ATM, provided a 'unique identifier' which implicated the accused.⁷⁸ As in *Tang*, she positively identified the accused as the person of interest.⁷⁹ In her reports, Dr Sutisno set out tables which listed similarities, including some labelled as 'definitive resemblance'. Her tables also recorded several differences between facial and body features in the police and ATM images. These 'unmatched features' included the general body build and amount of buccal fat in the cheeks. Unlike the similarities though, the differences were characterised as either superficial or mutable. In her testimony, these differences were explained away by the effluxion of time.

77 [2006] NSWSC 658 ('*Jung*').

78 Claims about the value of ears as 'identifiers'—unique or otherwise—were also prominent in *R v Alrekabi* [2007] NSWDC 110.

79 Dr Sutisno's report may have been completed before *Tang* was handed down.

In *Jung*, Dr Sutisno's evidence was, once again, primarily based on morphological comparisons. She seems to have been – notwithstanding the claims about the ears – attempting to avoid some of the complications caused by earlier anthropometrical efforts and the conspicuous lack of information about the distribution of facial and body features. Justice Hall explained this approach:

Dr Sutisno's analysis is ... one based upon what is referred to as "*morphological analysis*" to determine visual similarities or differences. Her opinion as to "*similarities*" between the person on the ATM photographs and the forensic photographs is based on such an analysis. *Dr Sutisno describes her methodology as avoiding statistical matching or population samples and issues of probability. "It is purely a qualitative assessment of relative similarities and differences between images."*⁸⁰

Interestingly, the absence of measurements and quantification seems to have been a deliberate strategy.⁸¹

In order to obtain some sense of the limits to her expert opinion evidence, it is useful to list some of the concessions made by Dr Sutisno during the *voir dire*. They include:

She was not aware of anyone in Australia who has validated her methods.

She agreed that ... one of the problems with this type of [morphological] analysis was that facial features can be altered by photographic angles.

She had not taken part in any study of camera angles, or the operation of cameras to do with motion blur, lighting and the translation of data format.

She did not have knowledge of the actual angle that the lens was at to see the face of the man at the ATM. ...

She did not measure the angle of the head of the person depicted in the ATM photograph, although that can be done. ...

She did not keep the notes of measurements that she carried out, conceding it would be useful to be able to have dimensions of, say, a person's left ear, when comparing another human ear.⁸²

The only professional support for Dr Sutisno's morphological approach was her account of a telephone conversation with a Japanese professor.⁸³ Remarkably, notwithstanding the numerous concessions, Dr Sutisno 'was not prepared to admit that she could make a mistake using these techniques'.⁸⁴

The prosecution lawyers, sensitive to the recent appeal in *Tang*, indicated that they would only rely on 'evidence of similarities' and would not introduce opinions – like those expressed in Dr Sutisno's reports – about identity during

80 *Jung* [2006] NSWSC 658, [24] (Hall J) (emphasis added). See also *R v Kaliyanda* [2006] NSWSC (Unreported, Hislop J, 17 October 2006) and *US v Hines* 55 F Supp 2d 62 (D Mass 1999).

81 Here, somewhat curiously, Dr Sutisno's method is described in a way where it seems to be dictated by a desire to avoid the difficulties created by quantification and 'statistical matching'. It does not, as we shall see, resolve questions about the meaning and value of purported similarities (and differences) based on qualitative comparisons. See Simon Cole, *Suspect Identities: A History of Fingerprinting and Criminal Identification* (2001).

82 *Jung* [2006] NSWSC 658, [9]. But consider *US v Johnson* 114 F 3d 808 (8th Cir 1997) and Kitty Hauser, 'A Garment in the Dock; or, how the FBI Illuminated the Prehistory of a Pair of Denim Jeans' (2004) 9 *Journal of Material Culture* 293.

83 *Jung* [2006] NSWSC 658, [24].

84 *Ibid.*

the trial.⁸⁵ In effect, the prosecution sought to rely on Dr Sutisno's 'qualitative assessment of relative similarities and differences between images'.⁸⁶

In response to Dr Sutisno's reports and testimony, the defence called two witnesses to rebut her opinion evidence at the *voir dire* and trial. The first, Glenn Porter, was a former forensic photographer with the Australian Federal Police. He expressed reservations about the possibility of establishing identity based on the ATM photographs. Porter highlighted a range of what were described as artefacts, aberrations and limitations with the images relied upon by Dr Sutisno. These included:

- (a) Barrel distortion, a distortion of the image, with a barreling effect from centre to edges (fish eye perspective).
- (b) Poor dynamic range – a narrow dynamic range on the monochromatic photographs showing various levels of grey. This has the effect of reducing the level of detail and contours.
- (c) Poor exposure in critical areas, especially in the areas of faces of people depicted.
- (d) Poor level of resolution.
- (e) Lack of fine detail.
- (f) Motion blur where the subjects are moving to obtain money from the ATM.
- (g) Some subjects are out of focus (depth of field).
- (h) Poor focus overall.
- (i) High level of "noise" – which effects digital images – reduces the level of resolution.
- (j) Lack of highlight or shadow detail.⁸⁷

Porter also referred to the absence of information about 'view, distance and lighting' in the forensic photographs taken of the accused. He 'emphasised that similarity of image perceptive [sic] and similar distances between lens and subject in relation to the ATM photographs and the forensic photographs were important to obtain accuracy'.⁸⁸

The other expert, Dr G.A. Doran, was a dental surgeon and consultant forensic anatomist. Dr Doran's evidence suggested that without better images, knowledge of camera angles and measurements, any identification or comparison would be 'very unreliable'.⁸⁹

⁸⁵ Ibid [28].

⁸⁶ Ibid [24].

⁸⁷ Ibid [11]. All this before we even begin to consider the possibility of image manipulation: see *US v Frabizio* 445 F Supp 2d 152 (D Mass 2006) and Glenn Porter, 'Visual Culture in Forensic Science' (2007) 39 *Australian Journal of Forensic Sciences* 81.

⁸⁸ *Jung* [2006] NSWSC 658, [12].

⁸⁹ Ibid [16].

Dr Doran, in his evidence, stated that with facial mapping, one looks at a number of features or regular well-known anatomical identifying features. By ascertaining the size, dimension and general morphology in shape, size and form, one can then establish parameters that can be used for comparison.

He stated that conclusions as to comparisons are very dependent on the actual physical photographic techniques and methods and angles of photography. A horizontal plane is usually quite accurate, whereas a vertical plane can be difficult. If a face is tilted in any way, that can affect accuracy.

In relation to the ATM photographs in these proceedings, he considered that many of them were at various angles and quite blurred and out of focus. Dr Doran identified difficulties with the photographs used as a basis for comparison in these proceedings. Having only seen the ATM photographs, he said the problem in assessing them would be the angle of the camera, the blurriness or the degree of focus and quality of the photographs as well as the various angles which the individual's head has been turned in the various photographs. He considered it was very difficult in assessing portions or parts of the human face in terms of large, medium and small. That was one of the difficulties he had in reading Dr Sutisno's reports. He stated that not using anthropometric measurements, one had removed the one quantifiable parameter used in facial identification and comparison. Without such measurements, he said it was "*very difficult to say what is large or what is medium*". Such descriptions were very qualitative and very subjective. He did not consider that they would be reliable. The descriptions would be valid and correct if they were used in conjunction with quantitative methods.⁹⁰

The two defence experts expressed concerns about the quality of images and the limited information relied upon by Dr Sutisno in her analysis (and identification).⁹¹ It was their contention that the information and methods were insufficient to ground the kinds of comparisons (and identifications) she was offering.

In determining whether Dr Sutisno's evidence was admissible, Hall J explained that the ATM photographs were 'not self-evident and are such that the jury would require the assistance of an expert such as Dr Sutisno'.⁹² On this basis the 'morphological analysis involving opinion evidence as to similarities ... passes the test of relevance'.⁹³

In terms of the admissibility of the expert opinion evidence – that is, the question of whether Dr Sutisno possessed 'specialised knowledge based on ... training, study or experience' – Hall J referred to her academic qualifications, recent practice as a forensic anatomist and 'practical experience in facial identification'.⁹⁴

The evidence does establish that Dr Sutisno does have specialised knowledge based on study and experience in relation to facial characteristics in the context of issues concerned with establishing identification both of deceased persons and otherwise. Such has become a recognised field for expert analysis, albeit of fairly recent

90 Ibid [13]–[15].

91 Other experts, including Dr Richard Kemp (an experimental psychologist) and Professor Maciej Henneberg (an anatomist and physical anthropologist), had criticised Dr Sutisno's opinion evidence in earlier proceedings. See, eg, *R v Murdoch (No 4)* (2005) 195 FLR 421.

92 *Jung* [2006] NSWSC 658, [45]. The ability to testify and need for assistance arose from the decision in *Smith v R* (2001) 206 CLR 650. For a detailed theoretical discussion see Katherine Giber, *Captive Images: Race, Crime, Photography* (2007).

93 *Jung* [2006] NSWSC 658, [50].

94 Ibid [51].

origin. The opinion she has expressed as to similarities is an opinion within that field of specialised knowledge.⁹⁵

The fact that Dr Sutisno was not an expert in photography or the use of software imaging programs did not detract from the admissibility of her opinion evidence. For the trial judge: 'The quality, suitability and reliability of the photographic images are no doubt important questions going to the reliability of the factual material Dr Sutisno has relied upon in the course of and for the purposes of her anthropological analysis'.⁹⁶ These, however, were issues for the jury.

The question of whether the opinion expressed by Dr Sutisno was wholly or substantially based on her specialised knowledge was answered in the affirmative:

The opinion as to anatomical similarities expressed by Dr Sutisno in her reports is, in my opinion, substantially based on the specialised knowledge to which I have earlier referred and the application of such knowledge to the photographic images in question.⁹⁷

Explaining this aspect of the decision, Hall J drew support from *Makita*, particularly the emphasis on the need to explain the basis, or reasoning behind, the opinion evidence:

However adequate or inadequate the photographic materials utilised by Sutisno for the purpose of her analysis, the evidence on the voir dire does not establish that she has failed to disclose the factual material she has utilised (the photographic images), the nature of the methodology that she has employed and the type of analysis described in her reports (morphological analysis). I have carefully reviewed the reports and her evidence in order to determine whether it may properly be said that, having regard to the specific principles governing admissibility of expert evidence as identified by Heydon, JA. in *Makita* (supra) and as summarised above, Dr Sutisno's evidence complies with the requirements for admissibility.

...

Dr Sutisno's opinion has been challenged, inter alia, upon the basis that she has not applied or sought measurable data for the purposes of her comparative analysis. Dr Sutisno, in her reports, proffers an explanation as to why this has not been done. *Whether that explanation is sound is, of course, a matter of fact.* The fact that she has elected to adopt a qualitative form of analysis does not, per se, render it inadmissible, provided that the provisions of s.79 and what I might refer to as the *Makita* principles, have otherwise been satisfied.⁹⁸

Dr Sutisno had explained how she compared police photographs of the accused with images of the person of interest recorded by the ATM camera. Her PhD in anatomy and the comparison of the images, rather than information about the validity and accuracy of her qualitative approach, rendered the Crown's highly incriminating (and highly prejudicial) expert opinion evidence admissible.

The final consideration for the *voir dire* was the application of section 137 to Dr Sutisno's de facto identification evidence. Acknowledging the existence of

95 Ibid [55] (emphasis added).

96 Ibid [58].

97 Ibid [59].

98 Ibid [62]–[65] (emphasis added). 'Matters of fact', of course, are for the tribunal of fact.

circumstances in which a judge might withhold ‘issues of credibility or reliability’ from the jury – where the evidence was irrelevant and therefore ‘it would not be open to the jury to conclude that the evidence could rationally affect the assessment of the probability of the existence of the fact in issue’ – Hall J embraced *Shamouil*:

The Chief Justice also observed that before the *Evidence Act* the Christie direction to exclude evidence at common law for which s.137 is a replacement, did not involve considerations of reliability of the evidence.

...

In the present matter, evidence is not evidence as to observation by a lay observer as to identity or similarity. *The context in which the present matter is being considered is one in which there is a disputed question of expert opinion evidence in which there is likely to be a divergence of views on the expert analysis of morphological features.* The question of reliability or credibility is to be evaluated in that context.

Accordingly, in undertaking the balancing exercise required by s137, *it is not appropriate for the purposes of the present application, for a determination to be made as to the quality or reliability of the evidence as those are matters for jury determination.*

The probative value of the opinion evidence of Dr Sutisno, *if accepted*, would be as relevant circumstantial evidence of similarities relevant to the question of identity. *If the jury were to accept* Dr Sutisno’s analysis and opinion based on it, then it would constitute important evidence in the Crown case.⁹⁹

For Hall J, credibility and reliability were not part of the balancing exercise required by section 137. Once the expert evidence was deemed admissible, the fact that the experts disagreed, regardless of the substance or significance of their disagreement, provided a reason for allowing their evidence to go before the jury. It would be for a jury to resolve the conflicts between them.

Justice Hall then turned to consider, drawing on *Yates* and *Lisoff*, whether there was a real risk of unfair prejudice, such as the jury adopting an illegitimate form of reasoning or giving Dr Sutisno’s opinion evidence undue weight. Interestingly, his anxieties were not dispelled by evidence of reliability or evidence that the jury would not engage in illegitimate reasoning or give the evidence too much weight. Rather, for Hall J, potential dangers were averted by the participation of the accused’s rebuttal experts:

The accused led evidence from two expert witnesses to whom I have already referred, namely, Mr Porter and Dr Doran. It is not irrelevant in the balancing exercise under s.137 in relation to the issue of unfair prejudice for some account to be taken of the fact that the accused has available to him such expert evidence which bears directly upon the analysis undertaken by Dr Sutisno.¹⁰⁰

Significantly, the trial judge seems to have been relieved by the presence of experts called by the defence. Their participation appears to dispel doubts about unfairness or the risk of unfairness created by the admission of Dr Sutisno’s incriminating opinion evidence.

99 Ibid [73]–[78] (emphasis added). ‘Direction’ should, presumably, read ‘discretion’.

100 Ibid [82].

Finally, quoting from the decision in *Lisoff*, Hall J explained that the complexity of expert evidence was not an appropriate reason for inferring that a jury “‘might” be affected in a prejudicial way’.¹⁰¹ He continued:

I do not consider, *on balance*, that it can be said that there is a real risk of unfair prejudice based upon what was asserted by senior counsel for the accused, namely, that there may be undue reliance by the jury on the expert evidence of Dr Sutisno and that there may be difficulty in them comprehending and giving effect to the evidence as to the significance of perspectives, angles, distances and other matters relevant to photograph images.¹⁰²

For Hall J, the conflict between the experts was not so ‘extraordinary’ that ‘a careful and sensible jury, *properly directed* as to the relevant law and as to the relevant evidence, could not decide in a reasoned and responsible way’.¹⁰³

Dr Sutisno was allowed to testify and now, it seems, facial mapping has become a legally recognised form of ‘specialised knowledge’ in the courts of NSW.

A Discussion

There is much to say about the admissibility determination in *Jung*. At no stage was Hall J concerned with the reliability of Dr Sutisno’s expert opinion evidence. The fact that her techniques could have been validated was not considered significant. Because facial mapping evidence is so epistemologically flimsy, the case provides a particularly useful illustration. The inability (or unwillingness) to exclude facial mapping evidence using sections 79 and 137 (and section 135) demonstrates the weaknesses of our current approach to expert opinion evidence adduced by the prosecution.

In addition to the concerns expressed by Porter and Dr Doran, there are many other serious issues which threaten the reliability and probative value of the facial mapping (or *de facto* identification) evidence.¹⁰⁴ First, there are no large databases recording the distribution of anatomical features or the relationships between anatomical features. Even if we temporarily ignore the problems raised by focusing and distortion, angles and distances, lenses, lighting and depth of field, the conversion of three dimensional features to two dimensions – especially given the limited information recorded by many security cameras – we really cannot say much about apparent or purported *similarities* (and *matches*). The situation is analogous to early DNA results and population statistics.

Commenting on this problem in *R v Lucas*, Hampel J, like Shannon QC before him in the inquiry into the conviction of Edward Charles Splatt, expressed

101 Ibid [83].

102 Ibid [85] (emphasis added).

103 Ibid [84] (emphasis added).

104 A more extensive assessment was undertaken in Gary Edmond and Richard Kemp, ‘Suspect Science? Facial Mapping, CCTV and Legal Identification’ (Paper presented at the UNSW School of Law Seminar Series, Sydney, 24 October 2007; the University of Melbourne Law & Society Conference, Melbourne, 29 November 2007; the Cornell University Department of Science and Technology Studies, Ithaca, 25 February 2008).

concern that similarities or consistencies might be equated with identity (or origin).¹⁰⁵

There is no way the jury can properly weigh the value of such evidence if there is no evidence before it as the frequency of a match in the general population. I agree with Mr Shannon's observations and *I think that there is in this case the danger than consistency could assume the colour of identity, or at least of probability*. ... It follows that the evidence of the DNA testing, its results and conclusions is not admissible because it lacks sufficient probative value compared with its possible prejudicial effect.¹⁰⁶

Justice Hampel and Commissioner Shannon were both apprehensive lest 'expressions of conclusions by the scientists' provide "the bridge over which the jury could step in passing from the path of similarities to the separate rock of commonality of source of origin".¹⁰⁷ Shannon QC characterised this reasoning as 'highly dangerous' and it would seem to be one of the kinds of improper reasoning, or misuses of evidence, that section 137 was designed to prevent. Interestingly, in *Lucas*, the DNA-based identification evidence was excluded even though the basic biological techniques were relatively stable and routinely used in mainstream biomedical research.¹⁰⁸

Morphological (and anthropometrical) facial mapping techniques, by way of comparison, have not been validated, published or accepted by other scientists.¹⁰⁹ In *Tang*, Dr Sutisno testified about 'unique identifiers' and in *Jung* about 'unique identifiers' and 'definitive resemblance' even though she presented no information about the frequency or distribution of facial and body features among the Australian population or its sub-groups. In the absence of data on the distribution and frequency of facial and body characteristics (and their interrelations), these claims about similarities between anatomical features tell us little about identity and engender the dangers described by Hampel J and Shannon QC.¹¹⁰

Notwithstanding these limitations, it is common for facial mappers to testify in very confident terms. Dr Sutisno testified that she was certain about her *identification* and was unwilling during cross-examination, to concede even the

105 *R v Lucas* [1992] 2 VR 109 ('*Lucas*'). See also the approach to 'matching' ballistics in *R v Ngo* [2001] NSWCCA 1021, [1]–[7] (Dunford J).

106 *Lucas* [1992] 2 VR 109, 118 (emphasis added).

107 South Australia, Royal Commission Concerning the Conviction of Edward Charles Splatt, *Report* (1984) 39.

108 Jay Aronson, *Genetic Witness: Science, Law, and Controversy in the Making of DNA Profiling* (2007); Michael Lynch, 'Science Above All Else' in Gary Edmond (ed), *Expertise in Regulation and Law* (2004) 121.

109 Krista Kleinberg et al, 'Failure of Anthropometry as a Facial Identification Technique Using High-Quality Photographs' (2007) 52 *Journal of Forensic Science* 779; Ruth Costigan, 'Identification From CCTV: The Risk of Injustice' [2007] *Criminal Law Review* 591; Tony Ward, 'Liars, Damned Liars and Expert Witnesses: Video Identification and the Law of Expert Evidence' (1999) 12 *Security Journal* 29.

110 Here, it is useful to compare the reasoning in *Tang* at [143]: 'In the absence of any kind of objective standard or data base which is capable of leading to a quantification of probabilities, such as is given in the context of DNA evidence, evidence of similarity may have a cumulative effect, by reason of the number of points of similarity.' Dr Sutisno, and the judges, seem to allow the preference for a morphological or qualitative – ie, comparative – approach to circumvent the need to attend to the distribution and significance of similarities as well as the degree of similarity.

possibility of error. This might be considered as remarkable as it is disturbing. Other identification *experts* have testified at the level of 95 per cent confidence even though their opinions were not based on any credible statistical analysis of the distribution of anatomical features among relevant populations.¹¹¹ Facial mappers, and other *experts* giving identification evidence, have been willing to make positive identifications and, perhaps remarkably, it has been appellate judges who have required them to limit their opinions to similarities.

Rather than address 'reliability', judges have prevented facial mappers from positively identifying the accused as the person of interest. Facial mappers are restricted to the language of similarity. Yet, they can testify in the clearest of terms that in their opinion there are no meaningful differences between a suspect and an offender. This represents a de facto form of identification evidence sometimes characterised as circumstantial identification evidence. In practice, restricting the scope of the expert's testimony may have made it more difficult for the defence to expose the over-confidence and methodological apathy of the facial mappers. Judicial intervention has forced the expert's conclusions to be presented in a more conservative fashion even though the underlying techniques remain the same.¹¹² The judicial 'solution' trivialises reliability and displaces questions about what apparent similarities actually reveal.

Of interest, experimentally-based research suggests that anthropometric and morphological comparisons are far more complicated, and far less reliable, than facial mappers routinely acknowledge (or seem to appreciate). Studies in related areas suggest that facial mapping techniques, including morphological approaches, are likely to have high error rates. There is little evidence, for example, that even highly trained experts – like anatomists or psychologists – are better than average (ie, than members of the jury) at making identifications from images in conditions where the persons of interest are unfamiliar.¹¹³ Even minor changes in clothing, appearance or expression confound the accuracy of comparisons. According to Bruce et al, 'superficial impressions of resemblance or dissimilarity between face images can be highly misleading'.¹¹⁴ Notably, it is not only experts who encounter difficulties. Experiments with computerised identification systems have experienced problems improving upon modest

111 *R v Alrekabi* [2007] NSWDC 110.

112 Nevertheless, it may be unwise to ask facial mappers about identification as a means of attacking their methods and hubris. This may merely introduce highly prejudicial personal beliefs without advancing the methodological critique.

113 Richard Kemp et al, 'When Seeing Should Not Be Believing: Photographs, Credit Cards and Fraud' (1997) 11 *Applied Cognitive Psychology* 211; Mike Burton et al, 'Face Recognition in Poor-Quality Video: Evidence From Security Surveillance' (1999) 10 *Psychological Science* 243; Zoe Henderson et al, 'Matching the Faces of Robbers Captured on Video' (2001) 15 *Applied Cognitive Psychology* 445; Michael Bromby and Hayley Ness, 'Over-observed? What is the Quality of CCTV in this New Digital Legal World?' (2006) 20 *International Review of Law, Computers & Technology* 105.

114 Vicki Bruce et al, 'Verification of Face Identities From Images Captured on Video' (1999) 5 *Journal of Experimental Psychology: Applied* 339, 358.

success rates in circumstances where the images are temporally proximate, full frontal, and of high quality.¹¹⁵

Further, the way facial mappers are brought into investigations and their close relations with the investigators provide additional grounds for concern. One of the procedural problems besetting facial mapping evidence is that ordinarily the expert is sent one set of photographs associated with a crime and one set of photographs of a suspect – not infrequently in the same envelope – and asked to comment. The implication is, of course, obvious.¹¹⁶ While methodologically undesirable, this kind of introduction *might* be less troublesome if there were standardised techniques which had been independently validated and were independently reviewed. Unfortunately, this is not the case. The upshot is that experts with untested techniques are exposed to prejudicial information that might have a direct bearing on their analysis and opinions. Just to reinforce the seriousness of this issue, empirical studies and experience with miscarriage of justice cases suggest that even in areas with established protocols and standardised techniques, extraneous information – such as police suspicions – can influence the outcome of expert analyses. One recent study found that when latent fingerprint examiners were told about an investigation they changed their assessment of whether fingerprint samples matched, in ways that were consistent with the extraneous information.¹¹⁷

Another of the methodological problems with facial mapping evidence, namely its highly selective nature, flows from the expectations placed on the *experts*. Facial mappers may spend hours poring over tapes and images searching for apparent similarities – that is, *cherry-picking*.¹¹⁸ Because of the way they enter the investigation, these experts are disinclined to actively search for differences or consider apparent differences as definitive.¹¹⁹ Hundreds or even thousands of images – the stills from a security video – may yield just a few similarities. On the basis of such surveys, facial mappers tend to testify that there were only similarities and no differences – or no differences that could not be explained away – between the person in the reference images supplied by the police and the images associated with the crime. Why the many images which do not provide similarities should be ignored or trivialised is never adequately explained. Why should images that appear to present similarities be preferred

115 Bridget Mallon, “‘Every Breath You Take, Every Move You Make, I’ll Be Watching You’: The Use of Face Recognition Technology” (2003) 48 *Villanova Law Review* 955; Jonathan Phillips et al, *FRVT 2006 and ICE 2006 Large-Scale Results* (2007) National Institute of Standards and Technology <<http://www.frvt.org/FRVT2006/docs/FRVT2006andICE2006LargeScaleReport.pdf>> at 24 March 2008. Thanks to Ashley Rose for these references.

116 Procedurally, this resembles the way identification parades and photographic arrays were conducted before they were subjected to judicial and statutory regulation. See, eg, *Evidence Act 1995* (NSW) Part 3.9.

117 Itiel Dror, David Charlton and Ailsa Peron, ‘Contextual Information Renders Experts Vulnerable to Making Erroneous Identifications’ (2006) 156 *Forensic Science International* 74.

118 They may even modify images for these purposes.

119 Facial mappers sometimes advert to instances where they have *excluded* suspects as an indication of their independence and the efficacy of their techniques. Neither of these claims necessarily follows.

over those that are unclear or might suggest dissimilarity? There are no analytical or statistical techniques to explain these prejudicial preferences.¹²⁰

Third, facial (and body) mapping is not ‘knowledge’, ‘specialised knowledge’ or an established ‘field of expertise’.¹²¹ There are no training schools or institutes, specialist journals, or standardised techniques. None of the various techniques seem to have been validated.¹²² There is no credible explanation of how face and body mappers are able to overcome the limited information available, uncertainties associated with lighting, camera angles, lenses and blurriness, or able to compare high quality still photographs of the accused with moving – often in a staccato-like manner – video images and stills of very low quality.¹²³ We have no meaningful information about the capabilities of facial mappers or their error rates. The few existing papers and chapters relating to facial (and body) mapping – and similar techniques under alternative names – tend to be speculative or propositional.¹²⁴ Curiously, this handful of papers and chapters, often published in marginal forensic science journals, non-peer reviewed texts and even law journals, is sometimes presented, and accepted, as authoritative support for facial mapping expertise or ‘specialised knowledge’.

Fourth, there is considerable *slippage* between training in a recognised field like anatomy (or physical anthropology) or having experience in forensic facial reconstruction (from decomposed remains) and an ability to make reliable comparisons between two-dimensional images. Training in anatomy and possession of a specialised vocabulary do not enable a person to make reliable identifications, to reliably identify similarities and differences, or to somehow personally transcend the many technical obstacles. The fact that we do not know if apparent similarities are real or artefacts of the angles, lighting, and lenses (or the contortion of the unknown person’s face and body or a disguise) only compounds the difficulties. Rather than scrutinise what it is in a person’s training, study or experience that allows them to give evidence as an exception to the opinion rule, judges tend to provide university-trained experts with considerable leeway. In consequence, we have highly credentialed experts giving

120 Geoffrey Bowker and Susan Star, *Sorting Things Out: Classification and Its Consequences* (1999).

121 Here, judges (as co-producers) are actively involved in the construction of a legally recognisable ‘field of expertise’ or type of ‘specialised knowledge’. Judges often underestimate their role in the social legitimisation of knowledge and expertise. See Sheila Jasanoff (ed), *States of Knowledge: The Co-Production of Science and Social Order* (2004).

122 Richard Kemp and Kevin Coulson, ‘Facial Mapping and Forensic Photographic Comparison: An International Survey of Experts’ (2008) (forthcoming); Maciej Henneberg, ‘Facial Mapping, Body Mapping and the Duties of an Expert Witness’ (Paper presented at the Public Defenders Annual Criminal Law Conference, Sydney, 3 March 2007) <http://www.lawlink.nsw.gov.au/lawlink/pdo/ll_pdo.nsf/pages/PDO_facialmapping> at 24 March 2008.

123 *R v Drollett* [2005] NSWCCA 356, [63] (Simpson J, McClellan CJ at CL and Rothman J concurring).

124 In *Tang* at [72], Spigelman CJ reproduced the bibliography from Dr Sutisno’s Report. The Report makes multiple references to a three-page article from a popular law journal, a short chapter from a forensic science text and a 20 year old article on photographic superimposition. The texts are: Michael Bromby, ‘At Face Value? The Use of Facial Mapping and CCTV Image Analysis for Identification’ (2003) 153 *New Law Journal: Expert Witness Supplement* 302; Mehmet Iscan, ‘Introduction of Techniques for Photographic Comparison: Potential and Problems’ in Mehmet Iscan and Richard Helmer (eds), *Forensic Analysis of the Skull* (1993) 57; George Maat, ‘The Positioning and Magnification of Faces and Skulls for Photographic Superimposition’ (1989) 41 *Forensic Science International* 225.

their personal opinions about the identification of offenders, which are not based on ‘specialised knowledge’ and have little, if any, foundation in their formal training.

Sometimes judges tend to elide or – as in *Jung* – trivialise the issue of ‘specialised knowledge’ by focusing primarily on whether the ‘opinion’ is linked to the ‘training, study or experience’.¹²⁵ We can represent this schematically:

training, study or experience → opinion

If, in this alternative approach to admissibility, the opinions of facial mappers are predicated upon ‘study’ (or studies), then we should be shown those studies and the results. To the extent that ‘experience’ is invoked, it tends to be of a very general nature – eg, 20 years as an anatomist examining thousands of bodies – or based on the provision of similar evidence in earlier cases. Experience as an anatomist does not make opinions about identification or descriptions of similarities and differences based on the comparison of images reliable. Similarly, the fact that an expert has previously testified in court, or that a similar technique has been admitted into another court, does not make untested techniques valid. This last point is particularly important. Judges should be very careful when citing earlier judgments, or admissibility decisions from foreign jurisdictions, to support the admission of expert opinion evidence that is not ‘knowledge’ or credibly referable to any ‘training, study or experience’.¹²⁶ A research degree in anatomy and experience as an anatomist does not overcome the problems with population databases, the comparison of images, and the absence of validation studies. Rather, these limitations are precisely why validation studies and caution are required. Indeed, this is one of the reasons why we have limits on opinion evidence and protections like section 137.

Finally, the fact that facial mappers, and other identification *experts* have not undertaken validation studies – like blind trials and accuracy testing – when these could readily be organised should induce anxiety. Because courts have been so ready to admit their opinion evidence, experts and investigators have not been obliged to demonstrate that they can actually do what they claim to be able to do. Courts try to manage some of the obvious difficulties – as with the limitations imposed by Spigelman CJ in *Tang* – but there is little evidence that apparent

125 See, eg, Gaudron J in *Veleviski v R* (2002) 187 ALR 233, [80].

126 NSW judges considering facial mapping evidence frequently refer to cases like *The Queen v Murdoch* (No 4) (2005) 195 FLR 421 (Martin CJ) where Dr Sutisno’s positive identification evidence was admitted. See also *Murdoch v The Queen* [2007] NTCCA 1. They also advert to English cases where facial mapping evidence is admissible as positive identification and police officers are allowed to make positive identifications on the basis of repeatedly watching CCTV recordings and looking at photographs. The leading English decision in this regard is *Attorney General’s Reference No 2 of 2002* [2002] EWCA Crim 2373. This legal authority has not all gone one way, however. Several English decisions, of which *R v Gray* [2003] EWCA Crim 1001, [16] is the prime example, have been highly disparaging of facial mapping evidence. In Australia, in the years before *Tang* and *Jung*, several judges of the District Court of NSW actively excluded facial mapping evidence. An instructive example is in *R v BLM* [2005] DC (Unreported, Blanch DCJ, 14 September 2005).

similarities drawn from security and surveillance images provide a sound basis for expert identification.¹²⁷

It would not be difficult to test, or dramatically improve, the processes employed by investigators when they seek the assistance of these experts.¹²⁸ If facial mappers are actually able to make reliable anthropometric and morphological comparisons, then there is no reason why the courts should not impose or require more rigorous procedures. It might prove interesting, for example, to ask two independent facial mappers (or anatomists) to each identify the facial (and/or body) features visible in a single set of images using a standardised set of parameters. One expert could be given images of the suspect and the other could be assigned images of the offender from the security recording. Third parties could then compare the descriptions provided through these independent assessments. If there were not detailed descriptions of similarities drawn from the two sets of images then – regardless of whether facial mapping techniques are capable of assisting identification, and regardless of the frequency of facial and body features among populations and sub-populations – the particular evidence would not be capable of supporting the prosecution case. To the extent that any description offers no more than vague or non-specific similarities, the opinion evidence would be of very limited probative value.¹²⁹ This kind of procedure would provide means to prevent experts offering biased – whether conscious or unconscious – and incriminating *ipse dixit*.¹³⁰ In the absence of validation and reliability studies this could provide a first order prophylactic. Even this, however, does not overcome the problem of whether apparent similarities are artefacts, coincidental or even meaningful.

Overall, we have no information about the accuracy of facial mapping evidence or techniques. All that we have are good reasons – some methodological, some procedural and some sociological – for approaching facial mapping evidence and even evidence of similarities and differences with considerable scepticism.

127 Vicki Bruce et al, 'Matching Identities of Familiar and Unfamiliar Faces Caught on CCTV Images' (2001) 7 *Journal of Experimental Psychology: Applied* 207. Many of the problems associated with reliably identifying unfamiliar people are not attributable to the quality of the images. There is empirical evidence that familiarity can make a difference to identification even where the quality of the images is poor. And, that people who are not familiar with persons of interest often make mistakes when the images are of high quality. Our jurisprudence, perhaps unsurprisingly, has allowed lay familiars to make identifications from security and surveillance images. Though judges have, it is fair to say, encountered difficulties explaining this approach, particularly around the fact/opinion dichotomy. See *R v Smith* (1999) NSWLR 419 (Sheller JA); *R v Smith* (2001) 206 CLR 650 (Kirby J); *R v Marsh* [2005] NSWCCA 331 (Studdert J, Kirby and Howie JJ concurring); *R v Drollett* [2005] NSWCCA 356 (Simpson J).

128 Changing the process may have an impact on the probative value of the evidence. Consider *Blick* [2000] NSWCCA 61, [25]; *R v Truong* (1996) 86 A Crim R 188, 195 (Miles CJ).

129 See, eg, *R v Dann* [2000] NSWCCA 185.

130 The close links between forensic scientists, police and the prosecution are frequently trivialised or overlooked. See Carol Jones, *Expert Witnesses: Science, Medicine and the Practice of Law* (1994). Interestingly, in civil proceedings and in relation to the evidence adduced by criminal defendants, judges are far more likely to express concerns about bias and partisanship. But see Gary Edmond, 'Judging Surveys: Experts, Empirical Evidence and Law Reform' (2005) 33 *Federal Law Review* 95.

Jung is informative because it concerns novel and controversial opinion evidence. The evidence was admitted notwithstanding the controversy, the centrality of identity, and the fact that identifications based on CCTV and security images are proliferating. The prevailing approach to sections 79 and 137 affords evidence from the state's investigative agencies too easy access to the courts.¹³¹ Moreover, *Jung* would seem to present the kind of situation, given the unknown probative value of the expert's opinion, where there is a very real risk of the jury overvaluing her evidence. Indeed, to the extent that reliability is not demonstrated, any reliance placed on the opinion evidence may not only be unfair but irrational. In what has come to be a rather ritualistic incantation, the trial judge in *Jung* expressed relief that rebuttal experts, cross-examination and appropriate directions on the law and the evidence would provide safeguards against 'illegitimate reasoning' and placing 'undue weight' on the state's inculpatory expert evidence. In the absence of evidence about the validity and accuracy of Dr Sutisno's techniques, this might be considered unduly optimistic.

In a system of justice that claims to take admissibility jurisprudence and the need for fair trials seriously, it might be thought that sections 79 and 137 would provide the trial judge with the means and inclination to exclude unreliable expert opinion evidence or expert opinion evidence of unknown reliability where there is a real danger than the jury might misuse or overvalue it. Recent cases, however, like *Jung* and *Tang*, suggest that judges do not always use these admissibility rules and protections in ways that might actually provide protection to the accused.¹³²

Now we move to consider some of the implications of this over-inclusive admissibility jurisprudence.

V PROCEDURAL AND SYSTEMIC PROBLEMS WITH EPISTEMOLOGICAL INDIFFERENCE

Expert opinion evidence of questionable reliability is regularly admitted in criminal trials. This subordinates concerns about accuracy and fairness to the primacy of the tribunal of fact. Instead, we should have a system requiring the state to demonstrate that experts can actually do what they claim when proffering their incriminating opinions. Our admissibility standards should be indexed to reliability. If the state cannot demonstrate that incriminating expert opinion evidence is based on techniques that are valid and reliable, then that evidence should not be presented to the jury. Just because the jury has constitutional primacy as the tribunal of fact does not mean that they should be arbiters of all expert opinion evidence and all conflicts between experts. In adversarial systems, the trial judge is in a reasonable position to determine whether the state's expert

131 Consider the apparent *visual verballing* disclosed in *R v Rix* [2005] NSWCCA 31.

132 See, eg, *R v Leung* (1999) 47 NSWLR 405 (Simpson J, Spigelman CJ and Sperling J concurring); *R v Li* (2003) 140 A Crim R 288 (Ipp JA, Whealy and Howie JJ concurring); *R v Madigan* [2005] NSWCCA 170, [107] (Wood CJ at CL, Grove and Hoebein JJ concurring); *R v Riscuta* [2003] NSWCCA 6 (Heydon JA, Hulme J and Carruthers AJ concurring); *R v Gao* [2003] NSWCCA 390, [23] (Greg James J, Sully and Adams JJ concurring).

opinion evidence is sufficiently reliable for admissibility as well as whether its probative value *actually* outweighs any real danger of unfair prejudice. Our trial judges should be more willing to scrutinise and exclude incriminating expert opinion evidence.¹³³

In this section it is my intention to raise a series of procedural, institutional and social problems caused or accentuated by epistemological indifference in order to reinforce the utility of a reliability threshold.

A Cross-Examination (on *ipse dixit*) and Rebuttal Experts

In the foregoing discussion we have encountered trial and appellate judges explaining the decision to admit the state's incriminating expert opinion evidence, drawing confidence from the availability of cross-examination and the possibility of rebuttal expertise.¹³⁴ As desirable as cross-examination may be, there is, unfortunately, limited empirical evidence of its efficacy.¹³⁵ There is no evidence that cross-examination by defence lawyers consistently or meaningfully exposes the very real limitations in some kinds of expert opinion evidence, or overcomes the very real dangers of unfair prejudice caused by admitting unreliable expert opinion evidence and expert opinion evidence of unknown reliability. Moreover, cross-examination and rebuttal expertise might not be particularly effective in overcoming some of the structural and institutional advantages accruing to the prosecution and its expert witnesses.

Although cross-examination is a much celebrated and useful forensic resource, its value in relation to expert opinion evidence, and particularly the opinion evidence of experienced forensic scientists called by the prosecution, might be practically limited. Part of the problem seems to be the actual organisation of the criminal proceedings and the *insulation* of the state's forensic scientists. Experts called by the state are typically presented, and present themselves, as independent and impartial experts. These forensic scientists usually have ways of insulating their performances and experience even where they work in close proximity with the police investigation and have access to inadmissible information and the personal suspicions of investigators. When it comes to defence and rebuttal experts, they are in a position which is structurally different and far more ambiguous.

These experts often appear as something resembling a necessary evil. Called to respond to the inculpatory expert evidence adduced by the prosecution, they are routinely characterised as partisan experts engaged by the defence to challenge what is presented as the painstaking and meticulous work undertaken by investigators working for the state. Rather than present opinions that are carefully integrated into a cohesive narrative, rebuttal experts tend to be critical. Their

133 Andrew Ashworth, 'Excluding Evidence As Protecting Rights' [1977] *Criminal Law Review* 723.

134 See also *The Queen v Murdoch (No 4)* (2005) 195 FLR 421, [91], [95] (Martin CJ).

135 Cross-examination remains a vital forensic tool. There are, however, questions about its effectiveness and consistency, especially against experienced witnesses. See generally Sheila Jasanoff, *Science at the Bar: Law, Science and Technology in America* (1995); Michael Lynch and David Bogen, *The Spectacle of History: Speech, Text and Memory at the Iran-Contra Hearings* (1996); Gary Edmond, 'Science in Court' (1998) 20 *Sydney Law Review* 361, 391–400.

criticisms and their emphasis on protocols, good practice and uncertainties in the evidence adduced by the prosecution often appear pedantic or impractical and render them vulnerable to marginalisation as *ivory tower*-types – that is, remote from the real-world exigencies of forensic scientific practice and investigations.¹³⁶ There are no guarantees that an independent expert called by the defence will perform as well as the forensic scientist(s) called by the prosecution, or appear as impartial or credible regardless of the quality of their evidence (or critique).

Just because cross-examination *might* be able to expose problems with a technique or an expert *might* be found who is willing to testify about weaknesses in the forensic scientific evidence, does not mean that these should provide the primary bulwark against unreliable and prejudicial expert opinion evidence adduced by the prosecution. Facilitating cross-examination or allowing the defence to call rebuttal expertise does not make a trial fair. Structural *symmetry* is not the same as substantial fairness.¹³⁷

One of the difficulties for a defendant confronted with incriminating expert opinion evidence is the need to demonstrate unreliability and/or discredit the forensic scientists called by the prosecution. This difficulty is accentuated by the fact that these efforts take place in a setting where the reliability of the expert opinion evidence adduced by the prosecution is implied. Implicitly, experts called by the prosecution have the imprimatur of their scientific institution, the prosecution and the court. Rather than require the expert opinion evidence to earn this imprimatur, impecunious defendants are individually expected to demonstrate unreliability (and engender incredulity). In effect, the defence is required to show that the state's experts are mistaken, incompetent or fraudulent. The practical need to demonstrate unreliability or incompetence places the defence at a serious disadvantage because even powerful and – what in other settings might be – unassailable critiques can be parried on the basis that they are motivated or the particular challenge goes only to weight – an issue for the jury – rather than admissibility – an issue for the trial judge. Judges can manage their own performance by recognising limitations with the expert opinion evidence and even expressing personal doubts about its reliability and probative value while admitting the evidence and deferring to the tribunal of fact.

Where experts are allowed to present opinions, which are not grounded in their actual training or experience and are not shown to be based on reliable techniques, it can be quite difficult to challenge their evidence. Even though the kinds of problems raised by Porter and Dr Doran as well as those developed in

136 This was how the prosecution challenged the evidence of Professor Barry Boettcher, acting as a rebuttal expert against the blood evidence, at the Chamberlain trial and appeal. See Gary Edmond, 'Law, Science and Narrative: Helping the "Facts" to Speak For Themselves' (1999) 23 *Southern Illinois University Law Journal* 555. See also *Bropho v The State of Australia* [2007] WADC 77 (Goetze DCJ).

137 Sometimes procedures are not even formally symmetrical. There have been instances where an investigator who had listened repeatedly to surveillance recordings was allowed to give ad hoc expert opinion evidence about the identity of the speaker and, remarkably, in the same case an apparently qualified expert, called by the defence, was prevented from testifying, and even from criticising the methods used by the lay investigator, because she had not personally analysed the recordings. See *R v Madigan* [2005] NSWCCA 170, [107] (Wood CJ at CL, Grove and Hoeben JJ concurring).

Part IVA ('Discussion') can be developed through cross-examination and the testimony of rebuttal experts, it can be very difficult to effectively counter the opinion of an expert when it is linked to his or her long and often diffuse experience or based around some ephemeral exposure during his or her formal training. We have already encountered judicial analysis of facial mapping evidence. Why should a juror be particularly concerned about the intricacies of the distribution of facial features when the prosecution witness has a PhD in anatomy, a position at a research-intensive university, claims to be able to transcend problems with angles, lenses and lighting through years of experience, and is obviously satisfied that there are no differences between the accused and the person in the incriminating images? These kinds of predicates tell us little, if anything, about reliability. Yet they seem to make claims about similarity and difference (as well as identity) highly persuasive for a variety of audiences. In order to shake the opinion of an expert called by the prosecution – particularly where the evidence is essentially *ipse dixit* – it is usually incumbent on the defence to challenge the expert's experience or credibility as well as the reliability of their techniques. There are few guarantees that even competent counsel will be able to usefully develop and convey the many serious problems with some kinds of opinion evidence and, in the absence of previous forensic misadventure, it may be difficult to make credibility assaults look like anything more than defence counsel going through the motions. The unreliability of the evidence and the resulting unfair prejudice is not repaired, and may not even be adequately addressed, by (the possibility of) skilled cross-examination or expert disagreement.

Placing faith in the efficacy of cross-examination and rebuttal experts also imposes considerable resource burdens on the defence. Where a specific type of expert evidence or opinion – like the testimony of facial mappers – re-appears, sundry defence lawyers are required to challenge the admissibility of the opinions and then the reliability of the opinions in each proceeding. The unified prosecution, the police and the expert(s), in contrast, learn from each forensic encounter.¹³⁸ Rather than impose a reliability threshold as a prophylactic, admitting unreliable expert opinion evidence transfers responsibility for demonstrating *unreliability* onto the defence. It shifts costs from the prosecution and investigators onto defendants and legal aid. The upshot is that the defence is required to challenge the prosecution's expert opinion evidence in a setting where there is implicit support for its value. This represents a very inefficient use of the state's limited legal aid budget – as well as court time – and produces inconsistencies between trials as lawyers with quite different skills, abilities and predilections, as well as clients with varying resources, differentially expose limitations with the state's incriminating expert opinion evidence. These issues reinforce the desirability of requiring the state to demonstrate the reliability of its techniques rather than leaving it to the vagaries of individual criminal trials.

138 Stefan Timmermans, *Postmortem: How Medical Examiners Explain Suspicious Deaths* (2006); Jones, above n 130.

Cross-examination and defence experts provide an appealing, if deceptive, symmetry and a useful rhetorical resource to boost claims about the fairness of criminal processes. There is, however, little evidence that they can consistently and effectively expose limitations in incriminating expert opinion evidence or facilitate a balanced presentation of expert disagreement.

B Abandoning Expert Conflict to the Jury

Whenever two experts disagree on an issue there is a tendency to treat it as a genuine conflict of expert opinion that is best resolved by a jury.¹³⁹ Judges, as the decisions in *Tang* and *Jung* illustrate, have a tendency to conceive almost all expert disagreements, including circumstances where defence experts are challenging the very existence of ‘specialised knowledge’ or raising fundamental criticisms about techniques and capabilities, as conflicts that should be resolved by a jury.¹⁴⁰

Classifying expert disagreement in this way relieves judges of the need to take professional responsibility for admissibility determinations and the exercise of discretions. Overly-inclusive approaches to admissibility can actually create controversy and disagreement. Judges seem to underestimate how expert evidence jurisprudence may actually contribute to controversy in the courtroom as well as the need for the defence to secure critical expert opinions. The failure to exclude unreliable expert opinion evidence actually creates the need for cross-examination and rebuttal experts.¹⁴¹ Ironically, the possibility of cross-examination and the possible participation of rebuttal experts – rather than any evidence about their effectiveness – allows judges to express confidence in the fairness of the trial.

Even if we accept that lay juries are capable of handling expert disagreements, there are good reasons not to place all disagreements before them. Sections 79

139 *R v Duke* [1979] 22 SASR 46, 47–8 (King CJ) provides a typical approach to this division and reinforces the widespread confidence in the jury’s capabilities. Perhaps the most insightful commentary in this regard was written by Dixon J in *Sinclair v R* (1946) 73 CLR 316, 337–8. This approach has a long pedigree, as some influential critics attest. See Learned Hand, ‘Historical and Practical Considerations Regarding Expert Testimony’ (1901) 15 *Harvard Law Review* 40.

140 This essay should not be understood as a critique of jury competence. Problems with expert opinion evidence are not always lost on the jury. During the trial in *Tang* the jury asked a particularly encouraging question. In relation to Dr Sutisno’s facial and body mapping analysis they inquired:

Was there any photo anthropometry performed in comparing the surveillance images and forensic photos of the accused, what were the results? How accurate is morphology analysis as a technique? What percentage of cases are correct matches of persons versus incorrect matches? Could we please ask Dr Sutisno how many matching morphological features she needs to form the opinion that two photos are the same person, what would be the minimum?

In response, Dr Sutisno was permitted to make the following statement before the jury:

From my own experience and from what I’ve studied I believe it can be fairly accurate in narrowing it down to whether or not they are one and the same person. ... Well, enough to be able to provide in results that would be sufficient to show that there are matches or no matches. Basically to show that they are of the same person or disprove it altogether and say that they’re not of the same person.

Then, during cross-examination, ‘Dr Sutisno rejected the suggestion that there was a degree of subjectivity in her assessment’. This response to a jury question about the accuracy of the technique does not get far beyond subjective impressions. See *Tang* (2006) 65 NSWLR 681, [74]–[78] (emphasis added).

141 Jasanoff, above n 121.

and 137 should have a more exclusionary role. If the prosecution cannot show that its expert opinion evidence is reliable then a judge should exclude it. Lay jurors should not be obliged to assess assertion, speculation or even sincere educated guesses because judges have become indifferent to the reliability and probative value of expert opinion evidence adduced by the prosecution.¹⁴²

C Form Over Substance: Judicial Directions and Warnings

Without wanting to devote too much space to directions and warnings, what we can say by way of summary is that almost all of the available research suggests that judicial instructions are difficult to follow and probably ineffective.¹⁴³ Continuing confidence in directions and warnings seems to be misguided.¹⁴⁴ Trial and appellate judges concerned with securing formally fair trials have a tendency to exaggerate the value of instructions, directions and warnings. To admit the prosecution's unreliable expert evidence or expert opinion evidence of unknown reliability on the basis that cross-examination (and possibly defence experts) will expose limitations, and that instructions will clearly convey how the jury should approach expert opinion evidence and expert disagreement, might be thought, at this stage of the day, to be incredibly optimistic.

Prevailing approaches have a tendency to shield judges, especially appellate courts, from some of the regulatory implications of their jurisprudence and decision making. In effect, by investing too much confidence in cross-examination, rebuttal experts, directions and warnings, judges bestow inappropriate advantages on investigators, prosecutors and forensic scientists.

142 It might be argued that limiting the information available to the jury is inconsistent with the (commitment to) gradual liberalisation in evidence law, sometimes described as *free proof*. See, eg, Mirjan Damaska, *Evidence Law Adrift* (1997); Dennis Galligan, 'More Scepticism About Scepticism' (1988) 8 *Oxford Journal of Legal Studies* 249, 255–8. While I am not necessarily opposed to this general trajectory, there are some important qualifications. In recent years, in jurisdictions with civil juries, there has not been a more inclusive response to expert opinion evidence adduced by plaintiffs. Perhaps more substantially, free proof does not entitle a party to produce irrelevant evidence and, as this essay contends, should not entitle the state to adduce expert evidence that is probably unreliable. Commitment to a more *laissez-faire* approach to admissibility should not trivialise the implications for the defence or the expectations imposed on lay juries. In adversarial trials, where the parties are responsible for adducing evidence, there are always – and always will be – practical and strategic constraints on what a jury sees and hears. Free proof should not be simply equated with better proof or more accurate fact-finding. Moreover, it may be that the original commitment to free proof was based on a more limited sociological appreciation of the practical constraints on adversarial criminal trials. These constraints may have been exacerbated by changes to the respective resources available to the prosecution, investigative agencies and the defence. Recent resourcing trends seem to have made more permissive approaches to incriminating expert evidence less desirable.

143 See Joel Lieberman and Bruce Sales, 'The Effectiveness of Jury Instructions' in Walter Abbott and John Batt (eds), *A Handbook of Jury Research* (1999) 18.1; James Ogloff and Gordon Rose, 'The Comprehension of Judicial Instructions' in Neil Brewer and Kipling Williams (eds), *Psychology and Law: An Empirical Perspective* (2005) 407.

144 Australian judges, on average, have been highly resistant to methodologically credible studies of jurors and jury decision-making as well as studies of mock jurors. They have also been remarkably resistant to facilitating empirical studies of their courts. Even the current reference on 'Jury directions in criminal trials' by the NSWLRC is not undertaking any primary empirical studies of juror responses during actual trials.

Standards that take into account the reliability of expert evidence, in contrast, require state and investigative agencies to actually test the assumptions, techniques and opinions routinely relied upon to establish guilt. They also raise the quality of techniques and evidence used in investigations and prosecutions.

There will be no loss of legal legitimacy if judges exclude incriminating expert opinion evidence that is either unreliable or of unknown reliability. Almost all the institutional risks run the other way.

D Necessity (and Centrality) of the Evidence

One of the weakest justifications for admitting expert opinion evidence is *necessity*. In the cases considered above we can identify judicial appeals to the necessity of the expert opinion evidence adduced by the prosecution. In *Tang*, Spigelman CJ explained that the poor quality of the security videos was such that the jury needed the help of an expert: ‘... the quality of the photographs derived from the videotape was such that the comparison of those stills with the photographs of the Appellant could not be left for the jury to undertake for itself.’¹⁴⁵ In *Jung*, the jury were said to ‘require the assistance of an expert such as Dr Sutisno’.¹⁴⁶

Admitting expert opinion evidence on the grounds of necessity, or the centrality of the identification evidence, tends to displace (and trivialise) issues of reliability. If the techniques and assumptions underpinning the expert’s opinion are not demonstrably reliable, then in no sense can the evidence be necessary or central.

E Separating the Issue of ‘Reliability’ from the Other Incriminating Evidence

Our judges have a tendency to rationalise the admission of evidence on the basis of its relationship with other evidence. When it comes to expert opinion evidence, however, this places the cart before the horse. When determining the value of expert opinion evidence it is necessary to ascertain the reliability of the technique(s) before the opinions are buttressed with other evidence. The fact that there is other admissible, and even compelling, evidence pointing to the guilt of the accused does not necessarily add to the reliability of the expert opinion evidence. Assessment of the reliability of expert opinion evidence should be *independent* of the other evidence – however incriminating. The overall strength of a case tells us nothing about the validity or reliability of the techniques relied upon by forensic scientists and police.

Independence operates at yet another level. Often expert opinion evidence is presented by the prosecution as independent, and therefore corroborative, evidence. In practice, however, this opinion evidence is not always genuinely independent. In *Tang*, for example, the police had located the accused’s

¹⁴⁵ *Tang* (2006) 65 NSWLR 681, [120].

¹⁴⁶ *Jung* [2006] NSWSC 658, [45].

fingerprint on a packet of cigarettes taken during the robbery.¹⁴⁷ The facial mapping expert seems to have known about the existence of the fingerprint as well as the identity (and Asian appearance) of two of the three persons of interest appearing in the security video based on admissions. For the purposes of the investigation she was asked whether Tang was the *unknown* offender. To characterise her evidence as independent support may be deeply misleading. Knowledge of the fingerprint, and even the implication that Tang was the prime police suspect must, in the absence of information about the reliability of Dr Sutisno's identification techniques, make her opinions unsettling. It also makes claims about the strength of the circumstantial case against the accused much weaker than the way it was presented to the jury.¹⁴⁸ (This last issue highlights one problem with over-reliance on instructions.)

By focusing on the reliability of expert opinion evidence judges could untangle some of these interactions and apparent synergies. A circumstantial case should not be used to augment expert opinion evidence that is either unreliable or of unknown reliability.

F 'Criteria enabling evaluation of ... validity'

It is important to stress, in relation to an expert witness' methodology, that following or comprehending the underlying process should not be equated with reliability or even an adequate *basis*. In *Makita*, Heydon JA described the 'prime duty of experts' as the need to 'furnish the trier of fact with criteria enabling evaluation of the validity of the expert's conclusions'.¹⁴⁹ Regrettably, this did not lead to an assessment of reliability, or even technical consideration of validity, in *Makita*, *Tang* or *Jung*.

A clear example of this difficulty arises from Hall J's remarks in *Jung*:

... the evidence on the voir dire does not establish that she has failed to disclose the factual material she has utilised (the photographic images), the nature of the methodology that she has employed and the type of analysis described in her reports (morphological analysis). ... Dr Sutisno's evidence complies with the requirements for admissibility.¹⁵⁰

Here, Hall J substitutes some of the concerns from *Makita* and *Tang*, notably the ability to follow the expert's reasoning, for evidence of the opinion having an adequate basis for admissibility. The ability to describe a process does not, however, make it reliable, especially where the process has been subjected to destabilising and unanswered methodological criticisms. In this example, knowing how Dr Sutisno went about comparing the images does not make the process valid or reliable. Without information about the frequency and distribution of attributes, or formalised means of overcoming the many distortions, limitations and selection biases and their effects, the ability to understand the process used by the witness may be of very limited value. This is

147 *Tang* (2006) 65 NSWLR 681, [99]. *Tang* offered an explanation for the fingerprint that, if accepted, might have provided a relatively innocent explanation.

148 *Ibid* [157].

149 *Makita* (2001) 52 NSWLR 705, [59].

150 *Jung* [2006] NSWSC 658, [62]. See also *R v Li* (2003) 140 A Crim R 288, [84]–[85].

another reason why reliability is such an important consideration and more fundamental than the ability to understand the process or underlying reasoning relied upon by the witness.

G Eliminating Ad Hoc ‘Expertise’

In passing, it seems appropriate to make a few remarks about the expansion of common law exceptions to the opinion rule *after* the introduction of the *Evidence Act*. Even though section 79 would seem to cover the field, since 1995 judges in NSW have expanded the concept of the ad hoc expert (also known as the expert ad hoc).¹⁵¹ The concept was originally imported from New Zealand and applied in circumstances where it enabled investigators and interpreters to make transcripts of covert voice recordings as an *aid* for the jury.¹⁵² More recently, judges have expanded it to enable investigators and interpreters to make positive identifications based on repeated exposure to sound recordings – sometimes in circumstances where different languages are spoken in the samples.¹⁵³ These developments are curious given that the *Evidence Act* explicitly allows for transcripts prepared from recordings to be used as evidence of a tape’s contents.¹⁵⁴ Nevertheless, as *Tang* illustrates, the concept of the ad hoc expert can be employed to allow *experts* without ‘specialised knowledge’ to proffer incriminating opinion evidence about the identity of a person associated with a criminal act.¹⁵⁵

Significantly, recourse to ad hoc expertise is only ever used to admit the prosecution’s evidence – usually where the witness is not appropriately qualified to give an opinion or the opinion is not based on ‘specialised knowledge’.¹⁵⁶ Perhaps predictably, where judges create or feel compelled to develop exceptions to the opinion rule that are not to be found in the *Evidence Act*, they never apply sections 137 and 135 to address the very real danger of unfair prejudice to the accused.

Recourse to ad hoc expertise should stop and the category should be abandoned. Judges should not be creating exceptions to the opinion rule or expanding the common law to enable the prosecution to adduce unreliable opinion evidence in circumstances not permitted by section 79.

151 For a more detailed discussion see Gary Edmond and Mehera San Roque, ‘Petit Truths: Ad Hoc Experts and Legal Legitimacy’ (Paper presented at the University of Melbourne Law & Society Conference, Melbourne, 29 November 2007).

152 *R v Menzies* [1982] 1 NZLR 40; *Butera v Director of Public Prosecutions (Vic)* (1987) 164 CLR 180.

153 See, eg, *R v Leung* (1999) 47 NSWLR 405; *R v Li* (2003) 140 A Crim R 288; *R v Madigan* [2005] NSWCCA 170; *Korgbara v The Queen* (2007) 170 A Crim R 568 (McColl JA, James J concurring).

154 *Evidence Act 1995* (NSW) s 48. See *R v Cassar* [1999] NSWSC 436 (Sperling J).

155 Ad hoc expertise is often recognised by courts where the relevant established fields and disciplines have been unable to develop reliable techniques such as in voice analysis or computer-assisted identification. See Roderick Munday, ‘Videotape Evidence and the Advent of the Expert Ad Hoc’ (1995) 159 *Justice of the Peace* 547.

156 Ad hoc experts are not required to abide by the rules and codes governing the performance of expert witnesses, used especially in civil litigation. Consider the *Uniform Civil Procedure Rules 2005* (NSW) sch 7.

H Accuracy of Verdicts and De Novo Review

One benefit arising from an emphasis on reliability is that in principle it facilitates the *de novo* review of admissibility determinations and the application of exclusionary discretions by appellate courts.¹⁵⁷ Historically, appellate courts have been quick to defer to the many advantages accruing to the trial judge. On the admissibility and particularly the reliability of expert evidence, however, there would seem to be few reasons for such deference. An expert's demeanour and performance, or (favourable) experience in previous proceedings, should be granted little, if any, significance. Trial judges and appellate courts should be looking for evidence of testing, validation and accuracy rates, as well as germane studies published in authoritative literatures. The burden of demonstrating reliability should lie with the prosecution. If dissatisfied in terms of reliability or fairness, trial judges and appellate courts should be willing to exclude the prosecution's expert opinion evidence. These assessments are not dependent on any special advantages available at the trial.¹⁵⁸

I Inconsistency Between Civil and Criminal Justice

Another illuminating feature of this review is what it suggests about judicial ideologies. The cases seem to invert the kinds of approaches to expertise that we might expect from our criminal and civil justice systems.

In *Makita*, Heydon JA asked many questions of Associate Professor Morton's opinions and suggested that it is 'necessary to examine Professor Morton's report closely'.¹⁵⁹ For example:

... can it be said that Professor Morton's report goes beyond a series of ocular pronouncements? Does it usurp the function of the trier of fact? More vitally, did it furnish the trial judge with the necessary scientific criteria for testing the accuracy of its conclusions? Did it enable him to form his own independent judgment by applying the criteria furnished to the facts proved? Was it intelligible, convincing and tested? Did it go beyond a bare ipse dixit? Did it contain within itself material which could have convinced the trial judge of its fundamental soundness?¹⁶⁰

Justice of Appeal Heydon inquired about the level of acceptance behind Professor Morton's conclusions, whether certain aspects of his evidence might require even more testing, and whether his opinion usurped the 'function of the trier of fact'.¹⁶¹ Part of this extract is duplicated in *Tang*, but the stringency from the civil appeal is not reproduced in the criminal context. Interestingly, the last

157 Consider *Re Truscott* [2007] ONCA 575, [95]: 'the rules of evidence governing the admission of evidence in criminal proceedings are shaped primarily to facilitate the search for the truth. That search is not less important and no different when considering the admissibility of evidence offered on appeal.'

158 There may be some minor advantages, such as enabling the trial judge to observe experts interact on a concurrent evidence panel, but this can never be enough to overcome the failure to test and validate a technique or provide other indicia of reliability. Consider, eg, *Dobler v Halverson* [2007] NSWCA 335, [51] (Giles JA, Ipp and Basten JJA concurring); Gary Edmond, 'Secrets of the "Hot Tub": Expert Witnesses, Concurrent Evidence and Judge-Led Law Reform in Australia' (2008) 25 *Civil Justice Quarterly* 51–82.

159 *Makita* (2001) 52 NSWLR 705, [58].

160 *Ibid* [87].

161 *Ibid*.

sentence, concerned with ‘fundamental soundness’, was omitted and questions about ‘testing’ seem to have been practically elided.

While the point is not to endorse Associate Professor Morton’s evidence, or to question the New South Wales Court of Appeal’s finding, it is interesting to note that he had undertaken standardised testing and engaged, though perhaps not entirely successfully, with the existing Australian standards. Nevertheless, Associate Professor Morton’s opinion was rejected because his reasoning was characterised as unclear and overcome by judicial ‘common sense’. Dr Sutisno, in contrast, was not part of an established facial mapping fraternity, had not publicly validated her techniques, and yet her evidence was admitted and relied upon in a series of criminal prosecutions.

Whereas we might expect judges to impose a high standard on the state in criminal proceedings and be more sympathetic to the plight of individual plaintiffs in civil proceedings, we encounter in these decisions something approximating the opposite result. These cases and the jurisprudence around expert opinion evidence suggest a disturbing inversion of principle.¹⁶²

VI CONCLUSION: FROM HERE TO *RELIABILITY*?

In concluding, one of the first things to wonder about is the parochial aversion to *reliability*. The idea of *reliability* has not been central to common law jurisprudence associated with expert opinion evidence or opinion based upon ‘specialised knowledge’. We can perhaps appreciate that many judges are not skilled or experienced in assessing the reliability of different forms of expertise. Reticence and the difficulty of the task should not, however, deflect them from this important civic responsibility. If our criminal justice system aspires to be rational and fair, and not to convict the innocent, then reliability needs to take a more fundamental position in our jurisprudence.

The word ‘reliability’ does not feature in ss 79, 135 or 137 of the *Evidence Act*. Notwithstanding this omission, there have been a few occasions where senior members of our judiciary have intimated that section 79, and its common law equivalents, might actually require some assessment of *reliability*. Several High Court judges have explicitly recognised such a need. Part of the South Australian common law, requiring ‘a body of knowledge or experience which is sufficiently organised or recognised to be accepted as a reliable body of knowledge or experience’ was endorsed by Gaudron and Gummow JJ in *Osland v The Queen*.¹⁶³ Subsequently, in *HG*, Gaudron J indicated that this passage may have relevance to section 79:

The position at common law is that, if relevant, expert or opinion evidence is admissible with respect to matters about which ordinary persons are unable “to form a sound judgment ... without the assistance of [those] possessing special knowledge or experience ... *which is sufficiently organised or recognised to be accepted as a reliable body of knowledge or experience*”. There is no reason to

¹⁶² This inversion cannot be adequately explained by the participation of the jury in criminal proceedings.

¹⁶³ *Osland v The Queen* (1998) 197 CLR 316, [53] (Gaudron and Gummow JJ), [164]–[169] (Kirby J). The case cited is *R v Bonython* (1984) 38 SASR 45 at 46–7 (King CJ).

think that the expression “specialised knowledge” gives rise to a test which is in any respect narrower or more restrictive than the position at common law.¹⁶⁴

That approach was reiterated in *Velevski v The Queen*.¹⁶⁵ References to reliability can also be found in a number of prominent common law cases from State and Territory courts of appeal. These include *R v Gilmore*, *R v Carroll*, *Lewis v The Queen*, as well as *R v Bonython*.¹⁶⁶

At common law, the need for reliable expert opinion evidence, particularly forensic scientific evidence, has occasionally been quite prominent. This has been most conspicuous in the aftermath of notorious miscarriages of justice.¹⁶⁷ The appeal in *Lewis*, delivered in the wash-up of the Morling Royal Commission into the Chamberlain convictions, affords a useful illustration of the potential of a reliability threshold. The trial actually featured the testimony of some of the expert odontologists involved in the Chamberlain prosecution. Like facial mappers, these experts made positive identifications on the basis of quite limited information and non-standardised techniques. The *Lewis* decision is important because it provides a considered response to incriminating expert opinion evidence, in terms reminiscent of *Makita*, but also takes the issue of reliability seriously.

In *Lewis*, Maurice J explained that there was a need for the ‘Crown ... to carefully lay the ground for the reception of the opinions expressed by [the dental experts]. *It could only do this by proving the scientific reliability of the exercise they carried out*’.¹⁶⁸ Instead, at trial the Crown ‘chose to rely on the witness’ qualifications and experience in the field of forensic dentistry generally, and, in particular, upon the impressive curriculum vitae’.¹⁶⁹ Significantly, ‘[n]o experimental research was pointed to nor other studies which demonstrated the reliability of the deduction made by these witnesses’.¹⁷⁰ Justice Maurice continued:

Forensic evidence, especially if it goes to a vital issue implicating an accused person in the commission of an offence, may often have a prejudicial effect on the minds of a jury which far outweighs its probative value. ...

For my part I think that whenever the Crown wishes to rely upon forensic evidence the prosecutor has a clear duty, not just to his client, the Crown, but to the trial

164 *HG* (1999) 197 CLR 414, [58].

165 *Velevski v R* (2002) 187 ALR 233, [82] (Gaudron J).

166 *R v Gilmore* [1977] 2 NSWLR 935 (Street CJ, Lee and Ash JJ concurring); *R v Carroll* (1985) 19 A Crim R 410 (Kneipp J, Campbell CJ concurring); *Lewis v The Queen* (1987) 88 FLR 104 (Maurice J and Muirhead AJ) (‘*Lewis*’); *R v Bonython* (1984) 38 SASR 45 (King CJ, Bollen J concurring).

167 Prominent Australian examples include Lindy and Michael Chamberlain (the Morling Royal Commission) and Edward Charles Splatt (the Shannon Report). See also Clive Walker and Keir Starmer (eds), *Miscarriages of Justice* (1999); Richard Nobles and David Schiff, *Understanding Miscarriages of Justice* (2000); Barry Scheck, Peter Neufeld and Jim Dwyer, *Actual Innocence* (2000); Gary Edmond, ‘Misunderstanding the Uses of Scientific Evidence in High Profile Criminal Appeals: The Social Construction of Miscarriages of Justice’ (2002) 22 *Oxford Journal of Legal Studies* 53. In the civil sphere, judges and legislators have developed different ways to manage expert evidence, particularly the socio-legal problems attributed to expert opinion evidence adduced by plaintiffs. Cf section 50 of the *Civil Liability Act 2002* (NSW).

168 *Lewis* (1987) 88 FLR 104, 122 (emphasis added).

169 *Ibid*.

170 *Ibid* 122–3.

judge and the jury to acquaint them, in ordinary language, through the evidence he leads, with those aspects of the expert's discipline and methods necessary to put them in a position to make some sort of evaluation of the opinions he expresses. Where the evidence is of a comparatively novel kind, the duty resting on the Crown is even higher: *it should demonstrate its scientific reliability*. It is not an answer to considerations that dictate these things be done to say the defence may draw it out in cross-examination; that is an abdication of the Crown's primary function in a criminal prosecution.¹⁷¹

The *Lewis* decision makes explicit the need for the Crown to 'demonstrate' the 'scientific reliability' of novel expert opinion evidence.¹⁷² Surrogates, such as credentials, confidence and experience in court, cannot substitute for studies demonstrating the reliability of the techniques or the process of deduction. Similarly, leaving the issue for weight and cross-examination represents 'an abdication of the Crown's primary function'.

Another important and illuminating set of examples concerned with *reliability* are the common law decisions structuring the admissibility of DNA profiling evidence. Most of these refractory decisions revolved around novel techniques and the development of population statistics associated with early DNA testing. These decisions have obvious parallels with identification and comparison evidence associated with facial mapping, particularly with respect to novelty, issues of standardisation, the meaning of population statistics (ie, distribution), and the rapid proliferation of the technology. In cases such as *R v Elliott*, *R v Tran*, *R v Lucas*, *R v Green* and *R v Pantoja*, senior judges excluded the prosecution's expert evidence because they were not satisfied with its reliability or probative value.¹⁷³ These judges, interestingly, did not refer to the common law *tradition* of taking the evidence 'at its highest' and did not, on principle, accept the evidence as reliable simply because a jury might.¹⁷⁴ Similarly, the possibility of cross-examination, rebuttal experts, and directions did not prevent the exclusion of incriminating expert evidence even in circumstances where the overall case was considered compelling. Rather, these judges deemed expert evidence adduced by the prosecution inadmissible because they were concerned about its reliability, the capability of the jury to meaningfully assess it and the danger of unfair prejudice. This cautionary approach, particularly associated with novel proffers of expert opinion evidence, seems to have been abandoned in recent years. This is unfortunate because the exclusion of incriminating expert opinion evidence – particularly in the United States and England, but also in Australia – contributed to the refinement of DNA typing techniques, stimulated

171 Ibid 123–4 (emphasis added). This passage was endorsed in *R v Tran* (1990) 50 A Crim R 233, 242 (McInerney J).

172 Though, in principle, there is no reason why the obligation to demonstrate 'scientific reliability' should not be ongoing. If fresh doubts about a technique or approach emerge then the prosecution should be obliged to dispel them.

173 *R v Elliott* [1990] NSWSC (Unreported, Hunt J, 6 April 1990) 10; *R v Tran* (1990) 50 A Crim R 233, 242 (McInerney J); *R v Lucas* [1992] 2 VR 109 (Hampel J); *R v Green* [1993] NSWCCA (Unreported, Gleeson CJ, Cripps JA and Abadee J, 26 March 1993) 10 (Cripps JA, Gleeson CJ concurring); *R v Pantoja* (1996) 88 A Crim R 554 (Hunt CJ at CL, Hidden J concurring).

174 There is some doubt, as the DNA cases indicate, about the strength of this 'tradition'. See Eric Hobsbawm and Terence Ranger (eds), *The Invention of Tradition* (1992).

research into population genetics, and improved the way the evidence was presented.¹⁷⁵ Requiring the state to demonstrate reliability can actually change institutional behaviour, encourage further study, and improve the reliability and probative value of techniques and opinions.

Finally, without wanting to suggest that approaches from the United States and Canada provide simple solutions to our problems, it is worth noting that both US and Canadian Supreme Courts require parties to demonstrate that their expert opinion evidence is reliable.¹⁷⁶ In *Daubert*, a majority of the Supreme Court of the United States explained that Rule 702 of the *Federal Rules of Evidence* (1975) – which has an obvious resemblance to our section 79 – imposes a ‘relevance and reliability’ approach toward admissibility. Even before statutory revision made the need for reliability explicit, the Supreme Court read ‘reliability’ into the phrase ‘specialized knowledge’.¹⁷⁷ More recently, Canadian courts, notably the Supreme Court and Ontario Court of Criminal Appeal, have become increasingly insistent on the need to demonstrate the ‘reliability’ of incriminating expert opinion evidence. The exclusionary discretions provided a basis for these developments at least since the Supreme Court handed down *R v Mohan*. More recent decisions, such as *R v DD*, *R v J-LJ* and *Re Truscott*, have brought the need for reliable expert opinion evidence to centre stage.¹⁷⁸ This trend is perhaps most clearly expressed in a recent decision by Deschamp J. Writing for the majority in *R v Trochym*, her Honour explained:

Reliability is an essential component of admissibility. Whereas the degree of reliability required by courts may vary depending on the circumstances, evidence that is not *sufficiently reliable* is likely to undermine the fundamental fairness of the criminal process.¹⁷⁹

175 More recent admissibility decisions, where the DNA profiling evidence is considered admissible, tend to stress the validation or reliability of the techniques. In *R v Gallagher* (2001) NSWSC 462, [36], [62], [72], [114], [140], Barr J ruled that DNA typing was admissible under section 79 on the basis that the system was properly validated, reliable and accurate. See also *R v Karger* (2001) SASR 1. Cf *The Queen v Hoey* [2007] NICC 49, [64]–[65].

176 In earlier papers I have not encouraged the appropriation of the *Daubert* criteria or the ethos underlying *Daubert* and its progeny. The emphasis here is on the importance of reliability per se, without committing to a particular model of science or expertise. Consider Gary Edmond and David Mercer, ‘Recognising *Daubert*: What Judges Should Know About Falsificationism’ (1996) 5 *Expert Evidence* 29; Gary Edmond, *Pathological Science? Demonstrable Reliability and Expert Forensic Pathology Evidence* (2007) Inquiry Into Pediatric Forensic Pathology in Ontario <http://www.goudgeinquiry.ca/policy_research/pdf/Edmond_Paper.pdf> at 24 March 2008.

177 Rule 702 of the *Federal Rules of Evidence* (1975) originally stated: ‘If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise.’ It was amended in 2000 to read:

If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise, if (1) the testimony is based upon sufficient facts or data, (2) the testimony is the product of reliable principles and methods, and (3) the witness has applied the principles and methods reliably to the facts of the case.

178 *R v Mohan* [1994] 2 SCR 9; *R v DD* [2000] 2 SCR 275; *R v J-LJ* [2000] 2 SCR 600; *Re Truscott* [2007] ONCA 575.

179 *R v Trochym* [2007] 1 SCR 239, [27] (emphasis added). For an earlier discussion of the importance of reliability in the English context, see Mike Redmayne, *Expert Evidence and Criminal Justice* (2001).

The previous examples from the common law, the *Evidence Act* and foreign jurisdictions indicate that judges are not required to be indifferent to validity and reliability or the actual probative value of expert opinion evidence. Similarly, continuing confidence in the lay jury does not seem to preclude the exclusion of otherwise admissible and incriminating expert opinions.

A Taking ‘Specialised knowledge’, ‘Probative value’ and ‘Unfair prejudice’ Seriously

The failure to take ‘reliability’ seriously tends to prevent judges from excluding untrustworthy, if incriminating, expert opinion evidence adduced by the prosecution. The failure to take ‘probative value’ seriously has meant that ‘unfair prejudice’ to the accused is rarely considered in relation to expert opinion evidence.

Unreliable expert opinion evidence and expert opinion evidence of unknown reliability routinely contaminate criminal trials. Rather than exhibiting genuine interest in the substantial fairness of criminal trials by excluding some kinds of incriminating expert opinion evidence, senior judges have preferred to rely upon a range of traditional protections (of unknown value). These formal *safeguards* – which include cross-examination, rebuttal experts and directions – do not individually or collectively make a trial fair or overcome the very serious prejudice that may attend from allowing unreliable expert opinion evidence to go before a jury.

How then should we proceed? There appear to be two possible routes. First, we could revise the *Evidence Act* to make the need for reliability explicit. We might have an exception to the opinion rule which looks something like the following:

79 Exceptions: opinions based on specialised knowledge

- (1) If a person has specialised knowledge based on the person’s training, study or experience, the opinion rule does not apply to evidence of an opinion of that person that is wholly or substantially based on that knowledge.
- (2) Where an opinion based wholly or substantially on specialised knowledge is adduced by the prosecutor in criminal proceedings it must be demonstrably reliable.
- (3) Where an opinion based wholly or substantially on specialised knowledge is not demonstrably reliable and is adduced by a defendant and admitted, subsection (2) does not apply to an opinion adduced to prove that that evidence should not be accepted.

Revisions like this would make the need for *demonstrable reliability* less ambiguous.¹⁸⁰

Through section 79(3), this particular proposal would also prevent the accused benefiting inappropriately from a lower admissibility threshold. It would enable the prosecutor, or another party, to respond *in kind* to expert opinion evidence

¹⁸⁰ ‘Demonstrable reliability’ does not require ‘absolute certainty’ or ‘infallibility’. Rather, it requires the prosecution to provide *good grounds* or a *credible basis* for believing that on the balance of probabilities a technique or approach is reliable. See, eg, *R v Gilmore* [1977] 2 NSWLR 935, 939–40 (Street CJ).

‘adduced by a defendant’ that does not meet the reliability standard imposed on the state. It seems preferable to revise section 79 rather than section 137 so that assessment of the reliability of expert opinion evidence is not restricted to a balancing exercise between probative value and unfair prejudice. It is not, after all, the text of section 137 that requires re-consideration.

The alternative route is to aim for similar results through revision of the relevant jurisprudence. Because similar outcomes could be achieved through re-interpretation of the *Evidence Act*, particularly reading ‘reliability’ into ‘specialised knowledge’ and assessments of ‘probative value’, and making judges personally responsible for determining the probative value of the prosecution’s expert opinion evidence, it is my intention to comment briefly on this alternative.

B Reading ‘Reliability’ into ‘Specialised knowledge’

While the need for ‘specialised knowledge’ may be no more demanding than the previous common law ‘field of expertise’ requirement, there is no reason to imagine that it is not demanding or should not entail some consideration of ‘reliability’. If we return to the definition of ‘knowledge’ endorsed by Spigelman CJ in *Tang*, we can appreciate that there is no *prima facie* reason for insensitivity to reliability:

... ‘knowledge’ connotes more than subjective belief or unsupported speculation. The term ‘applies to any body of known facts or to any body of ideas inferred from such facts or accepted as truths on good grounds’.

Significantly, the Supreme Court of the United States has repeatedly read the need for ‘reliability’ into its interpretations of the phrase ‘scientific, technical and other specialized knowledge’ from the *Federal Rules of Evidence* (1975). The *Daubert* decision required the trial judge, acting in the capacity of *gatekeeper*, to ‘ensure that any and all scientific testimony or evidence admitted is not only relevant, but reliable’.¹⁸¹ That approach was affirmed and extended to ‘technical and other specialised knowledge’ in *Kumho Tire Co v Carmichael*.¹⁸²

There would seem to be no particular reason why ‘specialised knowledge’ in section 79 could not support *demonstrable reliability*.¹⁸³ Notwithstanding the decision of the CCA in *Tang*, in a rational system of justice it is not desirable to interpret ‘knowledge’ and certainly not ‘specialised knowledge’ without incorporating some notion of reliability. Section 79 was not, after all, designed to enable creation scientists and astrologers to testify with impunity. Increased sensitivity to reliability would help to ground ‘specialised knowledge’ – that is, make it a more serviceable admissibility criterion.

‘Reliability’ has a range of technical meanings and implications whereas ‘specialised knowledge’, without more, is quite vacuous.¹⁸⁴ ‘Reliability’ tends to

181 *Daubert*, 509 US 579, 589 (1993).

182 *Kumho Tire Co v Carmichael*, 526 US 137, 148 (1999). See Beecher-Monas, above n 2.

183 *Makita* (2001) 52 NSWLR 705, [85] (Heydon JA); *Tang* (2006) 65 NSWLR 681, [134], [150]–[152] (Spigelman CJ, Simpson and Adams JJ concurring, and *HG* (1999) 197 CLR 414, [41] (Gleeson CJ) all employ the language of ‘demonstration’, but not in relation to ‘reliability’.

184 Consider the terms in the long extract from *Makita* (2001) 52 NSWLR 705, [85]. Few, if any, have pre-existing or established meanings among communities of experts.

engender an expectation of validation whereas a focus on ‘specialised knowledge’ encourages an exercise in hermeneutics. Consequently, rather than try to ascertain whether an expert’s opinions are based on ‘specialised knowledge’, especially where the existence of the particular *specialty* is contested or has tenuous links to an established field, a judge could look to information about the reliability of the techniques and opinions. For example, a judge might be interested in the kinds of studies that have been undertaken, where they have been published, their limitations, and the response from other experts. Such an approach converts judicial interest in an abstraction – whether something is ‘specialised knowledge’ – with no proper answer, to an investigation into whether some technique ought to be considered sufficiently reliable to go before a jury.

From here, the bases – Spigelman CJ’s ‘two limbs’ – could be used to supplement admissibility determinations: to ascertain whether the opinion is grounded in the demonstrably reliable technique and whether the techniques can be clearly and credibly linked to *particular* training, study or experience. Such an approach would transform the operation of section 79. Thus, for forensic scientific evidence there would be three basic considerations which, in order of importance, would be:

- Reliability.* ‘Specialised knowledge’ should be informed by the need for *demonstrably reliable* opinion evidence—at least with respect to opinions adduced by the prosecution. Novel techniques should encourage close judicial scrutiny. Though, the existence of a field, along with earlier admissibility decisions, should not be used to circumvent continuing judicial interest in reliability.
- Basis 1.* The opinion must be wholly or substantially *based* on the ‘specialised knowledge’. That is, the opinion must be clearly derived from the demonstrably reliable technique.
- Basis 2.* The ‘specialised knowledge’ must be *based* on ‘training, study, or experience’. The opinion should be based on demonstrably reliable techniques which are clearly linked to a person’s ‘training, study and experience’. That is, the person with the opinion should be competent with the demonstrably reliable technique from their training, study or experience.

This tripartite approach would inject reliability into the interpretation of section 79. Obviously, this is an interpretation that the text could sustain. All that is needed is for judges to make the need for reliability explicit in their admissibility jurisprudence.

One caveat is that this essay is concerned with expert opinion evidence *adduced by the prosecution*. The argument that this evidence, *in particular*, should be demonstrably reliable is not reflected in the text of section 79. Consequently, it may be that section 79 needs to be re-drafted to formally distinguish between the prosecution and the other parties. Alternatively, judges could informally impose a higher standard on the prosecution to reflect the

resources available to the state and reinforce its role as a model litigant. There are many good reasons why an individual plaintiff or criminal defendant should not be confronted with the same admissibility standard as that required of the state in serious criminal prosecutions. Also, it is important that a demonstrable reliability standard should not be used to discriminate *against* the accused (or plaintiffs in civil litigation). The need for demonstrably reliable expert evidence does not, however, depend on this asymmetry. Though, the failure to formally distinguish between plaintiffs, criminal defendants and the prosecution (or state) has been one of the main problems with admissibility jurisprudence in US federal courts in the aftermath of *Daubert*.¹⁸⁵

C (Actually) Determining ‘Probative value’

Regardless of what judges do with ‘specialised knowledge,’ the time has come to change the way they apply section 137 (s 135 and, for common law judges, the *Christie* discretion).

Trial judges should draw principled distinctions between lay evidence and expert opinion evidence. They must, in addition, be willing to make an assessment of the probative value of expert opinion evidence adduced by the prosecution. That is, trial judges should be attentive to the actual reliability of techniques so they can meaningfully assess ‘the extent to which the evidence *could rationally affect* the assessment of the probability of the existence of a fact in issue.’ This should be an objective test subject to *de novo* review by an appellate court.¹⁸⁶ Once the probative value of the expert opinion evidence adduced by the prosecution has been determined, a more meaningful balancing exercise can be undertaken. If the techniques and opinions are not demonstrably reliable then there will always be a very substantial risk that the jury will overvalue or misuse opinion evidence proffered by experts, police and investigators.

D Reliability?

Throughout, this essay has made relatively unproblematised reference to the concept of *reliability*. Although it is not my intention to suggest that recourse to ‘reliability’ will solve all the problems with expertise or provide a straightforward admissibility threshold, these are not reasons not to give

185 See Scientific Knowledge and Public Policy, *Daubert: The most influential Supreme Court Ruling You’ve Never Heard Of* (2003) <<http://www.defendingscience.org/upload/Daubert-The-Most-Influential-Supreme-Court-Decision-You-ve-Never-Heard-Of-2003.pdf>> at 24 March 2008; Lloyd Dixon, *Changes in the Standard for Admitting Expert Evidence in Federal Civil Cases Since the Daubert Decision* (2001); Simon Cole, ‘Grandfathering Evidence: Fingerprint Admissibility Rulings From *Jennings* to *Llera Plaza* and Back Again’ (2004) 41 *American Criminal Law Review* 1189–276; Gary Edmond and David Mercer, ‘*Daubert* and the Exclusionary Ethos: The Convergence of Corporate and Judicial Attitudes Towards the Admissibility of Expert Evidence in Mass Toxic Torts’ (2004) 26 *Law & Policy* 231.

186 According to James J in *R v SJRC* [2007] NSWCCA 142, [34] (Rothman and Harrison JJ concurring), on ‘appeal, an application of s 137 is to be regarded as analogous to the exercise of a discretion and, therefore, can be reviewed by an appellate court only in accordance with the principles stated in *House v The King* (1936) 55 CLR 499 at 504–5’. Cf *R v Cook* [2004] NSWCCA 52, [38] (Simpson J, Ipp JA and Adams J concurring).

‘reliability’ a prominent role in our admissibility jurisprudence. How should we approach admissibility? The answer seems to be to treat it in a flexible but principled manner.¹⁸⁷

Reliability entails trustworthiness. When trying to ascertain the trustworthiness or dependability of expert opinion evidence adduced by the prosecution, judges should consider whether there has been any attempt to validate or even improve the techniques and underlying assumptions. Testing is very important for forensic scientific evidence. While there may be disagreements over the adequacy, independence and significance of testing – obviously large scale testing, testing undertaken at arm’s length, and testing where the results are published in reputable peer reviewed journals, will generally be preferable – the failure to undertake validation studies or credibly respond to obvious limitations should ordinarily be damning.¹⁸⁸

Where a type of evidence, technique or procedure, whether DNA typing or identification based on sound recordings, CCTV or security images, is likely to be used repeatedly, there is an even greater need for the state (or prosecution) to test the techniques and demonstrate their reliability. *Ipse dixit*, or even educated guesses, are not appropriate when it comes to opinions based on ubiquitous and proliferating forms of evidence. To be admissible, the techniques supporting an expert’s opinions must be able to consistently and accurately do what is claimed for them. The prosecution must be able to satisfy the court, using evidence, that this is so.

Where we can actually measure validity and error rates, or determine that they have not been ascertained or even considered, it makes little sense to allow (or require) them to be determined by a jury on the basis of in-court conflict. Such an approach cannot be considered fair or efficient. Where vital information – such as the distribution of facial and body features – is not available, deference to the jury is incomprehensible.¹⁸⁹ Not all expert disagreement should be abandoned to the jury. The jury should be responsible for resolving disagreements around expert opinion evidence only after the prosecution has demonstrated that the incriminating expert opinion evidence satisfies a reliability threshold. This balances meaningful jury participation with the accused’s right to a fair trial.

Unfortunately, there are no simple algorithms or formulae for determining reliability or the adequacy of testing.¹⁹⁰ There are, however, many indicia that

187 Some commentators suggest that a more principled – that is, less rule-bound – approach would help to clarify practice. Consider John Braithwaite, ‘Rules and Principles: A Theory of Legal Certainty’ (2002) 27 *Australian Journal of Legal Philosophy* 47.

188 Jane Campbell Moriarty and Michael Saks, ‘Forensic Science: Grand Goals, Tragic Flaws, and Judicial Gatekeeping’ (2005) 44 *Judge’s Journal* 16; Michael Saks and Jonathan Koehler, ‘The Coming Paradigm Shift in Forensic Identification Science’ (2005) 309 *Science* 892. For some discussion of testing and its limitations, see Harry Collins, *Changing Order: Replication and Induction in Scientific Practice* (1992); Harry Collins and Trevor Pinch, *The Golem: What Everyone Should Know About Science* (1998).

189 Several judges have suggested that testing some types of expert opinion evidence might be difficult or impossible. In some cases this seems to be mistaken. Consider the remarks by Ipp JA in relation to ad hoc identification evidence from video recordings in *R v Li* (2003) 140 A Crim R 288, [111].

190 Collins, above n 188; Michael Mulkay and Nigel Gilbert, ‘Putting Philosophy to Work: Karl Popper’s Influence on Scientific Practice’ (1981) 11 *Philosophy of the Social Sciences* 389.

might prove helpful to judges.¹⁹¹ The real danger arises from treating any indicia inflexibly – like a checklist. Our judges need to develop more principled approaches to the reliability of expert opinion evidence and its role in different types of legal proceedings. In this task they might be assisted by a range of scientists, social scientists, legal scholars and meta-experts who are familiar with these kinds of issues or have experience in other jurisdictions and other kinds of policy and regulatory arenas.

Judges, prosecutors and experts should not remain indifferent to the reliability of expert opinion evidence. Failure to take reliability seriously is an abrogation of their respective duties. It undermines the rationality and fairness of criminal processes and threatens the veracity of outcomes. Expert opinion evidence adduced by the prosecution should not be admitted in criminal proceedings if its reliability cannot be persuasively established. In addition, when considering the exercise of a discretion (such as section 135) or undertaking the balancing exercise required by section 137, judges should determine the probative value of expert opinion evidence so that they actually afford some protection to the accused.

E Ritual or Rationality?

All this leaves us with two pressing questions. *In a criminal prosecution, who should bear the risk of unreliable expert opinion evidence? Can a rational system of criminal justice allow the state to adduce incriminating expert opinion evidence that is either unreliable or of unknown reliability?*

As things stand, the accused bears the risks of our legal system's disinterest in the reliability of expert opinion evidence and indifference to the effectiveness of its safeguards. It is hard to imagine more emasculated exclusionary discretions or a less principled admissibility jurisprudence.

¹⁹¹ See Edmond, *Pathological Science?*, above n 176.