The Empirical Psychological Science behind Ad Hoc Expert Voice Identification Evidence

RYAN CHAN[*]

Abstract

This article explores the common law creation of the ad hoc expert with respect to voice identification evidence. It considers the developments of the law, including the Uniform Evidence Law regime, in conjunction with an analysis of the empirical psychological studies on voice identification. The author contends that the broad application of the ad hoc expert provision, in relation to voice identification in this regime, has permitted incriminating but unreliable evidence into the criminal trial. It is clear that courts need to reconsider their approach to this specific area of evidence law in order to preserve the sanctity of the fair trial.

I INTRODUCTION

Australian courts have frequently allowed police officers and other investigators, such as translators and interpreters, to give opinion evidence on voice identification,[1] despite the fact that these witnesses often have little demonstrable 'skill, training or experience' in this forensic field.[2] This is possible through the exception of the 'ad hoc expert', where witnesses are deemed as experts often by virtue of 'displaced exposure'—usually repeated and remote listening to surveillance tapes in question.[3] Originally borne from common law, the ad hoc expert is often considered as a category of expert under s 79 of the Uniform Evidence Law regime (the 'Uniform Evidence Laws'),[4] which provides: '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'. Section 79 and its accommodation of the ad hoc expert serves as an exception to the well-known opinion rule, in which opinion evidence is not admissible to prove

---

[*] Master of Commercial Law candidate, University of Melbourne; BEcon/LLB (Hons), University of Queensland.

[1] In most cases, the courts are considering whether to allow witnesses to give evidence of their opinion as to whether a voice captured on a surveillance tape is the same as the voice of the accused. See, eg, *R v Butera* (1987) 164 CLR 180 ('*Butera*'); *R v Leung* (1999) 47 NSWLR 405 ('*Leung*'); *Li v The Queen* (2003) 139 A Crim R 281 ('*Li*'); *R v Riscuta* [2003] NSWCCA 6; *R v El-Kheir* [2004] NSWCCA 461.

[2] See *Evidence Act 1995* (Cth) ss 76, 79.

[3] See, eg, *Leung* (n 1); *Li* (n 1).

[4] The Uniform Evidence Laws comprises several Australian statutes: *Evidence Act 1995* (Cth); *Evidence Act 2011* (ACT); *Evidence Act 1995* (NSW); *Evidence (National Uniform Legislation) Act 2011* (NT); *Evidence Act 2001* (Tas); *Evidence Act 2008* (Vic).

the existence of a fact about the existence of which the opinion was expressed.[5] Over the years, ad hoc expert evidence on voice identification has been widely criticised due to its lack of commitment to scientific methods.[6] Without formal training or credible validity and reliability underpinning voice identification techniques, such opinion evidence poses problems to the fair trial.

This article contends that the decision to admit such incriminating opinion evidence of unknown probative value only serves to prejudice the fair criminal trial, and that traditional trial mechanisms of cross-examination and jury directions are insufficient to negate this danger. Part II will describe the Australian developments of the use of ad hoc experts for voice identification evidence. Part III will then provide an overview of the psychological research surrounding the reliability of voice identification evidence, and the efficacy of procedural trial safeguards. This part concludes that without formal training in the field of voice identification and comparison, opinion evidence of this kind is arguably no more than speculation. It is contended that ad hoc expert voice identification evidence should have no place in the criminal trial, and that greater collaboration between the law and expertise in phonetic, linguistic, and psychological forensic science is needed to provide light on how to better handle covert audio recordings as admissible evidence.

II THE AD HOC EXPERT

The ad hoc expert concept first arose from relatively restricted origins, in which its use was primarily confined to the production of transcripts from indistinct covert recordings as an interpretive aid for juries. However, over the years, its scope has since been significantly broadened to allow for the admission of voice identification evidence. Its beginnings can be traced to a New Zealand appeal case, *R v Menzies* ('*Menzies*'),[7] which involved the admissibility of a transcript of covert voice recordings prepared by a detective who had repeatedly listened to them. The recordings were of poor quality, to the extent that it was difficult to understand the content of the tapes. At first instance, the trial judge refused to admit the transcript as evidence but allowed for it to assist the jury in their comprehension of the tapes. The New Zealand Court of Appeal further held that the tapes were of such low intelligibility that it justified the use of an 'expert' to assist the jury. More relevantly, because the detective had listened to these tapes repeatedly, he had thereby 'acquired a special expertise' enabling him to

---

[5] See Uniform Evidence Laws s 76.
[6] See, eg, Gary Edmond, Kristy Martire and Mehera San Roque, 'Unsound Law: Issues With ('Expert') Voice Comparison Evidence' (2011) 35 *Melbourne University Law Review* 52, 53; Kristy Martire and Gary Edmond, 'Rethinking Expert Opinion Evidence' (2017) 40 *Melbourne University Law Review* 967; Gary Edmond and Mehera San Roque, 'Quasi-Justice: Ad Hoc Expertise and Identification Evidence' (2009) 33 *Criminal Law Journal* 8.
[7] [1982] 1 NZLR 41 ('*Menzies*').

craft the transcript. As per Cooke J: 'He may be a temporary expert in the sense that by repeated listening to the tapes he has qualified himself ad hoc'.[8] Importantly, the Court of Appeal qualified their decision by saying that such a transcript would not be appropriate for all jury trials, and that the transcript was to be classed only as an 'aid to the jury', rather than independent evidence of the tapes' content. Nonetheless, the concept of the ad hoc expert was born.

Close to a decade later, the ad hoc expert concept emerged in Australian courts in *Butera v DPP (Vic)* ('*Butera*'), which similarly to *Menzies* discussed the admissibility of transcripts derived from covert recordings of a heroin trafficking investigation. Here, several interpreters were used to prepare a written, translated transcript from recordings of conversations comprised of English, Thai, Punjabi, and Malayan languages. Citing *Menzies*, the majority stated:

> The translations are the respective renditions in English of what the interpreters heard in repeated playings of the tape. The respective translations are the product of the expertise which the interpreters brought to the task: they became ad hoc experts as to what was recorded by repeatedly listening to the tape being played over and they were experts in the languages to be translated.[9]

Similar to *Menzies*, *Butera*'s central issue was the evidentiary character of the transcripts, in which it held that the translated transcripts were to remain only as an interpretive aid for the jury.

The case of *Eastman v The Queen* ('*Eastman*') echoed the sentiments of *Menzies* and *Butera*,[10] but importantly, demonstrated an interaction of the common law ad hoc expert concept with the Uniform Evidence Laws. The court referred to s 48(1)(c) of the *Evidence Act 1995* (Cth), holding that this provision allowed for the admissibility of transcripts as evidence of the contents of audio tapes, but only where such content was unclear or contested.[11] In *R v Cassar* ('*Cassar*'), Sperling J sought to summarise the intersection between *Menzies, Butera, Eastman*, and the Uniform Evidence Laws.[12] A summary of his finding is outlined below:

a) A document that purports to be a transcript of words recorded on a tape is admissible to prove the conversation: [*Evidence Act 1995* (NSW)] s 48(1)(c);

b) No oral or other evidence is necessary to validate such a transcript, it being sufficient that it purports to be a transcript of the words: s 48(1)(c);

---

[8] Ibid 49 (Cooke J, with McMullin, Somers and Richmond JJ agreeing).
[9] *Butera* (n 1) 187–8 (Mason CJ, Brennan and Deane JJ).
[10] (1997) 76 FCR 9 ('*Eastman*').
[11] Ibid 110.
[12] [1999] NSWSC 436 ('*Cassar*').

c)   Where a tape is indistinct, a transcript may be used to assist the jury in the perception and understanding of what is recorded on the tape: *Butera* at 187;

d)   Where a tape is indistinct, a transcript made by an "ad hoc expert", being a person qualified only by having listened to the tape many times, may be used for this purpose. That is particularly so where the tape needs to be played over repeatedly before the words uttered could be made out unaided: *Menzies* at 49, cited in *Butera* at 188;

e)   If there is doubt or disagreement whether the transcript accurately deciphers the sounds captured on the tape, the transcript should be used only as an aide-memoire, I take that to mean that the jury is to give priority to what they hear (or do not hear) on the tape, if that is not consistent with what appears in the transcript: *Butera* at 188;

f)   The jury may have the transcript before them when the tape is played over in court: *Eastman* at 200;

g)   The jury should be informed, when the transcript is tendered, as to the use which they may make of it: *Eastman* at 220;

h)   A transcript may be rejected or its use limited pursuant to ss 135–137.[13]

In short, the operation of s 48 essentially permits the admissibility of a transcript as an interpretive aid for the jury, and as evidence of the content of the tape itself. The statutory provision does away with the common law need to endow the producer of the transcript as an ad hoc expert.[14] Yet despite this, the scope of the ad hoc expert has continued to endure and expand.

The case of *R v Leung* ('*Leung*') marked the transition from the use of ad hoc expertise in assisting the jury in comprehending covert audio surveillance tapes, to the use of positive identification of voices. To summarise, Mr Fung was an accredited interpreter working with the Australian Federal Police. He had been given several clandestine recordings which contained conversations in various Chinese languages. Mr Fung not only produced translated transcripts, but also expressed his opinion as to who he believed the identity of the speakers were in some of these tapes. This opinion was later repeated in evidence. At first instance, there was a debate upon the statutory restrictions on expert opinion evidence, particularly s 79 of the Uniform Evidence Laws, and it was conceded that Mr Fung's opinion was not based on 'specialised knowledge' or 'training, study, or experience'.[15] Moreover, during cross examination, Mr Fung also stated that 'he was not a voice expert, but had done his best to identify the voices'.[16] The Court of Criminal Appeal held, drawing from *Butera, Menzies,* and *Eastman*, that the concept of the ad hoc expert 'continues to have application under the New South Wales

---

[13] Ibid [7].
[14] Edmond and San Roque (n 6) 10.
[15] In *Leung*, Simpson J held that s 79 was 'sufficiently wide to accommodate' the common law notion of the ad hoc expert as referred to in *Butera* and *Menzies*: *Leung* (n 1) [40].
[16] Ibid [21].

evidentiary regime'.[17] However, they ultimately decided that because of various factors, such as Mr Fung's concession and the cross-lingual element of identification, that Mr Fung's opinion could not be characterised as one of an ad hoc expert, but was a lay opinion. It is important to note here, though, that by virtue of Mr Fung's repeated listening of the covert tapes, the Court was willing to consider whether he was within the realm of an ad hoc expert, which would have allowed him to make a positive identification.[18]

Shortly after, *Li v The Queen* (*'Li'*) went further than *Leung*, and unequivocally settled the use of the ad hoc expert exception in voice identification evidence. Similar to *Leung*, Mr Chan, an experienced and accredited interpreter, was asked to listen to a number of surveillance tapes, consisting of conversations in Cantonese and Mandarin. Mr Chan transcribed the contents of these tapes and also provided an opinion of the identity of the speaker in these tapes, denoting him as 'M1'. A year later, Mr Chan was asked to listen to a part of the appellant's police interview, which was conducted in English, and identify whether M1 was the appellant. Prior to making this comparison, he relistened to the original tapes, conceding 'that it might have been only once',[19] and identified the appellant, Li, as M1.

The appellant contested the evidence on various grounds, which included: that Mr Chan was not a voice identification expert and gave a lay person's opinion with respect to the similarity of voices between tapes; that there were no 'special features of the voice'; that people's voice on the phone are different to when speaking face to face; and that Mr Chan had 'no training, knowledge, or experience in comparing voices in speaking in English with those speaking in Cantonese'.[20] It was also submitted that the evidence of Mr Chan was 'tainted' as he must have known that the appellant was a suspect when he compared the appellant's voice on the record of interview with the voice of M1 on the five original tapes.[21] Due to these considerations, the appellant argued that under s 137 of the *Evidence Act 1995* (NSW), the evidence should have been excluded as the danger of unfair prejudice against the accused outweighed its probative value.[22] However, Ipp J (Whealy and Howie JJ agreeing) drew upon *Leung* and relied upon the concept of ad hoc expert to dismiss the appellant's arguments. Because Mr Chan had repeatedly listened to the tapes, Ipp J held that Mr Chan had acquired a 'familiarity' with the voice of M1,

---

[17] Ibid [40]–[67].
[18] Edmond and San Roque (n 6) 11.
[19] *Li* (n 1) [36].
[20] Ibid [45].
[21] Ibid [58].
[22] *Evidence Act 1995* (NSW) s 137: '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 accused.'

qualifying him as an ad hoc expert to give a reliable and valid opinion about the identity of the voice.[23] On this basis, his Honour rejected that only voice identification experts ought to give evidence in respect to voice identification:

> The very many hours that Mr Chan spent listening to and working on the five tapes with a view to identifying the words spoken by M1 qualified him as an ad hoc expert in the characteristics of M1's voice ... Mr Chan did not have to become a voice recognition expert to become an ad hoc expert within the meaning of *R v Leung*. In any event, his years of practising as an interpreter and translator would have been of great assistance to him as it meant he had considerable familiarity with the language spoken and with voices speaking Cantonese and English.[24]

Justice Ipp also found that the cross-lingual element in the comparison of tapes that Mr Chan was required to undertake *did not* detract from his ability to express a reliable opinion.[25] Ipp J ultimately held that by way of s 79 of the Uniform Evidence Laws, the voice identification evidence was properly admitted into trial.

Justice Ipp's findings continue to be readily accepted by the courts, and *Li* and *Leung* still operate as authoritative law, having been cited in numerous jurisdictions in recent years. In New South Wales, both *Nguyen v R* and *R v Nasrallah* have affirmed the admissibility of voice identification evidence from investigators through the ad hoc expert concept.[26] This is similarly reflected in South Australia in *R v Phan*.[27] Victorian courts, however, have diverged from using s 79 and the ad hoc expert exception as a pathway to admitting voice identification evidence. In *Kheir v R* and *Tran v The Queen; Chang v The Queen*, it was held that in Victoria, 'identity evidence is not treated as a matter requiring proof of expertise, whether ad hoc or otherwise', and that voice identification evidence is 'not to be regarded as a field of expertise about which only experts could give evidence'.[28] However, this difference is slight, and Victorian courts are still willing to allow investigators, by virtue of their repeated listening of particular voice recordings, to provide voice identification evidence through s 78, rather than s 79 of the Uniform Evidence Laws. Section 78 conventionally allows for eyewitnesses to express opinions when necessary to understand their perceptions.[29] While the debate between the

---

[23] *Li* (n 1) [39], [42], [48], [51].
[24] Ibid [42], [48], [51].
[25] Ibid [56].
[26] *Nguyen v R* [2017] NSWCCA 4, [41]–[51]; *R v Nasrallah* [2015] NSWCCA 188, [10]–[47].
[27] *R v Phan* [2017] SASCFC 70, [72]–[73].
[28] *Kheir v R* [2014] VSCA 200, [61]–[62]; *Tran v The Queen; Chang v The Queen* [2016] VSCA 79, [60].
[29] Uniform Evidence Laws s 78 provides for the lay opinion exception to the opinion rule:

application of these sections is outside the scope of this article, it is worthwhile to note that using s 78 to admit such opinion identification evidence has been criticised as wholly inappropriate:

> It is not appropriate to use s 78 to enable those, but especially investigators, who did not directly perceive a matter or event, to express their opinions in criminal proceedings. Section 78 should be restricted to direct sensory witnesses. It should not be used to admit the opinions of those listening and watching displaced by time or space.[30]

Rather, the opinions of those with expertise or purported expertise should be regulated only by s 79.[31] It is clear that the concept of the ad hoc expert continues to serve as an available avenue for legal practitioners seeking to adduce voice identification opinion evidence from police officers and other investigators, such as interpreters and translators.

III THE SCIENCE BEHIND VOICE IDENTIFICATION

It is respectfully submitted that the findings by Ipp J are incorrect. The 'technique' of repeated listening to tapes does not guarantee that their opinions are any more than subjective belief or unsupported speculation.[32] In the absence of evidence about the validity and reliability of this technique, it is erroneous to believe that such opinions have been inferred from 'good grounds'.[33] Moreover, there is a plethora of psychological research indicating our unconscious processes and conscious judgments are easily influenced by various factors, regardless of the technique in question.[34]

A *Voice Identification Studies*

In testing unfamiliar voices,[35] one notable study exposed participants to either 30 or 70 seconds of a previously unknown voice. It was found that

---

The opinion rule does not apply to evidence of an opinion expressed by a person if: (a) the opinion is based on what the person saw, heard or otherwise perceived about a matter or event, and; (b) evidence of the opinion is necessary to obtain an adequate account or understanding of the person's perception of the matter or event.

See also *R v Whyte* [2006] NSWCCA 75; *Lithgow City Council v Jackson* (2011) 244 CLR 352; Gary Edmond, 'Regulating Forensic Science and Medicine Evidence at Trial: It's Time for a Wall, a Gate and Some Gatekeeping' (2020) 94 *Australian Law Journal* 427, 432.

[30] Edmond (n 29) 432.

[31] Ibid.

[32] Edmond and San Roque (n 6) 21.

[33] Ibid.

[34] For a further discussion in respect to the dichotomy between automatic and unconscious thinking (System 1) and conscious and controlled thinking (System 2), see Daniel Kahneman and Amos Tversky, 'Prospect Theory: An Analysis of Decision Under Risk' (1979) 47(2) *Econometrica: Journal of the Econometric Society* 263.

[35] The term 'familiar' in empirical voice identification literature is used to denote a threshold of perception whereby something or someone becomes recognisable or identifiable. A person's voice is therefore 'familiar' when the listener of that voice is able to attribute a

listeners could only identify the target voice in 42 % of instances presented. More alarmingly, in the target-absent condition, listeners identified another previously unheard, innocent voice as the target voice 51 percent of the time.[36] Several studies have reported similar findings in respect to listeners and unfamiliar voices.[37]

Voice identification studies with familiar voices have yielded conflicting findings. While some studies have found that familiar voices are recognised with a high degree of accuracy by listeners,[38] other investigations have presented contrasting results.[39] In one study, 20 men belonging to a social fraternity were given voice samples of varying duration from nine men of their fraternity ('familiar') and 11 unknown men ('unfamiliar'). It was reported that only 60% of listeners could recognise all nine familiar voices as members of their fraternity; moreover, six of the 11 unfamiliar voices were wrongly identified as members of their fraternity.[40] It is particularly concerning that even when listeners are given samples of familiar voices, and voices comparably more familiar than the samples the ad hoc experts in *Li* were tasked with, voice identification is not consistently accurate.

Various factors have also been shown to diminish the accuracy of voice identifications. For instance, if the quality of the speech is poor, such as if the voice is heard through a telephone, whispered, or part of a low-quality

name to that voice, or link that voice to prior exposure, with a high degree of accuracy. Conversely, 'unfamiliar' is the exact opposite; usually voices by unknown people. People may never achieve this precise level of familiarity with a voice unless they have been repeatedly exposed to it in a variety of different occasions (typically family members and close friends). See, eg, Kanae Amino and Takayuki Ara, 'Effects of Linguistic Contents on Perceptual Speaker Identification: Comparison of Familiar and Unknown Speaker Identifications' (2009) 30 *Acoustical Science and Technology* 89; Diana Roupas Van Lancker et al, 'Phonagnosia: A Dissociation between Familiar and Unfamiliar Voices' (1988) 24 *Cortex* 195.

[36] José Kerstholt et al, 'Earwitnesses: Effects of Speech Duration, Retention Interval and Acoustic Environment' (2004) 18 *Applied Cognitive Psychology* 327.

[37] A Daniel Yarmey et al, 'Commonsense Beliefs and the Identification of Familiar Voices' (2001) 15 *Applied Cognitive Psychology* 283; José Kerstholt et al, 'Earwitnesses: Effects of Accent, Retention and Telephone' (2006) 20 *Applied Cognitive Psychology* 187; Brian Clifford, 'Voice Identification by Human Listeners: On Earwitness Reliability' (1980) 4 *Law and Human Behaviour* 373.

[38] Evelyn Abberton and Adrian Fourcin, 'Intonation and Speaker Identification' (1978) 21 *Language and Speech* 305; Harry Hollien et al, 'Perceptual Identifications of Voices Under Normal, Stress and Disguise Speaking Conditions' (1982) 10 *Journal of Phonetics* 139.

[39] Bonnie Bartholomeus, 'Voice Identification by Nursery School Children' 27 *Canadian Journal of Psychology* 464; Daniel Read and Fergus Craik, 'Earwitness Identification: Some Influences on Voice Recognition' (1995) 1 *Journal of Experimental Psychology: Applied* 6.

[40] Alvin Goldstein and June Chance, 'Voice Recognition: The Effects of Faces, Temporal Distribution of Practice and Social Distance' (Working Paper, Midwestern Psychology Association, 1985). See also Phil Rose and Sally Duncan, 'Naïve Auditory Identification and Discrimination of Similar Voices by Familiar Listeners' (2013) 2(1) *International Journal of Speech Language and the Law* 1.

recording.[41] Other factors include the exposure time,[42] if the pitch of the voice had been altered,[43] or if there is a delay between the original exposure and subsequent identification.[44] On this latter point, it has been found that listeners were able to identify voices at a peak rate of 49% after a delay of one week, but only 8% after three weeks.[45]

In light of the scientific evidence above, it is clear that the interpreter's and investigator's opinion in *Li* should not have been admitted as evidence. First and foremost, the technique of repeatedly listening to audio surveillance tapes has not been meaningfully tested, and any opinion drawn from it lacks valid and reliable substance. There is little evidence that police, translators and interpreters, and even linguists perform better than average or are particularly accurate at comparisons across the many different conditions confronting earwitnesses and listeners.[46] Moreover, to become familiar enough with a voice to produce relatively high accuracy in identification,[47] one must have listened to the voice for a significant amount of time in a variety of different contexts. This familiarity does not seem to be achieved by repeated listening of audio surveillance tapes.

Such tapes in *Li* were also characterised by several factors which may diminish accuracy in voice identification, such as low quality or telephone recordings, alteration of pitch, and short duration of tapes.[48] The fact that the interpreter made the identification a whole year after the exposure to the tapes is also concerning, given that studies have shown significant declines in accuracy rates after only three weeks.[49]

The cross-lingual voice comparison which the investigators were required to undertake further jeopardised the accuracy of their findings. Similar to the voice identification studies outlined above, cross-lingual voice comparison studies have demonstrated low accuracy rates. In one study by

---

[41] A Daniel Yarmey, 'Earwitness Speaker Identification' (1995) 1 *Psychology, Public Policy, and Law* 792; Yarmey et al (n 37); Tara Orchard and A Daniel Yarmey, 'The Effects of Whispers, Voice-Sample Duration, and Voice Distinctiveness on Criminal Speaker Identification' (1995) 9 *Applied Cognitive Psychology* 249.

[42] Susan Cook and John Wilding, 'Earwitness Testimony: Never Mind the Variety, Hear the Length' (1997) 11 *Applied Cognitive Psychology* 249; Susan Cook and John Wilding, 'Earwitness Testimony: Effects of Exposure and Attention on the Face Overshadowing Effect' (2001) 92 *British Journal of Psychology* 617.

[43] Howard Saslove and A Daniel Yarmey, 'Long-Term Auditory Memory: Speaker Identification' (1980) 65 *Journal of Applied Psychology* 111.

[44] Lori van Wallendael et al, '"Earwitness" Voice Recognition: Factors Affecting Accuracy and Impact on Jurors' (1994) 8 *Applied Cognitive Psychology* 661.

[45] Clifford (n 37).

[46] Edmond, Martire and San Roque (n 6) 102.

[47] It is contended that accuracy rates below 90 percent are not sufficient in forensic contexts given the serious consequences associated with an error: Ibid 91.

[48] *Li* (n 1) [70].

[49] The term 'significant' for the purposes of this paper is used in the context of psychological research, which is commonly defined as results occurring due to chance at a probability of below 5%.

Goggin, it was reported that accurate identification rates ranged between 12% and 35% for listeners making identifications across languages.[50] Another study found an accuracy rate of around 47%, with a false alarm rate over 67%, even when the second language was familiar to listeners.[51] The accuracy of cross-lingual identification is also influenced by a variety of factors, such as the types of languages being compared and, most relevantly, the listener's familiarity with the speaker's voice.[52]

B *Confirmation Bias*

Confirmation bias describes situations where investigators are more inclined to view evidence aligning with their preconceived expectations, rather than at its face value.[53] In the context of identification evidence, investigators who are provided with clear clues that others believe the source and target voice originate from the same person, there is an increased likelihood that that investigator will declare a match between the two pieces of evidence, despite the voices originating from different speakers. This phenomenon has been found in forensic finger print examiners, in which they were given extraneous, inaccurate information about evidence and consequently produced mistaken conclusions.[54] Contextual cues affect highly skilled experts, influencing decision-making at an unconscious level, and even formal training and experience will not necessarily protect them from making errors.[55] The implication of this is that it is incredibly difficult to meaningfully cross-examine these issues.[56] We are oblivious to the unconscious processes of our brain,[57] and the situational forces that profoundly impact upon our behaviour, and thus cannot report upon it.[58]

---

[50] Judith Goggin et al, 'The Role of Language Familiarity in Voice Identification' (1991) 19 *Psychology* 448.

[51] Axelle Philippon et al, 'Earwitness Identification Performance: The Effect of Language, Target, Deliberate Strategies and Indirect Measures' (2007) 21 *Applied Cognitive Psychology* 539.

[52] Olaf Köster and Niels Schiller, 'Different Influences on the Native Language of a Listener on Speaker Recognition' (1997) 4 *Forensic Linguistics* 18; Kirk Sullivan and Frank Schlicting, 'Speaker Discrimination in a Foreign Language: First Language Environment, Second Language Learners' (2007) 7 *Forensic Linguistics* 95.

[53] Itiel Dror, 'Human Expert Performance in Forensic Decision Making: Seven Different Sources of Bias' (2017) 49(5) *Australian Journal of Forensic Science* 541; Gretchen Chapman and Eric Johnson, 'Incorporating the Irrelevant: Anchors in Judgments and Belief and Value' in Thomas Gilovich, Dale Griffin and Daniel Kahneman (eds), *Heuristics and Biases: The Psychology of Intuitive Judgment* (Cambridge University Press, 2002) 120, 133.

[54] Itiel Dror et al, 'When Emotions Get the Better of Us: The Effect of Contextual Top-Down Processing on Matching Fingerprints' (2005) 19 *Applied Cognitive Psychology* 799; Itiel Dror, David Charlton and Alisa Péron, 'Contextual Information Renders Experts Vulnerable to Making Erroneous Identifications' (2006) 156 *Forensic Science International* 74.

[55] Edmond, Martire and San Roque (n 6) 89.

[56] Ibid; *Nguyen v The Queen* (2002) 26 WAR 59, 87 [124] (Anderson J).

[57] Richard Nisbett and Timothy Wilson, 'Telling More Than We Can Know: Verbal Reports on Mental Processes' (1977) 84(3) *Psychological Rev* 231.

[58] Ibid.

In *Li*, the fact that Mr Chan listened to the original five tapes prior to making a positive identification of the appellant in the recorded police interview gives rise to a concern of such confirmation bias. As correctly identified by the appellant, the effect of relistening to the original five tapes primed Mr Chan's expectation that Li was a suspect when undertaking this voice comparison. However, such bias was quickly dismissed by Ipp J on the basis that it does not affect voice identification evidence.[59] This is wholly incorrect. There is simply no empirical evidence to suggest voice identification evidence is less susceptible to this danger than finger print or photographic identification evidence.[60] Ipp J's reasoning appears to mirror the Dunning-Kruger effect, a cognitive bias where people mistakenly believe they are capable in assessing the validity and reliability of evidence, despite the vast psychological evidence suggesting otherwise.[61]

IV Misplaced Judicial Reliance on Trial Safeguards

The current preference of Australian judges for managing the potential dangers of incriminating voice evidence is by way of careful jury directions.[62] There have been various studies conducted concerning the efficacy of jury directions, but perhaps the most glaring problem relates to the judge him/herself.

According to Martire and Edmond, courts typically impair their analysis of the reliability and validity of expert evidence through the focus on individual cases rather than on empirical, scientific literature.[63] Studies that test the performance levels of purported experts under various conditions (such as voice identification ability by laymen), have usually been unreasonably dismissed by the courts, even though these studies may

---

[59] *Li* (n 1) [58]–[60].

[60] A Daniel Yarmey, 'The Psychology of Speaker Identification and Earwitness Memory' in Lindsay et al (eds), *The Handbook of Eyewitness Psychology, Vol. 2. Memory for People* (Laurence Erlbaum Associates Publishers, 2007) 101.

[61] David Dunning and Justin Kruger, 'Unskilled and Unaware of it: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments' (1999) 77 *Journal of Personality and Social Psychology* 1121; Jon Hanson and David Yosifon, 'The Situation: An Introduction to the Situational Character, Critical Realism, Power Economics, and Deep Capture' (2003) 152 *University of Pennsylvania Law Review* 129.

[62] Chantelle Baguley, Blake McKimmie and Barbara Masser, 'Deconstructing the Simplification of Jury Instructions: How Simplifying the Features of Complexity Affects Jurors' Application of Instructions' (2017) 41(3) *Law and Human Behaviour* 284; Edmond, Martire and San Roque (n 6) 98. Jury instructions assist juries to decide legally correct verdicts because they outline the process juries should follow to evaluate the evidence and decide their verdict. Instructions should reduce the likelihood that juries rely on irrelevant information or biases to decide their verdict.

[63] Martire and Edmond (n 6) 987.

provide helpful, informed inferences to be drawn on the evidence's probative value.[64]

How are judges, in their instructions, able to convey to the jury the dangers of ad hoc voice identification evidence if they do not even understand them themselves?[65] Such ignorance may be partially explained by what psychologists call the 'curse of knowledge'. Because judges have had greater exposure to false confessions and ad hoc expert evidence, they may fail to empathise with the jury's understanding of these issues and be inclined to overestimate their own ability to explain these complex phenomena to juries.[66]

Turning to jury directions themselves, research has found that jury directions are complex, and that jurors face difficulty understanding them.[67] Indeed, it has been shown that mock jurors only understand 50 to 70% of instructions, meaning that they will inevitably rely on mental shortcuts or heuristic cues, such as the source of the information or other extraneous material, rather than the content of the information itself.[68] Various studies have also suggested that jury directions are futile, as jurors are unable and often unwilling to forget incriminating testimonies heard.[69] More alarmingly, directions to disregard inadmissible evidence can sometimes magnify the impact erroneous evidence has on jurors' decision making.[70]

---

[64] Ibid. See, eg, *Li* (n 1) [106]–[111] (Ipp JA); Veronica Stinson, Jennifer Devenport and David Kravitz, 'How Effective is the Motion-to-Suppress Safeguard? Judges' Perceptions of the Suggestiveness and Fairness of Biased Lineup Procedures' (1997) 82(2) *Journal of Applied Psychoogy* 211.

[65] See Edmond, Martire and San Roque (n 6) 103; Martire and Edmond (n 6) 987.

[66] Michael Saks and Barbara Spellman, *The Psychological Foundations of Evidence Law* (New York University Press, 2016) 19–22; David Dunning et al, 'Flawed Self-Assessment: Implications for Health, Education, and the Workplace' (2004) 5(3) *Psychological Science in the Public Interest* 69.

[67] Brian Cutler, Hedy Dexter and Steven Penrod, 'Nonadversarial Methods for Improving Juror Sensitivity to Eyewitness Evidence' (1990) 20 *Journal of Applied Social Psychology* 1197; Amiram Elwork, Bruce Sales and James Alfini, 'Towards Understandable Jury Instructions' (1982) 65 *Judicature* 432; Alan Reifman, Spencer Gusick and Phoebe Ellsworth, 'Real Jurors' Understanding of the Law in Real Cases' (1992) 16 *Law and Human Behaviour* 539; Bradley Saxton, 'How well do Jurors Understand Jury Instructions? A Field Test using Real Juries and Real Trials in Wyoming' (1998) 33 *Land and Water Law Review* 59. This is problematic as jurors will only apply directions to the extent that they understand them.

[68] Baguley, McKimmie and Masser (n 62).

[69] Steven Fein, Allison McCloskey and Thomas Tomlinson, 'Can the Jury Disregard that Information? The Use of Suspicion to Reduce the Prejudicial Effects of Pretrial Publicity and Inadmissible Testimony' (1997) 23 *Personality and Social Psychology Bulletin* 1215; Saul Kassin and Samuel Sommers, 'Inadmissible Testimony, Instructions to Disregard and the Jury: Substantive Versus Procedural Considerations' (1997) 23 *Personality and Social Psychology Bulletin* 1046.

[70] John Reinard and Rodney Reynolds, 'The Effects of Inadmissible Testimony Objections and Rulings on Jury Decisions' (1978) 14 *Journal of the American Forensic Association* 91;

The safeguard of cross-examination is also commonly relied upon in discovering any dangers surrounding expert evidence.[71] Proponents of cross-examination assume three points: first, that lawyers are aware of the technical detail and limitations of the evidence presented; second, that they are capable of effectively conveying such limitations to a lay jury, and that they have the resources to do so; third, that jurors are sensitive to these same factors when evaluating evidence and rendering verdicts.[72] As expressed above, lawyers are generally not cognisant of the empirical research surrounding voice identification evidence, let alone the notion that the expert's claim should be validly and reliably tested.[73] Intuitively, jurors would have even less of an understanding or comprehension of such empirical research if it were presented.[74]

In a study conducted by Devenport and colleagues, legal practitioners and jurors were assessed on their sensitivity to major biases capable of diminishing eyewitness's accuracy in positive line-up identifications. These included instruction bias and presentation bias. Instruction bias is present when the eyewitness is told or is implied that the culprit is in the line-up before them. Presentation bias is found in simultaneously presented line-ups as opposed to sequentially presented line-ups. In simultaneously presented line-ups, all line-up members are presented at one time and eyewitnesses compare them with one another before making an identification decision. In contrast, sequentially presented line-ups allow the eyewitness to view only one line-up member at a time, and the eyewitness must then make an identification decision before viewing the next line up member. Moreover, in sequentially presented line-ups, eyewitnesses are not allowed to view previously viewed line-up members and are generally not informed as to how many line-up members he or she will be seeing. For many years, researchers have reliably demonstrated that there are significantly higher rates of false identifications made by eyewitnesses in line-ups containing these biases.[75] In this study, it was found that when watching a pre-recorded criminal trial, legal practitioners and jurors were not sensitive towards the deleterious effects resulting from

---

J Tanford, 'The Law and Psychology of Jury Instructions' (1990) 69 *Nebraska Law Review* 71.

[71] Gary Edmond et al, 'How to Cross-Examine Forensic Scientists: A Guide for Lawyers' (2014) 39 *Australian Bar Review* 174.

[72] Jennifer Devenport et al, 'How Effective Are the Cross-Examination and Expert Testimony Safeguards? Jurors' Perceptions of the Suggestiveness and Fairness of Biased Lineup Procedures' (2002) 87(6) *Journal of Applied Psychology* 1042.

[73] Ibid.

[74] Ibid.

[75] See, eg, Jennifer Devenport and Ronald Fisher, 'The Effect of Authority and Social Influence on Eyewitness Suggestibility and Person Recognition' (1996) 11 *Journal of Police and Criminal Psychology* 35; Saul Kassin et al, 'On the "General Acceptance" of Eyewitness Testimony Research' (2001) 56 *American Psychologist* 405; Janat Parker and Virginia Ryan, 'An Attempt to Reduce Guessing Behaviour in Children's and Adults' Eyewitness Identifications' (2003) 17 *Law and Human Behavior* 11; Gary Wells, 'What Do We Know About Eyewitness Identification?' (1993) 48 *American Psychologist* 553.

instruction and presentation bias, leading the researchers to conclude that cross-examination as a safeguard may not be adequate in ameliorating such risks. While it is acknowledged that voice identification and line-up identification are different fields, the study demonstrates that if legal practitioners are not able to recognise and articulate these biases in their cross-examination of identification evidence, it is highly unlikely jurors will be made aware or convinced of its harmful effects on the ad hoc experts' identification accuracy.[76] Finally, to reiterate, there remain major difficulties in cross-examining expert witnesses upon the biases and contextual cues which affect their unconscious processes and overall judgments.

A chorus of commentators have suggested an alternative safeguard, being the provision of expert testimony by a qualified research psychologist.[77] However, the evidence supporting this proposition is mixed, and cannot be unequivocally said to improve jurors' ability in evaluating earwitness testimony.[78] Indeed, a recent study by Kemp and Martire found that the presence of an expert witness in a mock trial scenario had no significant effect on juror sensitivity to eyewitness accuracy.[79] There are also strong policy reasons against using psychologist's testimony as a safeguard.[80]

V DIALOGUE WITH LAW AND FORENSIC SCIENCE

In recent years, there has been growing recognition towards the evidentiary concerns surrounding indistinct covert recordings. In 2017, there was a 'Call to Action' from the Australian Linguistics Society, Australian Institute of Interpreters and Translators, Australasian Speech Science and Technology Association, and the Applied Linguistics Association of Australia, for judicial reform towards the use of covert recordings in four main areas: transcription of English, translation of languages other than English, attribution of utterances to speakers, and 'enhancing' of indistinct

---

[76] Devenport et al (n 72). See also Cutler, Dexter and Penrod (n 67).

[77] Brian Cutler, Hedy Dexter and Steven Penrod, 'Expert Testimony and Jury Decision Making: An Empirical Analysis' (1989) 7(2) *Behavioral Sciences and the Law* 215; Devenport et al (n 72).

[78] Michael Leippe, 'The Case for Expert Testimony About Eyewitness Memory' (1995) 1(4) *Psychology, Public Policy, and Law* 909; R Lindsay, Gary Wells and Fergus O'Connor, 'Mock-Juror Belief of Accurate and Inaccurate Eyewitnesses: A Replication and Extension' (1989) 13(3) *Law and Human Behavior* 333.

[79] Kristy Martire and Richard Kemp, 'The Impact of Eyewitness Expert Evidence and Judicial Instruction on Juror Ability to Evaluate Eyewitness Testimony' (2009) 33 *Law and Human Behavior* 225.

[80] Experts carry an expensive cost in court time and fees. The burden should therefore fall on the State, who has a wealth of resources and the heavy obligation of proving guilt beyond reasonable doubt. It would be wrong for the defence to assume this responsibility. See also Edmond, Martire and San Roque (n 6) 97.

audio.[81] Indeed, in addition to voice identification evidence, there have been numerous studies to indicate that other evidence in relation to covert recordings, such as transcripts and enhanced recordings as prepared by police investigators, interpreters, and security agencies are similarly prone to being unreliable and misleading towards juries.[82]

The 'Call to Action' was specifically addressed to the Australasian Institute of Judicial Administration and was eventually taken up by the Judicial Commission for Cultural Diversity ('JCCD'). At the end of 2019, a workshop in relation to the 'Call to Action' was held, consisting of numerous linguists, four judges of the JCCD, and representatives of police and public prosecutors from Australian jurisdictions. It was reported that judges of the JCCD accepted the problems with the handling of covert recordings and agreed to facilitate collaborative research aiming to improve the procedures.[83] While official specifics of this future collaboration are yet to be revealed, it is encouraging to observe a willingness from both the judiciary and relevant scientific fields to improve the reliability of admissible evidence stemming from indistinct covert recordings.

VI CONCLUSION

This article argues that voice identification opinions formed from non-qualified experts, such as police detectives and interpreters, should be rejected. There is an absolute necessity for techniques to be submitted to scientific evaluation, and if this is forgotten, there are very serious dangers attaching to this kind of evidence. Incriminating opinions should always be underpinned by techniques and approaches which are scientifically valid and reliable. Therefore, it is contended that the courts should not rely on the cases of *Li* and *Leung* as authority in deciding whether to allow ad hoc

---

[81] Forensic Transcription Australia, *A Call to Action: Australian Linguists Call on the Judiciary to Reform Practices for Using Covert Recordings as Evidence in Court* (2020) <https://forensictranscription.com.au/a-call-to-action/>; Australian Linguistic Society, *ALS Newsletter November 2019* (November 2019)
<https://als.asn.au/newsdetail/32063/als_newsletter_november_2019>.

[82] See, eg, Peter French and Helen Fraser, 'Why "Ad Hoc Experts" Should Not Provide Transcripts of Indistinct Forensic Audio, and a Proposal for a Better Approach' (2018) 42 *Criminal Law Journal* 298; Helen Fraser, 'Enhancing Forensic Audio: What if all that really gets Enhanced is the Credibility of a Misleading Transcript?' (2020) 52(4) *Australian Journal of Forensic Sciences* 465; Helen Fraser, 'Real Forensic Experts Should Pay More Attention to the Dangers Posed by Ad Hoc Experts' (2017) 50(2) *Australian Journal of Forensic Sciences* 125; Kate Burridge, 'The Dark Side of Mondegreens: How a Simple Mishearing can Lead to Wrongful Conviction', *The Conversation* (Web Page, June 2017) <https://theconversation.com/the-dark-side-of-mondegreens-how-a-simple-mishearing-can-lead-to-wrongful-conviction-78466>.

[83] 'Research Hub for Language in Forensic Evidence' *The University of Melbourne* (Web Page, 2020) <https://arts.unimelb.edu.au/school-of-languages-and-linguistics/our-research/research-centres-hubs-and-units/research-hub-for-language-in-forensic-evidence>; Australian Linguistic Society (n 81).

experts to produce opinion identification evidence. Moreover, s 79 of the Uniform Evidence Laws should no longer serve as a statutory basis for adducing ad hoc expertise voice identification evidence. Such evidence is not based on 'specialised knowledge' and is therefore nothing more than a subjective opinion of unknown reliability.[84] However, it is acknowledged that indistinct covert recordings serve an important role in criminal investigations and it is essential that there continues to be frequent collaboration between the judiciary and experts in phonetics, linguistics and psychology in reviewing and creating rigorous practices for admissible evidence stemming from audio recordings.

---

[84] Edmond and San Roque (n 6) 33.