

The Demeanour Fallacy

ROBERT FISHER*

Judges, juries and arbitrators have to decide whether to believe witnesses. It is popularly believed that studying a witness's demeanour helps in that exercise. Many legal methods and principles have been built on that assumption.

Scientific studies show the assumption to be false. It is not possible to tell whether speakers are telling the truth by watching and listening to them.

The legal world has yet to respond. Judges and other professional fact-finders need to understand that their traditional reliance on demeanour assessment is misconceived. Juries should be expressly warned against it. Appellate Courts should change their approach to it. There are good reasons for appellate reluctance to interfere with decisions at first instance but they do not include credibility.

Popular belief

It is popularly believed that studying a speaker's demeanour will help the observer to decide whether or not the speaker is lying.¹ As an American commentator put it in 1949:²

All of us know that, in every-day life, the way a man behaves when he tells a story — his intonations, his fidgetings or composure, his yawns, the use of his eyes, his air of candour or of evasiveness — may furnish valuable clues to his reliability.

The Courts, too, have traditionally assumed that liars look and sound shifty. To quote an English law lord, Lord Pearce:³

One thing is clear, not so much as a rule of law but rather as a working rule of common sense. A trial Judge has, except on rare occasions, a very great advantage over an appellate Court; evidence of a witness heard and seen has a very great advantage over a transcript of that evidence; and a Court of Appeal should never interfere unless it is satisfied both that the judgment ought not to stand and that the divergence of view between the trial Judge and the Court of Appeal has not been occasioned by any demeanour of the witnesses or truer atmosphere of the trial (which may have eluded an appellate Court) or by any other of those advantages which the trial Judge possesses.

Such is one of the premises on which the common law system of justice has been built. It appears to be at least a contributing factor in the common law requirements that critical evidence be given orally rather than in writing; that witnesses give their own evidence in the absence of special reasons justifying hearsay; that accusers personally confront accused; and, more recently, that witnesses should not be concealed behind a niqab.⁴ It underpins the deference that appellate courts pay to the supposed advantage of first instance

* * Hon Robert Fisher QC, LLD, arbitrator, mediator and judge of Courts of Appeal of Samoa, Cook Is and Tuvalu. My thanks are due to Augustine Choi for his research assistance with this article.

1 As, for example, in Allan Pease's long-standing *Body Language: How to read others' thoughts by their gestures* (Camel Publishing, Sydney, 1981) at 17. For recent impetus see the contemporary television series "Lie to Me" and "The Mentalist".

2 Jerome Frank *Courts on Trial: Myth and Reality in American Justice* (Princeton University Press, Princeton, 1949) at 21.

3 *Kinloch v Young* [1968] 2 Lloyd's Rep 403 at 431 per Lord Pearce.

judges and juries in seeing and hearing witnesses. Juries are encouraged to use an assumed facility for assessing the demeanour of witnesses for the purpose of deciding whom they believe.

No-one doubts that by one means or another fact-finders must decide who is telling the truth. Nor is there any lack of tools to help them in their task. The tools include the internal consistency of a witness's account; consistency with what the witness has said on other occasions; comparison with contemporaneous records and the other evidence in the case; the witness's verbal response when confronted with internal and external inconsistencies; the witness's general character and credibility irrespective of the current dispute; and the plausibility of the account given. Plausibility entails the need to consider how likely it is that people would act in the way the witness has suggested. Thus far is reasonably uncontroversial.⁵

The question is whether the fact-finder's tools also include the witness's demeanour. To that question the social sciences have a clear answer.

Scientific evaluation of demeanour assessment

The assumption that demeanour betrays lying has been repeatedly tested by social scientists.

The tests begin from the premise that speakers provide two sources of information — verbal and non-verbal. The verbal information is the meaning of the speaker's words. The words can be transcribed as text. Appellate courts and other reviewers can assess the internal consistency, consistency with other evidence, and plausibility, of the text. In evaluating the textual meaning, they are as well placed as those who had been present when the speaker uttered the words in question.

The words used (the verbal content) can be contrasted with the way in which they were uttered (non-verbal content). A speaker simultaneously manifests a package of non-verbal information which is normally subconscious. It consists of facial expressions, bodily movements and vocal characteristics. The vocal characteristics, or “paralinguistic cues”, include pitch, pace, volume, timbre, expression and tremors. It is this package of face, body and voice that lawyers describe as “demeanour”.⁶

4 For England see the ruling of Judge Murphy in *R v D EWCC Blackfriars*, 16 September 2013. For Canada see *NS v R* 2012 SCC 72, [2012] 3 SCR 726.

5 As to prior inconsistent statements Par Anders Granhag and Leif A Stromwall suggest that in some circumstances repetition of a rehearsed false story can produce more consistency than ad hoc reconstruction from memory: “Deception detection based on repeated interrogations” (2001) 6 *Legal and Criminological Psychology* 85.

6 Lord Bingham described demeanour as a witness's “conduct, manner, bearing, behaviour, delivery, inflexion; in short, anything which characterises his mode of giving evidence but does not appear in a transcript of what he actually said”: “Assessing Contentious Eyewitness Evidence: A Judicial View” in Anthony Heaton-Armstrong and others (eds) *Witness Testimony: Psychological, Investigative and Evidential Perspectives* (Oxford University Press, Oxford, 2006) ch 18 at 333.

Many experiments have been conducted to gauge the extent to which observation of demeanour helps when assessing veracity. The participants in the experiment are typically divided into four categories – the “subjects”, whose reactions are the subject of the experiment; “interviewers” who ask questions in the course of a dialogue that is observed or read by the subjects; “respondents” who answer questions in the course of that dialogue; and “experimenters” who subsequently survey the subjects to see what conclusions they came to.

By prior arrangement some of the respondents have been asked to lie while others are asked to tell the truth. The subjects are not party to that arrangement. In the simplest form of the experiment, two groups of subjects are asked to assess the truthfulness of the respondents. One group of subjects is permitted to see and hear the interviews, and thus to assess demeanour. The other group is confined to reading a transcript of the interviews.

Both groups are asked to decide which respondents were telling the truth. A comparison is then made to see whether those subjects who had enjoyed the opportunity to assess demeanour fare better than those confined to the transcript. Variants of the basic model focus on particular types of interview (e.g. between lawyer and witness or between investigator and suspect), particular settings (e.g. laboratory or courtroom), particular cues from respondents (e.g. visual, audio or both), and particular experience, attitudes and perceptions on the part of the subjects (e.g. students, jurors, lawyers, judges, law enforcement officers).

The result of these experiments can be summarised as follows:

- Behavioural cues popularly thought to be associated with lying - posture, head movements, shifty eyes, gaze aversion, fidgeting, and gesturing - have no correlation with dishonesty or lack of credibility.⁷
- Facial information and other paralinguistic cues are not effectively utilised in the detection of truth and lying when the speaker is attempting to conceal the lie.⁸
- Although there can be physical cues to deceit — for example, reduction in gestures, changes in pitch, and less vocal involvement — these are so subtle that they are not perceptible to the ordinary person. Changes in pitch, for example, can be measured only with sophisticated equipment.⁹

7 M Stone “Instant Lie Detection? Demeanour and Credibility in Criminal Trials” [1991] Crim LR 821 at 829. See also Heaton-Armstrong and others, above n 6, at 28–30.

8 Glenn Littlepage and Tony Pineault “Verbal, Facial, and Paralinguistic Cues to the Detection of Truth and Lying” (1978) 4 Pers Soc Psychol Bull 461 at 463.

9 Heaton-Armstrong and others, above n 6, at 28.

- Those few physical cues that are displayed by a person telling a lie are due to increased cognitive demand (that is, the concentration required in order to tell a false story), rather than nervousness.¹⁰
- Because nervousness is normal for a witness under interrogation, it is likely to result in “Othello error”. Othello error is the false interpretation of interrogation stress as insincerity.¹¹
- The witnesses most likely to be believed, whether or not truthful in fact, are those who appear confident and open, have a good memory for peripheral detail, and are attractive. Those who exhibit nervousness and hesitancy are less likely to be believed, especially if they appear unsavoury and unattractive.¹²
- Visual cues of an interview serve primarily to distract, lowering the proportion of accurate decisions by the observer. Interview situations in which a respondent is motivated to deceive may be more accurately judged when the interview is not directly observed; non-verbal information actually diminishes the ability to detect deceit.¹³
- Suspicious interviewers tend to view responses as deceptive because a suspicious interrogation distorts observers’ perceptions.¹⁴
- Due to the so-called “halo effect”, a perceived good or bad quality in a person will tend to colour all judgments pertaining to that person. Consequently, once a positive or negative impression is formed, this is likely to attach to all the evidence of that witness. People do not tend to differentiate between different parts of a testimony.¹⁵

10 Heaton-Armstrong and others, above n 6, at 29–30.

11 Paul Ekman *Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage* (Norton, 1985) at 170; and Charles F Bond Jr and William E Fahey “False suspicion and the misperception of deceit” (1987) 26 *British Journal of Social Psychology* 41 at 41.

12 Peter McClellan “Who is telling the truth? Psychology, common sense and the law” (2006) 80 *ALJ* 655 at 661–662.

13 Norman RF Maier and James A Thurber “Reliability of Judgments of Deception When an Interview is Watched, Heard, and Read” (1968) 21 *Personnel Psychology* 23 at 23.

14 Olin Guy Wellborn III “Demeanour” (1991) 76 *Cornell L Rev* 1075 at 1080.

15 McClellan, above n 12, at 662.

- Where respondents are induced to engage in real deception, subjects attempting to detect deception are unable to do so whether observing live, with video, with audio, or from transcript.¹⁶
- An analysis of psychological studies of deception detection consistently shows that most people cannot do better than chance in discerning lies under laboratory conditions. The face does not seem to give away deception cues and may provide misleading information. Most of the results fall in the 0.45 to 0.60 range with a chance level of 0.5.¹⁷ In other words the chance of a correct answer is no better than the toss of a coin.
- Although most people cannot do better than chance in detecting falsehoods, most confidently believe that they can.¹⁸

The result of these experiments has been summarised by an American commentator, Professor Olin Wellborn, in the following terms:¹⁹

Taken as a whole, the experimental evidence indicates that ordinary observers do not benefit from the opportunity to observe nonverbal behaviour in judging whether someone is lying. There is no evidence that facial behaviour is of any benefit; some evidence suggests that observation of facial behaviour diminishes the reliability of lie detection. Nor do paralinguistic cues appear to be of value; subjects who receive transcript consistently perform as well as or better than subjects who receive recordings of the respondent's voice. With respect to body cues, there is no persuasive evidence to support the hypothesis that lying is accompanied by distinctive body behaviour that others can discern.

Wellborn's summary of scientific findings was published in 1991. Despite an outpouring on the subject since, it is difficult to find anything to the contrary in either social science literature or academic legal commentary. In one study, observers who could both see and hear senders were found to be more accurate in detecting deceit than those who could only see, and also more accurate than those who could only hear, the sender.²⁰ But in the same study the authors' overall conclusion supported Wellborn's thesis, namely that "rates of lie detection vary within a narrow range ... within a few points of 50%".²¹ Again, the detection rate was found to be no better than a coin toss.

16 Gerald R Miller, Norman E Fontes and Arthur Konopka "The Effects of Videotaped Court Materials on Juror Response: Final Report" (Michigan State University, 1978) at 11–42.

17 Miron Zuckerman, Bella M DePaulo and Robert Rosenthal "Verbal and Nonverbal Communication of Deception" (1981) 14 *Advances in Experimental Social Psychology* 1 at 39–40.

18 Ekman *Telling Lies*, above n 11, at 162; and Glenn E Littlepage and Martin A Pineault "Detection of Deceptive Factual Statements from the Body and the Face" (1979) 5 *Pers Soc Psychol Bull* 325 at 328.

19 Wellborn, above n 14, at 1088.

20 Charles F Bond Jr and Bella M. DePaulo "Accuracy of Deception Judgments" (2006) 10 *Pers & Soc Psychol Rev* 214 at 226.

21 Bond and DePaulo, above n 20, at 231.

A recent article that appeared to suggest that science had relevantly moved on since 1991 turns out on closer inspection to be concerned with factors unconnected with demeanour (“private information”, verbal response to inconsistencies, information from third parties, physical evidence, solicited confessions, illogical answers, “truth bias” and “lie bias”).²² No-one doubts that on those topics science has progressed over the last 20 years or that they have an important bearing on a judge or jury’s ability to detect deceit. But those matters have nothing to do with demeanour. For present purposes the sole point is that attempting to assess demeanour does not help, a point which the commentator in question accepted.²³ Over the last 20 years the academic support for Wellborn’s conclusion has been widespread and, at least to this writer’s knowledge, unquestioned.²⁴

One should be cautious before jettisoning a popular belief which has been cherished by so many for so long. But it is hard to argue with science where the results have been independently repeated many times over the past 50 years. The overwhelming conclusion must be that demeanour is not a useful guide to veracity.²⁵ There are no observational advantages when assessing the honesty of a witness’s evidence. Those confined to reading the transcript will do just as well. Of particular concern is the fact that so many sincerely believe that they are capable of assessing veracity through demeanour. The widespread belief that they can is a popular fallacy.

None of this would matter if the demeanour fallacy remained in the laboratory. Unfortunately, misconceptions on the subject have far-reaching consequences. Three warrant further discussion – (1) understanding on the part of professional fact-finders, (2) jury methods, and (3) appellate deference to credibility findings at first instance.

(1) Understanding on the part of professional fact-finders

Professional fact-finders include judges, tribunals, arbitrators, and the police. Not all understand their own limitations. It is not at all unknown, for example, to find judges still basing their decision, in whole or in part, upon their assessment of a witness’s demeanour.

22 Max Minzner “Detecting Lies Using Demeanor, Bias, and Context” (2008) 29 *Cardozo L Rev* 2557.

23 Minzner, above n 22, at 2566. See also his conclusion that “[w]hile *liars do not give off demeanor cues*, they do tell stories that are less logical, less consistent, and contain few details than those of truth-tellers” (at 2569) and that “[w]hile *the newest results on lie detection support the now-traditional view in legal academia that demeanor is not a valuable tool in making credibility decisions*, they undermine the further conclusion that accurately detecting lies is impossible and, as a result, we should view credibility decisions by juries as no better than a coin flip” (at 2578) (emphasis added). Clearly they are not a coin flip. There are many ways of assessing credibility without having to rely on demeanor.

24 In support see Jeremy A Blumenthal “A Wipe of the Hands, A Lick of the Lips; The Validity of Demeanor Evidence in Assessing Witness Credibility” 72 *Neb L Rev* 1157 (1993); Bond and DePaulo, above n 20, at 231; and the huge collection of supportive literature collected in Minzner, above n 22, at 2563 note 40.

25 Failure to distinguish between demeanour and factors which have nothing to do with demeanour (“private information”, verbal response to inconsistencies, information from third parties, physical evidence, solicited confessions, illogical answers, “truth bias” and “lie bias”) undermines the suggestion by Minzner, above n 22, at 2564, that science has relevantly moved on since the Wellborn and Blumenthal articles: above n 14 and 24.

Most professional fact-finders receive professional training in some form. At least in the case of judges and arbitrators, professional training does not currently include warnings against reliance on demeanour.²⁶

Given that assistance from demeanour is illusory, one fears that there must have been many erroneous decisions by judges and others over the years.

(2) Jury methods

Society's most critical factual issues are delegated to juries. As the legal laity, they are routinely warned about the traps that lie ahead for the unwary. The warnings include such matters as suspect identification, delays in making sexual complaints, incriminating evidence from accomplices, failed alibis, the significance of proven lies and previous convictions.

When it comes to demeanour, however, New Zealand juries are given no warning at all. On the contrary, it is common for judges to encourage them to rely on it.²⁷ So a trial judge will commonly say something to the jury along these lines:

It is for you to decide what evidence you accept and what evidence you reject. In deciding whether or not to believe a witness, don't be afraid to fall back on your own experience in dealing with the people you encounter in everyday life. Each of you is already experienced in deciding whom and what to believe.

Of course such a direction is literally consistent with reliance on other sources such as internal consistency, contemporaneous documents, plausibility, and the like. But the average juror will almost certainly interpret a direction in those terms as encouragement to invoke an assumed facility for assessing honesty through demeanour. The false message is that through ordinary life experience, each of us has become equipped to assess veracity through demeanour.

The trap is compounded by the fact that people who do not show the emotional behaviour popularly expected of them are likely to be disbelieved.²⁸ For example juries think they know how a genuine rape victim will behave. Visibly upset rape complainants are seen as more credible than those who appear calm or relaxed.²⁹ Overlooked is the fact that an appearance of calmness may be no more than a mask used as a means of coping. Misunderstandings of this kind may help to explain those rape acquittals which otherwise seem to defy rational explanation.

26 The writer's experience is confined to training for judges and arbitrators. He cannot speak for other professional fact-finders but understands that the New Zealand Police are warned against reliance on demeanour.

27 Suzanne Blackwell "Child Sexual Abuse on Trial" (PhD thesis, University of Auckland, 2007) at 226 notes that in all nine of the child sexual abuse trials she examined the judges emphasised witness demeanour.

28 Mary R Rose, Janice Nadler and Jim Clark "Appropriately Upset? Emotional Norms and Perceptions of Crime Victims" (2006) 30 *Law Hum Behav* 203.

29 Elisabeth MacDonald and Yvette Tinsley (eds) "From 'Real Rape' to Real Justice: Prosecuting Rape in New Zealand" (Victoria University Press, Wellington, 2011) at 371; Blackwell, above n 27, at 202 and see generally Ellen Wessel and others "Credibility of the Emotional Witness: A Study of Ratings by Court Judges" (2006) 30 *Law Hum Behav* 221.

On the other hand there must have been countless witnesses who were believed against all odds simply because they knew how to present themselves. Psychopaths and confidence tricksters are notoriously convincing. So too repeat offenders, who might be expected to give their evidence with growing confidence as their history of apprehensions and trials grows.

It is hard to escape the conclusion that jury misconceptions about demeanour will also have produced many miscarriages of justice over the years.

(3) Appellate deference to first instance credibility findings

The legal system requires that on occasion the factual conclusions of a jury, trial judge, or tribunal, will be re-examined later by others. Typically the re-examination arises in the context of an appeal or judicial review. Occasionally it will be an extra-curial application for a pardon, for a referral back to the Court of Appeal, for compensation for wrongful imprisonment, or for some other such purpose. Decision-makers at the second stage can be usefully described as “transcript-reviewers” when they rely on a transcript rather than personal observation of the original witnesses.

A transcript is usually limited to the text, that is to say the verbal content, of testimony. What it does not normally include are the non-verbal cues. Very occasionally, non-verbal cues are expressly read into the transcript by judge or counsel (“witness uses hands to indicate a distance of about half a metre”). Most of the time, however, non-verbal cues go unrecorded.

Although not articulated in these terms, appellate courts have always assumed that unrecorded non-verbal cues might have played a crucial part in the assessment of oral evidence at first instance. The corollary is thought to be that appellate courts should refrain from intervening on matters of credibility except in the most exceptional of cases.³⁰ As the Supreme Court of New Zealand recently put it, “... the appeal court must of course take full account of the disadvantage it may well have in making an assessment of the honesty and reliability of witnesses on the sole basis of the transcript of the oral evidence.”³¹ The result is that “[t]raditionally, both in civil and criminal cases, credibility findings at first instance have been very difficult to impugn”.³²

All this has been built on the unexamined assumption that there are advantages in seeing and hearing witnesses. We have already noted that the assumption is groundless when assessing veracity – for that purpose personal observation of the witness brings no advantages at all. But might there be other advantages? Four candidates – (a) reliability, (b)

30 For recent examples, see *Austin, Nichols & Co Inc v Stichting Lodestar* [2007] NZSC 103, [2008] 2 NZLR 141 and *R v Matenga* [2009] NZSC 18, [2009] 3 NZLR 145, along with *Nalder & Biddle (Nelson) Ltd v C & F Fishing Ltd* [2007] 1 NZLR 721 (CA) at [88]; *Stemson v AMP General Insurance (NZ) Ltd* [2006] UKPC 30, [2007] 1 NZLR 289 at [12]; *R v Aranui* [2007] NZCA 354 at [19]; and *R v T* [2007] NZCA 296 at [25].

31 *R v Matenga*, above n 30.

32 *R v Munro* [2007] NZCA 510, [2008] 2 NZLR 87 at [76].

persuasiveness (c) meaning and (d) opportunity to ask questions – need to be considered in turn.

(a) *Non-verbal cues as a guide to reliability*

Sincere witnesses can still be wrong. Consequently the question whether to accept a witness's evidence turns on perceptions of its reliability as well as its honesty. Even if direct observation of the witness does not help in assessing honesty, might it help in assessing reliability?

One of the English Law Lords was convinced that it does. In 1919 Lord Shaw remarked that:³³

Witnesses without any conscious bias towards a conclusion may have in their demeanour, in their manner, in their hesitation, in the nuance of their expressions, in even the turns of the eyelid, left an impression upon the man who saw and heard them which can never be reproduced in the printed page.

Here too, science suggests otherwise. The issue has frequently been addressed in experiments testing eyewitness identifications. The results are consistent. A juror's opportunity to see and hear the identifying witness give evidence has no correlation whatsoever with the juror's detection of unreliability. That is so even where the witness is subjected to cross-examination by an experienced trial lawyer. In this context, jurors cannot outdo chance.³⁴ Personal observation of the witness holds no advantages when deciding whether or not to accept that witness's identification of a suspect.

That conclusion is not to be confused with witness confidence at the time that an identification is first carried out. Confidence, or lack of confidence, at that time is entirely relevant.³⁵ But whether a witness had been confident when he or she had first identified the suspect is a factual question which must be proved or disproved in the conventional manner. Evidence can usefully be called on that subject from the identifying witness, or from others who had been present at the time, or both. It has nothing to do with the confidence displayed by a witness while giving testimony in court. The presence or absence of confidence while testifying in court can be positively misleading.³⁶ The distinction has been highlighted in a perceptive judgment in the New Zealand Court of Appeal,³⁷ albeit then diluted in the New Zealand Supreme Court.³⁸

33 *Clarke v Edinburgh Tramways* [1919] SC (HL) 35 at 36 per Lord Shaw.

34 A survey of experimental findings is found in Wellborn, above n 14, at 1080–1091.

35 Siegfried Ludwig Sporer and others "Choosing, Confidence and Reliability: A Meta-Analysis of the Confidence-Reliability Relation in Eyewitness Identification Studies" (1995) 118 *Psychological Bulletin* 315; N Brewer and GL Wells "The Confidence-Accuracy Relationship in Eyewitness Identification: Effects of Lineup Instructions, Foil Similarity and Target-Absence Base Rates" (2006) 12 *Journal of Experimental Psychology Applied* 11; and Daniel B Wright and Elin M Skagerberg "Postidentification Feedback Affects Real Eyewitnesses" (2007) 18 *Psychological Science* 172, collected in *R v Edmonds* [2009] NZCA 303, [2010] 1 NZLR 762 at [117].

36 *R v Edmonds*, above n 35, at [118] and [119].

37 *R v Edmonds*, above n 35, at [117]–[119].

38 *Harney v Police* [2011] NZSC 107 at [33]: "*too much weight should not be given* to this factor [confidence] especially when it is not an expression of confidence at the time the identification was first made" (emphasis added). But the scientific evidence is far simpler: confidence at any time other than the first identification should be given no weight at all.

Whether or not a witness has correctly identified a suspect does not appear to differ from any other question of fact for present purposes. If demeanour information is of no use for that purpose, it seems reasonable to infer that it will be of no greater assistance in assessing any other question that turns on a witness's perception and memory.³⁹ In short, direct observation of a witness does not seem to help in assessing that witness's reliability.

(b) *Non-verbal cues as a guide to logical persuasiveness*

Closely related to the *reliability* of a reported fact is the *persuasiveness* of a reported opinion. The two overlap but for present purposes it is useful to distinguish between the reliability of a witness as to fact (did the witness really see Smith stab Brown?) and the rational persuasiveness of an expert's opinion (should one accept the view that the blood on Smith's knife matched that of Brown?). For present purposes the question is whether, in approaching the latter task, the fact-finder might enjoy demeanour advantages unshared by a subsequent transcript-reviewer.

Faced with the competing opinions of two experts, the fact-finder will need to decide which is the more persuasive. Considered rationally, one would expect persuasiveness to turn on the verbal content of the testimony (i.e. content derived from the meaning of the words used in relation to the issue together with the verbally reported details of the expert's skill and experience). It could scarcely turn on the expert's presentational skills and appearance (resonance of voice, physical attractiveness, clothing, confidence, colour of hair, skill with whiteboards, etc). Attributes in the latter category would not seem to have any logical connection with the soundness of an expert's opinion.

Unfortunately there is strong evidence that in choosing between competing experts, jurors are heavily influenced by irrelevant distractions and that the more arcane the issue, the more powerful those distractions will be.⁴⁰ The role of irrelevant distractions can be illustrated by reference to confidence. Confidence is one of the most significant presentational factors jurors rely on when assessing expert evidence.⁴¹ A survey of New Zealand juries revealed that 58 per cent of jurors thought that highly confident experts were "more believable"⁴² and a third even went so far as to say that "good eye contact with the jury makes an expert more believable".⁴³

The problem is the lack of any evidence to suggest a useful correlation between witness confidence, on the one hand, and the reliability of the opinion expressed, on the other. Few would want to rely on a witness's self-assessment of their own understanding.⁴⁴

39 Wellborn, above n 14, at 1091.

40 For excellent reviews of the New Zealand scene see Blackwell, above n 27; and Emily Henderson and Fred Seymour *Expert Witnesses Under Examination in the New Zealand Criminal and Family Courts* (University of Auckland Press, Auckland, 2013).

41 Henderson and Seymour, above n 40, at 12 and 13; Blackwell, above n 27, at 96–97; and Joseph Sanders "Jury deliberation in a complex case: *Havner v Merrell Dow Pharmaceuticals*" (1993) 16 Justice System Journal 45.

42 Blackwell, above n 27, at 222–223.

43 Blackwell, above n 27, at 223.

44 Henderson and Seymour, above n 40, at 14; and Cheryl Thomas *Are Juries Fair?* (Ministry of Justice (UK), Research Series 1/10, February 2010) at 39.

There is some correlation between hesitation pauses and either the difficulty of the topic,⁴⁵ or anxiety,⁴⁶ or both. Presumably confidence rises with frequency of appearances as an expert witness in successive cases. Yet frequency in the giving of evidence has no logical connection with the validity of the opinion expressed.

The role of such distractions can be illustrated by susceptibility to coaching. Experts can be coached in appearance and presentational skills. If the views of US lawyers are anything to go by, jurors are much more heavily influenced by presentation and appearance than the strength of the actual evidence itself.⁴⁷ This creates the opportunity for manipulating the jury by suggesting that the expert “dress smartly but not ostentatiously”,⁴⁸ appreciate that presentation is as important as actual expertise,⁴⁹ seek to appear confident and use an engaging manner,⁵⁰ give clear and simple explanations⁵¹ and avoid signs of emotion or argumentativeness.⁵² Yet none of these behaviours tells us anything useful about the validity of the opinion itself. It seems unreasonable to expect juries to distinguish between content and presentation without specific warnings on that topic.

Whether judges fare any better than juries in this respect is an open question. One might expect judges to be more alert to the fact that many expert witnesses are experienced performers. There is some support for the view that judges are less likely to be swayed by an expert’s confidence and other non-verbal cues.⁵³ However another study suggests that, at least in the United States, judges are no better than jurors at focusing on content rather than presentation.⁵⁴ In the view of a senior Australian Judge “the Judge is as likely to be persuaded by the more articulate and persuasive personality as by his or her own rational analysis of the conflicting opinions”.⁵⁵

Overall it is difficult to escape the conclusion that in assessing the persuasiveness of an expert opinion, non-verbal cues are more likely to be a hindrance than an aid. That is not

45 SR Rochester “The Significance of Pauses in Spontaneous Speech” (1973) 2 *Journal of Psycholinguistic Research* 51 at 70.

46 Rochester, above n 45, at 73.

47 Joseph Sanders “The Merits of the Paternalistic Justification for Restrictions on the Admissibility of Expert Evidence” (2003) 33 *Seton Hall L Rev* 881 at 913–914 as cited in Henderson and Seymour, above n 40, at 13 note 68.

48 Paul Newman “Giving a Performance” (2006) Feb CN 109 at 110 as cited in Henderson and Seymour, above n 40.

49 Nicola Cunningham “Emergency Physicians as Expert Witnesses: from frontline wise to court room woes” (2009) 21 *Emergency Medicine Australasia* 503 at 506. See also a study showing that jurors rated an expert’s evidence as more credible and comprehensible when presented in person than when read verbatim by a lawyer: Daniel Jacobovitch and others “Juror responses to direct and mediated presentations of expert testimony” (1977) 7 *Journal of Applied Social Psychology* 227.

50 Newman, above n 48, at 109.

51 Newman, above n 48, at 110.

52 Newman, above n 48, at 109–110.

53 Dawn McQuiston-Surrett and Michael J Saks “The testimony of forensic identification science: what expert witnesses say and what fact-finders hear” (2009) 33 *LHUMB* 436 at 450–451; and MacDonald and Tinsley, above n 29, at 242.

54 McQuiston-Surrett and Saks, above n 53, at 440 and 450–451.

55 Geoffrey L Davies “Current Issues — Expert Evidence: Court-Appointed Experts” (2004) 23 *CJQ* 367 at 370.

to suggest that there should be any departure from the current practice of presenting expert evidence in person. Current practice permits cross-examination, is more likely to hold a jury's attention, and is probably more comprehensible. What the studies do suggest, however, is that transcript-reviewers should not be deterred by any thought that jurors may have enjoyed a superior opportunity for assessing reliability. Nor is there any ground for believing that trial judges will perform better in that respect.

(c) *Non-verbal cues as a guide to meaning*

Up to this point the discussion has been confined to a witness's honesty, the reliability of his or her observations of fact, and the weight to be attached to his or her opinions. We have seen that in general a transcript-reviewer will be as well-placed to assess those matters as the original fact-finder.

A different topic altogether is a witness's meaning. It is one thing to decide whether a witness should be believed (was the witness lying when she said the car was speeding?), another altogether to decide what the witness meant (what did the witness mean by the word "speeding" when describing the driving?).

It was this distinction which the New Zealand Court of Appeal had in mind when it said in *Munro*:⁵⁶

Tone of voice is important in conveying meaning and this cannot be captured in a written transcript. For example, the concession "it's possible" could, depending on the tone, have a meaning ranging from "that's quite likely" to "yes, but pigs might fly". Pauses can be significant and are rarely shown in a transcript. Gestures and facial expressions can also add to meaning and are often not recorded in a transcript unless vital to a witness' answer and even then, when digital recording systems are in operation, only when specifically read into the record by counsel or the judge.

In other words non-verbal cues convey meaning,⁵⁷ even if they do not usefully convey whether the intended meaning should be accepted as genuine, reliable or persuasive.

Might this be a reason for deferring to the special advantages of the original fact-finder? The answer must be "yes" in certain cases. The cases will be those in which (i) the transcript has left what the witness meant unclear (by "possible" did she mean only that in theory nothing can ever be ruled out or did she mean that it was a realistic possibility?); (ii) it is conceivable that non-verbal cues could have assisted in resolving the ambiguity (a dismissive intonation might have indicated the former); and (iii) it is conceivable that removing the ambiguity could have played a significant part in the outcome (as, for example, where the question is whether the evidence permitted a finding of reasonable doubt).

All three of those conditions would have to be satisfied before lack of non-verbal cues could place a transcript-reviewer at a disadvantage. It seems unlikely that this will occur

56 *R v Munro*, above n 32.

57 It has even been claimed that nonverbal communication represents two-thirds of all communication: K Hogan and R Stubbs *Can't Get Through: 8 Barriers to Communication* (Pelican Publishing, Grenta, 2003) at 162.

often. The weight to be placed on testimony normally turns on whether a witness is to be believed, or can be relied upon, not what he or she means. As a senior English judge noted:⁵⁸

There are of course occasions when the simple language used by a witness, reproduced on the printed page, may be ambiguous although its meaning is, to anyone who hears the answer, apparent. For example, a witness may assent to a suggestion of cross-examining counsel with a shrug of the shoulders to indicate that though theoretically possible the suggestion is in practical terms absurd, or the assent may come after a long pause and in a manner indicating full acceptance. But responsible counsel do not allow a case to proceed on the basis of a single answer without making sure that the witness's position is clearly established, and were they to do so the judge would intervene to ensure that there was no room for doubt. The occasions on which the flavour of a witness's evidence is more accurately derived from the inflexion attached to a single answer, or his demeanour when giving it, than from the gist of a series of answers must, I suggest, be few.

(d) *The opportunity to ask questions*

The original fact-finder has the opportunity to ask clarifying questions. For all practical purposes a transcript-reviewer does not. Although appellate courts normally have the power to require a witness to answer further questions at the appeal hearing on the application of either party,⁵⁹ this is rarely used in practice.

Lack of opportunity to ask further questions does not appear to disadvantage an appellate court. There is no doubt that lie detection improves if the correct questions are asked.⁶⁰ But the point is that where the original fact-finder had used that opportunity, both the questions posed, and the answers given, will be recorded in the transcript. Since the two tribunals finish up with precisely the same information on which to base their decision, it is difficult to see how one could have any advantage over the other in that respect.

Conclusions over credibility deference

The question currently under consideration is whether, in the assessment of oral evidence, a fact-finder at first instance has advantages denied to a subsequent reviewer of the transcript. The scientific evidence on that subject can be summarised as follows:

- Where the question is whether the testimony was honest, reliable or persuasive, the transcript-reviewer is at no disadvantage. Personal presence while the witness gives evidence brings no benefits. On occasions it may even serve to distract from the content that matters.
- Where the issue is what the witness *meant* by his or her evidence, the transcript-reviewer may be at a disadvantage if, but only if, (i) the transcript has left what the witness meant unclear, (ii) non-verbal cues might have removed that ambiguity and (iii) it is conceivable that removing the ambiguity had played a significant part in arriving at the original fact-finder's conclusions.

58 Lord Bingham, above n 6, at 336.

59 As, for example, in the Court of Appeal (Civil) Rules 2005, r 45(1)(a).

60 See Aldert Vrij and Par Anders Granhag "Eliciting cues to deception and truth: What matters are the questions asked" (2012) 1 Journal of Applied Research in Memory and Cognition 110.

It follows that the question whether a transcript-reviewer should defer to a fact-finder's assessment of testimony at first instance cannot be answered in the abstract. An essential first step is to identify the particular issue that falls for decision in the particular case.

In nearly all cases - criminal and civil - the factual issues come down to the honesty and reliability of the witnesses and the inferences to be drawn from the evidence as a whole. Indeed in most criminal cases the factual issues can be reduced to veracity alone ("I saw the accused stab the deceased"; "The accused could not have stabbed the deceased as he was with me in Hamilton at the time"). In such cases the original fact-finder could not have derived any benefit from seeing and hearing the witness. A transcript-reviewer is as well-equipped to determine veracity as the original fact-finder.⁶¹ The same is true when it comes to the witness's reliability ("the man I saw running away looked like the accused").

The position is different when the issue is what the witness meant by his or her evidence ("When the car came down the road it was speeding"; "The odds of someone other than the accused having a matching DNA are 100,000 to one"). What a witness meant by his or her evidence is not a question which arises often on appeal or other form of review. But where it does, transcript-reviewers must ask themselves whether the original fact-finder may have gained a better insight through non-verbal cues and whether this might be a reason in itself for upholding the first instance decision.

To summarise, the fact that there were findings on credibility at first instance is not a legitimate reason for an appellate court to decline to intervene except in certain instances turning on meaning.

Are there other reasons for upholding original fact-finders?

It does not follow that more appeals should be allowed. Although this article is concerned with demeanour, it would be wrong to leave the impression that correcting the demeanour fallacy should encourage appellate courts to overturn original decisions more frequently.

Of course in an ideal world everyone would receive perfect justice, or failing that, at least the best justice that humans might be capable of delivering. But there are sound policy reasons for reluctance to interfere with first instance decisions even where, contrary to traditional dicta, the reasons have nothing to do with demeanour. The reasons are (a) respect for the constitutional role of juries, (b) the need to maintain public confidence in first instance judges, (d) the need to limit the scope of appeal hearings and (e) the need to promote finality.⁶² These will be considered in turn.

61 After concluding that a first instance judge or jury could not have derived assistance from demeanour, but surveying other factors from which they could not have derived any advantage compared with an appellate court, Minzner, above n 22, inexplicably concludes at 2579: "As a result, we do not know whether this evidence supports or undermines the notion of appellate deference." Significantly, he was unable to identify any factor which might have placed an appellate court at a disadvantage.

62 A curious feature of the New Zealand Supreme Court decision in *Austin, Nichols*, above n 30 is that none of these considerations received mention. ???

(a) Respect for the constitutional role of juries

Juries have an important role in the justice system as a symbol of democracy and a safeguard against arbitrary or oppressive government.⁶³ This alone is sufficient reason for judges to exercise caution before interfering with their factual conclusions. It is not necessary to resort to supposed jury advantages in assessing demeanour as a reason for reluctance to intervene.

(b) *Public confidence in first instance judges*

Trial judges are the engine of the court system. Respect for the system can be maintained only if the public respects their work. As has been noted:⁶⁴

The cases indulging in de novo review represented bad judicial policy. To accord no finality to the trier's findings undermines the resolution of disputes. As Professors Wright and Miller explained: "Even in instances where an appellate court is in as good a position to decide as the trial court, it should not disregard the trial court's finding, for to do so impairs confidence in the trial courts and multiplies appeals with attendant expense and delay."⁶⁵

An appellate court may always be "in as good a position to decide as the trial court," in the sense that the traditionally disparaged "cold record" may be as good a basis for decision, including judgments of credibility, as the trial court's traditionally exalted opportunity to see the witnesses. De novo review of facts is nevertheless a bad idea, and appellate court rejection of trial court findings should continue to be limited to instances of clear error.

It would do little for confidence in the work of trial judges if appellate courts placed little or no weight on decisions at first instance and started the reasoning process from the beginning.

(c) *Limiting the scope of appeal hearings*

A third reason for reluctance to intervene is efficiency.

Efficiency is not the way in which the concern has been traditionally expressed. The priority afforded to first instance findings has been said to rest, at least in part, on the fact that the trial judge "observes, as we cannot observe, the drift and conduct of the case; and also that he has impressed upon him by hearing every word the scope and nature of the evidence in a way that is denied to any Court of appeal."⁶⁶ To similar effect it has been claimed that "It is the tableau that constitutes the big advantage, the text with illustrations, rather than the demeanour of a particular witness."⁶⁷

On its surface, the argument is presented as one of appellate modesty: trial courts alone have the opportunity to see the whole picture. And it is true that first instance courts spend days, weeks, and sometimes months, wading through a myriad of factual detail. But this welter of material is scarcely beyond the reach of appellate courts if they wanted it.

63 Julia Tolmie and Warren Brookbanks *Criminal Justice in New Zealand* (LexisNexis, Wellington, 2007) at 200; New Zealand Law Commission "Juries in Criminal Trials: Part 1: A Discussion Paper" (Preliminary Paper 32 Wellington, NZLC, 1998) at 18; and *R v Nazif* [1987] 2 NZLR 122 (CA) at 127.

64 Wellborn, above n 14, at 1095–1096.

65 Charles Alan Wright and Arthur R Miller *Federal Practice and Procedure* (West Publishing, 1971) vol 9 at §2587.

66 *Kinlock v Young* [1911] SC (HL) 1 at 4 per Lord Loreburn.

67 Patrick Devlin (Lord Devlin) *The Judge* (Oxford University Press, Oxford, 1979) at 63.

Using the transcript, they could in theory descend to the same level of detail and comprehensiveness as the trial court.

The real point is not that appellate courts are unable to traverse the whole case again based on the trial record but the sheer cost and inefficiency of doing so. So it has been pointed out that although an appeal could be a rehearing in the literal sense:⁶⁸

[t]he result would be to increase very substantially the burden on appellate courts, to delay the *finis litium* which is held to be in the public interest, to increase uncertainty (because a litigant could never be sure on what version of the facts a point of law was to be decided) and to increase the already frightening costs of litigation. These are, without doubt, ills to be avoided if possible, and it would in my view be a respectable rule that every litigant should be entitled to a full contest on the facts at one level only and that the facts should be open to review thereafter only if some glaring and manifest error could be demonstrated.

Society cannot afford the limitless pursuit of justice. Duplicating an exercise already undertaken in the trial court would be an unwarranted cost to the parties and the public. It would be an inefficient allocation of judicial resources. In a word, a comprehensive re-examination of the evidence would be inefficient.

(d) *Promoting finality*

A fourth reason for giving first instance decisions presumptive force is finality. There is a public interest in bringing litigation to an end. As has been pointed out:⁶⁹

The underlying public interest is ... that there should be finality in litigation and that a party should not be twice vexed in the same matter. This public interest is reinforced by the current emphasis on efficiency and economy in the conduct of litigation, in the interests of the parties and the public as a whole.

This must apply equally to appeals. If there is no presumption (albeit rebuttable) that first instance decisions will be upheld, the appellant has little to lose by starting afresh before the appellate court. The odds of success would be no worse than they had been before the case began in the lower court.

That approach may be contrasted with the position if first instance decisions are given presumptive force. Raising the threshold before the respondent can win in the appellate court discourages those appeals based on nothing more than the hope that in an area of respectable disagreement there must always be a chance that a second decision-maker will take the other option.

Of course an appellate court might seek to justify intervention on the basis that its decision-makers offer superior skill, experience and numbers. But if it were thought that judges should be preferred to juries, that senior judges should be preferred to junior ones, or that three judges should be preferred to one, it would be more efficient to commence the proceedings before a forum structured in the preferred way to begin with. There is little point in having first instance courts unless their decisions continue to count for something when taken on appeal.

68 Lord Bingham, above n 6, at 336.

69 *Johnson v Gore Wood & Co* [2002] 2 AC 1 (HL) at 31 per Lord Bingham.

The efficient allocation of resources therefore suggests that first instance decisions should be given presumptive force on appeal. The first instance decision should not be interfered with unless shown to be clearly wrong.

Conclusions on appellate readiness to intervene

There will be many occasions when an appellate court should decline to interfere with a first instance decision. But the reasons for declining to intervene do not include relative advantages in assessing credibility. The real reasons are the need to respect the constitutional allocation of responsibilities between judges and juries, maintain public confidence in first instance judges, limit the scope of appeal hearings, and promote finality.

Appellate courts have often backed the wrong horse when determining their own review criteria. On the one hand, one of their stated grounds for declining to intervene - credibility⁷⁰ – is specious. On the other, they have shown a disturbing readiness to intervene without at least articulating the countervailing considerations just outlined.⁷¹

A glib response would be that so long as appellate courts are slow to interfere, the reasons for the slowness matter little. However the danger is that articulation of faulty review criteria will produce faulty outcomes. In cases where there had been no oral evidence, appellate courts may be jumping in too quickly, needlessly substituting their own opinion for that of the court at first instance. Equally there is a risk that in cases that had involved oral evidence, appellate courts will be too slow to intervene due to the mistaken assumption that fact-finders at first instance are better equipped to assess credibility.

Progress in recognizing the demeanour fallacy

Social scientists have known about the demeanour fallacy for well over 40 years.⁷² Although it has taken longer for the news to reach lawyers, there have been distinguished sceptics for many years. As long ago as 1974 an English judge confessed this:⁷³

I question whether the respect given to our findings of fact based on the demeanour of the witness is always deserved. I doubt my own ability, and sometimes that of other judges, to discern from a witness's demeanour, or the tone of his voice, whether he is telling the truth. He speaks hesitantly. Is this a mark of a cautious man, whose statements are for that reason to be respected, or is he taking time to fabricate? Is the emphatic witness putting on an act to deceive me, or is he speaking from the fullness of his heart, knowing that he is right? Is he likely to be more truthful if he looks me straight in the face than if he casts his eyes on the ground, perhaps from shyness or a natural timidity? For my part I rely on these considerations as little as I can help.

Similar doubts were expressed by a senior Australian figure in 1979:⁷⁴

70 For example, *R v Matenga*, above n 30.

71 For example, *Austin, Nichols*, above n 30; reaffirmed in *K v B* [2010] NZSC 112 at [31]–[32].

72 See, for example, A Daniel Yarmey *The Psychology of Eyewitness Testimony* (The Free Press, 1979) at 169.

73 MacKenna B (Justice MacKenna) “Discretion” (1974) 9 *The Irish Jurist* 1 at 10.

74 AM Gleeson (later Gleeson CJ) “Judging the Judges” (1979) 53 *ALJ* 344. To similar effect see Richard Eggleston *Evidence, Proof and Probability* (Fred B. Rothman & Co, Littleton, 1978) at 163.

Reasons for judgment which are replete with pointed references to the great advantage which the trial judge has had in making the personal acquaintance of the witnesses seem nowadays to be treated by appellate courts with a healthy measure of scepticism. What might be called the Pinocchio theory, according to which dishonesty on the part of a witness manifests itself in a manner that does not appear on the record but is readily discernible by anyone physically present, seems to be losing popularity.

And in 1982 an English Court of Appeal judge made this admission:⁷⁵

So the main job of the judge of first instance is to decide the facts. How does he do it? When there is a conflict of evidence between witnesses, some judges believe that they can tell whether a witness is telling the truth by looking at him and listening to him. I seldom believed that ...

The two notable features about these comments is that they were made without apparent knowledge of parallel advances in social science and further that the judges involved were careful to make their comments off the bench.

Then in 2003 the High Court of Australia gave a judgment which acknowledged the scientifically demonstrated fallibility of most people when trying to detect lies. The long-standing principles of witness credibility were still thought to be relevant but it was acknowledged that the scientific evidence “tends to reduce the occasions where those principles are seen as critical”.⁷⁶

Four years later it was New Zealand’s turn in *R v Munro*.⁷⁷ After a scholarly survey of scientific conclusions on the subject, the New Zealand Court of Appeal acknowledged the advantage of seeing and hearing witnesses for the purpose of augmenting meaning via non-verbal cues.⁷⁸ However as to veracity it concluded:⁷⁹

These empirical and psychological studies show that what was once considered the privileged domain of the jury cannot be justified by reference to a jury’s special ability in these areas. Errors in assessing credibility may well be common, and it cannot in all cases be presumed that the deliberative process has corrected them. All the above means that it is no longer possible to make blanket assertions about witness credibility being the province of the jury.

Munro is the first serious discussion of the point in New Zealand. However cautiously expressed, there is no disguising the change of direction.

That is not to suggest that the journey is complete. Two recent decisions of the Supreme Court of New Zealand show the distance still to travel. A matter of weeks after *Munro*, the Supreme Court gave its own decision in *Austin, Nichols & Co Inc v Stichting Lodestar*.⁸⁰ There, the Court was required to consider the basis on which appellate courts ought to review findings made at first instance. At several points,⁸¹ the Supreme Court

75 Patrick Browne (Browne LJ) “Judicial Reflections” (1982) 35 Current Legal Problems 5.

76 *Fox v Percy* [2003] HCA 22, (2003) 214 CLR 118 at [30]–[31] per Kirby J. Three years later Kirby J, with whom Gleeson CJ agreed, rejected “earlier appellate deference to erroneous fact-finding by primary judges ... [based on] the ‘subtle influence of demeanour’ that could have affected the primary judge’s conclusion, even though no express reference was made to such consideration”: *CSR Ltd v Della Maddalena* [2006] HCA 1, (2006) 224 ALR 1 at [19] and [23].

77 *R v Munro*, above n 32.

78 See the passage quoted from *R v Munro* in the text to above n 32.

79 *R v Munro*, above n 32, at [82] per Glazebrook, Chambers, Arnold and Wilson JJ.

80 *Austin, Nichols*, above n 30.

81 *Austin, Nichols*, above n 30, at [5], [13]–[14] and [17].

repeated as a given the traditional view that where credibility mattered, an appellate court should defer to the first instance judge's opportunity to assess demeanour. It was seemingly unaware of *Munro*, scientific findings on the subject, or the existence of other policy reasons for declining to intervene.⁸²

Austin was followed by *R v Matenga*⁸³ in which demeanour again fell for consideration in the Supreme Court. This time the context was the proviso to s 385(1) of the Crimes Act 1961. The proviso permits the Court of Appeal to uphold a verdict notwithstanding trial irregularities if it considers that no substantial miscarriage of justice occurred. Again the Supreme Court was seemingly unaware of *Munro* and scientific findings. As to credibility the Court contented itself with the following:⁸⁴

In coming to its conclusion concerning the inevitability of the verdict, the appeal court must of course take full account of the disadvantage it may well have in making an assessment of the honesty and reliability of witnesses on the sole basis of the transcript of the oral evidence.

What *Austin* and *Matenga* demonstrate is that centuries of cherished belief will not vanish overnight. But in fairness to the Supreme Court, no-one appearing there seems to have yet argued that it is time to revisit the role of witness demeanour. One might reasonably expect that when they do, the Court will respond.

A possible way forward

If science tells us that the honesty and reliability of a witness cannot be assessed by observing the way in which the witness gives his or her evidence, three consequences seem to follow.

(1) *Assessing veracity at first instance*

Those who are called upon to assess evidence must keep their own limitations in mind. A speaker's credibility cannot be usefully judged through his or her demeanour. This simple fact should be emphasised in the training of judges, arbitrators and other professional fact-finders.

The question is not whether a fact-finder should make a finding on credibility but the sources that can usefully be employed for that purpose. The true indicators of credibility were outlined earlier (essentially internal consistency, comparison with other evidence, general creditworthiness and plausibility).⁸⁵

It has generally been assumed that it is safe to leave it to the fact-finder's intuition to choose the sources that will assist in determining credibility. Unfortunately it is intuition that gave us reliance on demeanour. Deciding on credibility is a critical task for fact-finders. It is sufficiently important to warrant some level of instruction on the way to go about it.

82 See the above discussion of policy reasons for upholding first instance decisions.

83 *R v Matenga*, above n 30.

84 *R v Matenga*, above n 30, at [32].

85 Above page 2.

(2) *Jury warnings*

Far from encouraging reliance upon demeanour, we should expressly warn juries against it. It is true that jury directions are not always understood, retained and applied.⁸⁶ But this is hardly an excuse for the misleading nature of current jury directions on the subject.⁸⁷ Better a sound, if imperfectly applied, direction than a direction that is bound to mislead for those who seek to apply it.

As with all jury directions, warnings in relation to demeanour must be tailored to the facts of the particular case. For example guidance in assessing credibility will differ according to whether one is commenting on a witness unconnected with the parties, a sexual complainant, or an accused. However in general terms directions to a jury might usefully include something along these lines:

It is for you alone to decide what the facts are in this case.

In the course of deciding what facts you accept, you may need to decide who is telling the truth. In deciding whether or not to believe someone, a number of questions are usually thought to be helpful. Has the witness said anything inconsistent in her evidence to you or on some other occasion? How does her evidence compare with what other witnesses have said? How does it compare with what was noted in emails and other documents at the time? Is it likely that people would have behaved in the way she has suggested?

Those are the sort of questions that you might find helpful when deciding whether or not to believe a witness. But the one thing I must warn you against is relying on the way a witness looks and sounds when giving evidence. This can be a real trap. We all tend to think we know how honest and dishonest witnesses will look and sound. It is tempting to think that if a witness hesitates, looks down, mumbles, scratches her nose, or looks nervous, she must be lying. And equally it is tempting to think that someone confident, open, and quick in their answers will be telling the truth. But studies have shown time and time again that impressions of that kind are misleading. There can be any number of reasons for witnesses to present themselves in a particular way when they have to give evidence. Most of these reasons have nothing to do with their honesty and reliability.

So do not fall into the trap of trying to use the way a witness looked and sounded while giving evidence in this courtroom. Instead, rely on the consistency of what they said, how their evidence compares with the other evidence and the inherent likelihood of what they said.

(3) *Appellate deference to first instance findings*

The traditional appellate approach to assessment of oral testimony is misconceived. To decline to intervene on the ground that the trial judge or jury enjoyed special advantages in assessing the honesty and reliability of witnesses is to misunderstand what happens when testimony is received.

The fresh principles an appellate court ought to apply are as follows:

86 Warren Young, Neil Cameron and Yvette Tinsley *Juries in Criminal Trials: Part Two: A Summary of Research Findings* (NZLC PP37, v2, 1999) at 13 and 53–59; Gary Edmond “Is reliability sufficient? The Law Commission and expert evidence in international and interdisciplinary perspective: Part 1” (2012) 16 *E&P* 30 at 51 as cited in Young, Cameron and Tinsley at 19; MacDonald & Tinsley, above n 29, at 237–239 and 371–372; and Minzner, above n 22, at 2578 note 102.

87 As Minzner, above n 22, at 2578 seems to imply.

- The appellate court is as well-placed to assess honesty, reliability and persuasiveness as the original fact-finder. The traditional mantra that the original fact-finder had an advantage in seeing and hearing the witnesses has no scientific validity.
- Wherever the transcript is capable of more than one meaning, the appellate court must ask itself whether the original fact-finder might have been better placed to resolve the ambiguity. It is unlikely that this will arise often given that assessing testimony is usually concerned with credibility and reliability rather than meaning; the limited circumstances in which an appeal could turn on such an ambiguity; and the limited circumstances in which ambiguities could be resolved by non-verbal cues.
- There are strong policy reasons for declining to interfere with first instance decisions without good cause but they do not include credibility. The valid reasons are the need to respect the allocation of responsibilities between judges and juries, maintain public confidence in first instance judges, limit the scope of appeal hearings, and promote finality.

Conclusions

Credibility assessment is central to the work of judges, juries, tribunals, arbitrators and other decision-makers. There are ample tools with which to make that judgment. If they are insufficient, there is always the onus of proof to fall back on.

Unfortunately the one source that cannot be usefully resorted to is demeanour. Judges and other fact-finders should keep this in mind when listening to evidence. Juries should be expressly warned about it. Appellate courts should change their approach to it.

Traditionally, the credibility of witnesses, and appellate standards of review, have been the province of judges rather than legislators. One hopes that the Courts will respond to developments in understanding demeanour. If they do not, Parliament may need to step in, as it has been obliged to do on other occasions when the Courts have dragged their feet.⁸⁸

⁸⁸ As, for example, in the Evidence Act 1990 and its predecessors where the New Zealand Parliament intervened over such matters as propensity (s 40), spousal privilege (s 53 combined with omission from ss 54 to 59), corroboration (s 121), and delays in sexual complaints (s 127).