# Assessing Truthfulness on the Witness Stand: Eradicating Deeply Rooted Pseudoscientific Beliefs about Credibility Assessment by Triers of Fact

*Brent Snook, Meagan I. McCardle, Weyam Fahmy and John C. House*[*]

*The Supreme Court of Canada (SCC) ruled in R. v. S. (N.) (2012) that a witness' face provides useful cues to deceit that are important for trial fairness, and that the need to view a witness' face while testifying ''is too deeply rooted in the criminal justice system to be set aside absent compelling evidence'' (para, 27). In this commentary, we present compelling empirical evidence that (a) the vast majority of cues to deception are too faint for reliable deception detection, (b) most facial expressions and other non-verbal cues are unrelated to deception, (c) people, including professionals in legal disciplines (e.g., judges, lawyers), are unable to detect deception barely beyond chance levels, and (d) it is not possible to improve deception detection with training programs that focus exclusively on non-verbal cues. The fact that existing empirical data fails to align with the venerated belief surrounding deception detection raises serious concerns about the reliance on pseudoscientific beliefs in SCC decision-making.*

---

*La Cour suprême du Canada (CSC) a statué dans R. c. S. (N.) (2012) que le visage d'un témoin fournit des indices utiles pour détecter la tromperie qui sont importants pour l'équité du procès et que la nécessité de voir le visage d'un témoin durant son témoignage « est trop enracinée dans notre système de justice pénale pour qu'on l'écarte en l'absence d'une preuve convaincante » (paragraphe 27). Dans cet article, les auteurs présentent des preuves empiriques convaincantes selon lesquelles (a) la grande majorité des indices de tromperie sont trop faibles pour détecter cette dernière de façon fiable, (b) la plupart des expressions faciales et autres indices non verbaux ne sont pas liés à la tromperie, (c) les particuliers, y compris les professionnels des disciplines juridiques (p. ex. les juges, les avocats) ne sont pas en mesure de détecter la tromperie au-delà du niveau du hasard, et (d) il n'est pas possible d'améliorer la détection de la tromperie avec des programmes de formation axés exclusivement sur les indices non verbaux. Le fait que les données empiriques existantes ne parviennent pas à s'harmoniser avec la croyance fort répandue entourant la détection de la tromperie soulève de sérieuses inquiétudes quant à la confiance ressenti à l'égard des croyances pseudoscientifiques dans la prise de décision de la CSC.*

---

[*] Dr. Brent Snook: Professor, Department of Psychology, Memorial University of Newfoundland. Meagan I. McCardle and Weyam Fahmy: MSc Candidates, Department of Psychology, Memorial University of Newfoundland. John C. House: MOM, Msc, MLitt, Department of Psychology, Memorial University of Newfoundland.

There is a venerated belief in the Canadian legal system that a witness' credibility can be estimated — by lawyers and triers of fact — through an assessment of his or her demeanour and facial expressions.[1] This belief was made conspicuous in 2012 when the Supreme Court of Canada (SCC) issued a decision about whether or not a sexual assault victim should be permitted to wear her niqab while testifying.[2] During a preliminary inquiry and subsequent appeals, it was ruled that the witness had to remove her niqab because it inhibits an assessment of truthfulness by lawyers and triers of fact, and hence, threatens trial fairness. At the SCC hearing, counsel for the victim and supporting interveners (e.g., Canadian Civil Liberties Association) disagreed; they argued that the importance of seeing N. S.'s face for trial fairness is unfounded. The SCC had to reconcile competing rights that are protected by the *Charter of Rights and Freedoms* (hereafter referred to as the Charter). On one hand, wearing a niqab while testifying is fundamental to the freedom of religion which is protected by s. 2 of the Charter.[3] On the other hand, the accused has a right to a fair trial which is protected by ss. 7 and 11(d) of the Charter. In a six to one decision, the SCC ruled in favour of the accused's right to a fair trial, and ordered N.S. to remove her niqab while testifying in court.

Writing on behalf of the majority, the Chief Justice concluded that the niqab should be removed when evidence is likely to be contested and cross-examination is a possibility. In other words, she reasoned that the niqab may be worn if the evidence is inconsequential or uncontested by the accused.[4] Otherwise, N.S. would have to remove the niqab so that her face is visible to prevent serious risk to trial fairness. This opinion was based on the assumption that there is a strong relationship between seeing a witness' face (i.e., having access to demeanour evidence) and having a fair trial. For instance, the Chief Justice wrote that "non-verbal communication can provide the cross-examiner with valuable insights that may uncover uncertainty or deception, and assist in getting at the truth",[5] and that "covering a witness's face may also impede credibility assessment by the trier of fact, be it judge or jury."[6] Furthermore, she stated that "on the record before

---

[1]  Amna M. Qureshi, "Relying on Demeanour Evidence to Assess Credibility During Trial: A Critical Examination" (2014) 61:2 Crim. L.Q. 235.

[2]  *R. v. S. (N.)*, 2012 SCC 72, 2012 CarswellOnt 15763, 2012 CarswellOnt 15764, [2012] 3 S.C.R. 726, 290 C.C.C. (3d) 404, 98 C.R. (6th) 1.

[3]  See works by Natasha Bakht for a comprehensive review of this issue, e.g., Natasha Bakht, "What's in a Face? Demeanour Evidence in the Sexual Assault Context" Elizabeth Sheehy ed., *Sexual Assault Law, Practice and Activism in a post-Jane Doe Era* (Ottawa: University of Ottawa Press, 2010) forthcoming.

[4]  This rationale, while seemingly accommodating, is impractical because a challenge to any core witness' testimony (e.g., victim) will likely result in a cross-examination.

[5]  *S. (N.)*, *supra* note 2, at para. 24.

[6]  *Ibid.*, at para. 25.

us, I conclude that there is a strong connection between the ability to see the face of a witness and a fair trial."[7]

The Chief Justice also indicated that altering the custom of interpreting facial cues — to obtain demeanour evidence — during testimony would be warranted upon hearing compelling evidence that having the face covered would not compromise trial fairness. Specifically, she wrote that the importance of seeing the face for assessing credibility ". . .is too deeply rooted in the criminal justice system to be set aside absent compelling evidence."[8] She also stated, "In the absence of evidence showing that these beliefs, backed by centuries of practice, are unsubstantiated "myths" that should be excised from the law, we should not take such a radical step [of allowing the face to be covered]".[9] In this case, the only evidence challenging the belief that credibility can be assessed accurately from facial cues appears to have been a four-page unpublished article.

Two concurring Justices agreed with the conclusion of the majority but took a non-conditional stance, reasoning that a witness must always remove the niqab while testifying. Writing on their behalf, LeBel J. argued that a clear rule is required whereby a niqab should be permitted "in all cases or not at all".[10] Referring to jurisprudence, he concluded that the right to cross-examine and the ability to observe the face of a witness during cross-examination are important parts of allowing the accused to defend himself against criminal charges. He also noted that the consequences of any restrictions to seeing a witness' face weigh more heavily on the accused, and that the process should be balanced to work in the accused's favour (i.e., wearing the niqab obstructs the ability to determine the guilt or innocence of the accused).

The remaining Justice, Abella J., representing a dissenting opinion, reasoned that mandating the removal of the niqab would be harmful to the administration of justice (e.g., crime reporting, attendance at trial) and to religious rights, and should take precedence over trial fairness. Even though she advocated for the wearing of the niqab, her underlying reasoning failed to challenge the belief that the face and other aspects of physical demeanour are useful for detecting deceit on the witness stand. Her thoughts pertaining to deception detection during trials are captured in the following quote:

> "There can be no doubt that the assessment of a witness' demeanour is easier if it is based on being able to scrutinize the whole demeanour package — face, body language, voice, etc. Nor is there any doubt that historically and ideally, we *expect* to see a witness' face when he or she is testifying. That, however, is different from concluding that unless the entire package is available for scrutiny, a witness' credibility cannot adequately be weighed."[11]

---

[7]   *Ibid.*, at para. 27.

[8]   *Ibid.*

[9]   *Ibid.*, at para. 49.

[10]   *Ibid.*, at para. 69.

In other words, she subscribed to the idea that non-verbal cues are useful for detecting deception and that more non-verbal cues are better than fewer cues. She further reasoned that accurate deception detection is still possible when some cues to deception are unavailable.

A number of overarching assumptions regarding deception detection were present throughout the entire ruling. In particular, the *implicit* assumptions were that cues to deception exist (i.e., features of a witness' demeanour are able to reveal someone's state of mind), deception can be detected in facial cues (and other non-verbal cues), legal professionals can detect deception better than laypersons, and people can be trained to use non-verbal cues to detect deception.

For a comprehensive review of the jurisprudence pertaining to the strongly held belief — at all levels of the Canadian criminal justice system — that credibility can be by assessed accurately from a witness' demeanour, we recommend an article by Qureshi.[12] Although Qureshi cited deception detection research to support her arguments that demeanour evidence is unreliable and threatens the truth-seeking function of a trial, it is our intention to provide a greater level of specificity regarding the findings from that and other literature, and interpret their practical significance for assessing credibility on the witness stand.

## 1. DECEPTION DETECTION RESEARCH

As mentioned, it is important to note that the SCC acknowledged that its decisions are not necessarily cast in stone and may be revisited with the introduction of evidence contradicting previous rulings. For instance, the SCC indicated that, "practices based on myth that once stood in the way of conviction have been set aside"[13], and it is not averse to discarding present practices that are demonstrably wrong. In this case, the SCC was not provided with expert evidence on the extent to which seeing a witness' face relates to effective cross-examination, or on the extent to which a witness' credibility can be assessed accurately from facial cues.

In light of the SCC's apparent openness to aligning decisions with scientific findings, we present the results of a several meta-analyses of the deception detection research literature. There is a wealth of research on deception detection, thus, we restrict our commentary primarily to the cues of interest in the aforementioned SCC ruling — namely, non-verbal cues related the face, and other relevant cues (e.g., head movement, physical demeanour). Specifically, we present compelling empirical evidence that (a) the vast majority of cues to deception are too faint for reliable deception detection, (b) most facial expressions and other non-verbal cues are unrelated to deception, (c) people, including professionals in legal disciplines (e.g., judges, lawyers), are unable to

---

[11]   *Ibid.*, at para. 91.

[12]   Qureshi, *supra* note 1.

[13]   *S. (N.)*, *supra* note 2, at para. 37.

detect deception barely beyond chance levels, and (d) it is not possible to improve deception detection with training programs that focus exclusively on non-verbal cues.

## 2. ARE THERE ANY RELIABLE CUES FOR DETECTING DECEPTION?

To determine if people behave differently when they tell the truth or lie, researchers produced a quantitative summary of cues to deception.[14] Specifically, for each cue to deception (e.g., gaze aversion), the authors measured the size of the difference (i.e., effect size) between those telling the truth and those telling lies. The effect size statistic computed was *d*. For ease of interpretation, Cohen classified effect sizes as small ($d = 0.20$, or low practical significance), medium ($d = 0.50$, moderate practical significance), and large ($d = 0.8$, high practical significance). Using the analogy of height of teenage girls, a small effect size would be the equivalent of the minimal height difference between 15 and 16 year-old girls. A medium effect size would be the equivalent of a perceptible height difference between 14 and 18 year-old girls. A large effect size is the more substantial height difference between 13 and 18 year-old girls.[15] Although the authors identified 158 cues to deception in the empirical literature, only effect sizes for cues that were examined in three or more studies were computed so that they could obtain reasonably reliable estimates. In the end, using that inclusion criteria, only 88 cues to deception were focused upon.

The results of this meta-analysis failed to pinpoint any behaviours of high practical significance for detecting deceit (i.e., $d > 0.8$). Two cues (2.27%) were moderate-to-high in practical significance, but neither of which were non-verbal in nature. One cue pertained to verbal and vocal immediacy (i.e., speakers respond in ways that seem direct, relevant, clear, and personal rather than indirect, distancing, evasive, irrelevant, unclear, or impersonal) and the other cue pertained to cooperation (i.e., speaker seems cooperative, helpful, positive, and secure).

The authors found that 21 (23.86%) cues were low-to-moderate in practical significance. Only four of those cues pertained to non-verbal cues (i.e., nervous and tense overall, pupil dilation, relaxed external associations, chin raise), and only two of those pertained to facial expressions. It is worth noting that the two facial expressions of low-to-medium practical significance would still be detectable from a witness wearing a niqab (i.e., looking at the eyes and looking for a raise of a chin). The remaining 65 (73.86%) cues examined were found to have low practical significance for detecting deceit. Importantly, the vast majority of cues that could potentially be garnered from the removal of the niqab had negligible impact on differentiating truth-tellers from liars (e.g., pressing lips, facial expressiveness, eye-contact, gaze aversion, eye-shifts, brow

---

[14]   Bella M. DePaulo et al., "Cues to Deception" (2003) 129:1 Psychological Bull 74.

[15]   Jacob Cohen, *Statistical power analysis for the behavioral sciences*, 2nd ed (New York: Academic Press, 1988).

lowering, sneering, smiling, lip corner pull, eye muscle movements, blinking, facial fidgeting).

A subsequent meta-analysis also provides insights into the existence of reliable cues to deception.[16] This quantitative summary of 11 widely-used non-verbal cues to deception (across 54 studies and 153 effect sizes) showed that all six cues from the head area (i.e., blinks, eye contact, gaze aversion, head movements, nodding, smiling; all $ds$ < 0.20) and four of the five cues from the body area (i.e., adapters, illustrators, foot/leg movements, postural shifts; all $ds$ < 0.20) were barely perceptible. Only hand movements ($d$ = 0.38) had low-to-moderate practical significance for detecting deception. More specifically, they found that the number of hand movements decreased when lying — a finding that runs counter to the beliefs held by law enforcement, students, and the general public.[17]

At issue in *R. v. S. (N.)* (2012) was the wearing of a niqab while testifying in court. We examined the results from two meta-analyses and found that, while there are some cues to deception, most are imperceptible, and those that are perceptible are related to verbal behaviours. In direct opposition to the reasoning provided by the majority in *R. v. S. (N.)*., the best available empirical research suggests that the removal of the niqab would not provide access to any reliable cues to deception.

## 3. CAN PEOPLE DETECT DECEPTION?

To determine if people, without any training or aids, are able to determine when someone is telling a lie, one meta-analysis examined the ability of people to detect deceit from demeanour.[18] The blunt, albeit somewhat simplistic, answer to their question is that people are poor deception detectors. From an analysis of 206 documents and 24,483 judgements, it is estimated that the ability to judge truth-telling and lying is around 54%; which is akin to guessing. More precisely, people can only correctly classify 47% of lies as deceptive and 61% of truths as non-deceptive.

A more nuanced consideration of this meta-analysis reveals that people are more accurate at judging deception when observing audio ($d$ = 0.37) or audiovisual ($d$ = 0.44) materials than when observing only visual materials. The results suggest that the contribution of having access to visual information (e.g., facial expressions) adds little to the ability to detect deception — a finding that is consistent with previous research findings that the strongest cues to deception are

---

[16] Siegfried L. Sporer.& Barbara Schwandt, "Moderators of Nonverbal Indicators of Deception: A Meta-Analytic Synthesis" (2007) 13:1 Psychol. Pub. Pol'y & L. 1.

[17] E.g., Lucy Akehurst et al., "Lay Persons' and Police Officers' Beliefs Regarding Deceptive Behaviour" (1996) 10:6 Applied Cognitive Psychology 461.

[18] Charles F Bond & Bella DePaulo, "Accuracy of Deception Judgments" (2006) 10:3 Personality & Social Psychology Rev 214 [Bond & DePaulo, "Accuracy of Deception Judgements"].

auditory in a nature.[19] Of particular relevance to the current commentary, Bond and DePaulo found that the ability to detect deception is negligible ($ds < 0.20$, or low practical significance) when people only see the face, only see the body, or see both the face and body. They also found that experts — such as law enforcement officials, judges, psychiatrists, job interviewers, and auditors — are no better than laypeople (e.g., students) at detecting deception. Although non-significant, the results suggest that experts may be less accurate than non-experts in detecting deceit; experts are also less inclined than laypersons to believe people are truthful.[20]

A second meta-analysis on deception detection accuracy provides converging evidence that people are not proficient in differentiating between liars and truth-tellers.[21] Specifically, an analysis of 108 studies and 16,537 participants revealed an average deception detection accuracy of 54.50%. The results showed that professionals (e.g., police officers, judges, psychologists) are not substantially more accurate than naïve students in detecting deception (55.51% vs. 54.22%; guessing would be 50%). The authors also found that confidence in one's ability to detect deception, age, experience, education, and personality type were all unrelated to deception detection accuracy.[22]

Relatedly, a meta-analysis — based on 142 studies including 19,801 judgments — of individual differences in deception detection accuracy (and bias) revealed that people are only, on average, 54.05% accurate when discriminating between lies and truth-telling.[23] Importantly, the authors conclude that "several converging lines of evidence indicate that virtually all *individuals* are barely able to detect lies, and that real differences in detection ability are miniscule."[24]

## 4. CAN TRAINING IMPROVE THE ABILITY TO DETECT DECEPTION?

The question of whether training improves the ability to detect deception was also addressed in a meta-analysis. In the most comprehensive quantitative

---

[19] DePaulo et al., *supra* note 14; for similar arguments about the need to focus on 'what is said' when trying to detect deceit see Aldert Vrij, "Nonverbal dominance versus verbal accuracy in lie detection: A plea to change police practice" (2008) 35: 10 Criminal Justice & Behavior 1323.

[20] Bond & DePaulo, "Accuracy of Deception Judgements", *supra* note 18.

[21] Michael G. Aamodt & Heater Custer, "Who Can Best Catch a Liar? A Meta-Analysis of Individual Differences in Detecting Deception" (2006) 15: Spring Forensic Examiner 6.

[22] For research showing that there is a negligible relationship between confidence and accuracy when detecting deception see Bella M. DePaulo, et al, "The Accuracy-Confidence Correlation in the Detection of Deception"(1997) 1:4 Personality & Social Psychology Rev 346.

[23] Charles F. Bond & Bella DePaulo, "Individual Differences in Judging Deception: Accuracy and Bias" (2008) 134:4 Psychological Bull. 477.

[24] *Ibid.*, at 487.

summary available, researchers explored the effect of any type of training on detection accuracy (the ability to differentiate between truth and lies), lie accuracy, and truth accuracy.[25] This meta-analysis revealed a small-to-medium training effect for detection accuracy and lie detection accuracy, but a negligible effect for truth detection accuracy. Their results suggest that it may be possible to improve the ability of people to differentiate between truth-tellers and liars, and to detect lies if the training program focused on detecting lies. Their results also suggest that it is not possible to teach people to detect the truth, unless the training program focused on detecting the truth with verbal cues. Of particular importance to the current commentary was the finding that training programs that included nonverbal and/or para-verbal cues (e.g., vocal expressions such as pitch) tended to produce small or insignificant effects on the ability to detect deception.

## 5. THE DECLINE EFFECT

Notwithstanding the data presented above, those who believe in the human ability to detect deception and the validity of deception detecting training have a strong presence in the literature, and thus their criticisms cannot be ignored. The main criticism that "believers" put forth is that research in deception detection often relies on laboratory studies and "little white lies", which do not reflect the high-stakes reality of lies in real-world investigative contexts.[26] However, research studies that utilize "higher stakes" lies are a relatively new concept, which, due to the decline effect, could be a problem in itself. If the cues associated with higher stakes lies are fundamentally different than those from laboratory studies as some claim, those cues to deception are subject to what is called the *decline effect*. As Granhag and colleagues eloquently explain:

> Deception cues become weaker the longer they are studied. . . . The so-called *decline effect* is a tendency for scientific findings to decrease in strength over time. Such effects have been observed in medicine and biology (McMahon, Holly, Harrington, Roberts, & Green, 2008; Ozonoff, 2011), and here we seem to be observing a decline effect in non-verbal behaviour cues. There is currently no generally agreed-upon explanation for the decline effect. Schooler (2011) notes some prosaic explanations for the effect, such as regression to the mean (i.e. self-correction from extreme values over time) or publication bias, meaning that the publication of novel findings may be more likely if these effects are large, whereas follow-up studies on an established phenomenon may pass the bar of peer review even if the effects are smaller. . . In sum, research on non-verbal cues that distinguish truths from lies finds that

---

[25] Valerie Hauch et al, "Does Training Improve the Detection of Deception? A Meta-Analysis" (2014) 43:3 Communication Research 283.

[26] Clea Wright Whelan, Graham Wagstaff & Jacqueline M. Wheatcroft, "High Stakes Lies: Police and Non-Police Accuracy in Detecting Deception" (2015) 21:2 Psychology, Crime & L. 127.

> such cues tend to be weak and inconsistent and, to the extent there is empirical support for such cues, that support systematically diminishes as research accumulates.[27]

Altogether, their commentary explains that new findings in the field of deception detection based on single studies must be interpreted with caution. Specifically, the use of a single primary study (or a couple of primary studies) to refute the meta-analytic findings ought to be viewed with skepticism until a robust body of research has demonstrated that that deception detection is possible.

## 6. RESEARCH WITH VEILED WITNESSES

Granted concerns about the decline effect, recent research that directly tackles the issue of deception detection in veiled witnesses must be acknowledged.[28] Across two experiments, all participants were asked to watch 20 videos of different women providing eyewitness testimony regarding the theft of items from a purse (10 were truthful statements and 10 were false). Participants were also assigned randomly to one of three veiling conditions where all 20 eyewitnesses they watched were either wearing a niqab, a hijab, or were non-veiled. They were then asked to decide which of the eyewitnesses were telling the truth and which were lying. The results showed, in both experiments, that deception detection accuracy was slightly above chance levels. Of direct importance, however, was the finding that the deception detection accuracy was significantly higher (but small effect size) when the face was fully veiled (i.e., niqab condition) than when the face was fully exposed (i.e., no veil condition). In addition, there was also no significant difference in deception detection accuracy between the niqab condition and the hijab condition (i.e., only the hair was veiled). In sum, the only study that has explored the effect of veiling on deception detection accuracy found that covering the face while testifying may actually improve, albeit slightly, the ability to detect deception. Such findings converge with the conclusions from the meta-analytic studies that seeing the face does little to assist with deception detection accuracy.

## 7. CONCLUDING COMMENTS

The Justices alluded to the idea that the presentation of scientific evidence could have impacted their conclusion. Such a sentiment is perplexing because a vast amount of research on deception detection was available during the SCC's decision-making process. As reviewed above, several meta-analyses of the detection deception research showed that most cues to deception are too faint for

---

[27] Pär Anders Granhag, Aldert Vrij & Bruno Verschuere, *Detecting Deception: Current Challenges and Cognitive Approaches* (Chichester: John Wiley & Sons, 2015) at 44-45.

[28] Amy-May Leach et al., ''Less is More? Detecting Lies in Veiled Witnesses'' (2016) 40:4 L.. & Human Behavior 401.

reliable detection, most facial expressions and other non-verbal cues are unrelated to deception, legal professionals (and others) are unable to accurately detect deception beyond chance levels, and that training people to use non-verbal cues to improve their deception detection is unviable. With this evidence considered, the force of the conclusions of the SCC's majority ruling is diminished significantly.

We echo Qureshi's argument that these data suggest that nobody's credibility (or liberty) should rest solely, or even partially, on the assessment of non-verbal cues.[29] Given that it is Canada's final court of appeal, serious consideration needs to be given to ensuring that robust scientific findings are incorporated into fully informed decision-making by the Justices of the SCC, and equally important, that pseudoscientific beliefs are eradicated from SCC decision-making. In the case of *R. v. S. (N.)* (2012), without the scientific findings to consider, the SCC based their decision on what they had available to them — layperson beliefs and customs regarding demeanour "evidence". However, by doing so, layperson beliefs were used to override the complainant's constitutional right to freedom of religion; a right that may have otherwise been protected had the relevant scientific evidence been considered.

---

[29] Qureshi, *supra* note 1.