

To Veil or Not to Veil, Detecting Lies in The Courtroom: A Comment on Leach et al. (2016)

Vincent Denault ^{1,2}, Louise Marie Jupe ³, Olivier Dodier ^{4,5}, & Nicolas Rochat ⁶

¹ Université de Montréal, Montreal, Canada

² Centre d'études en sciences de la communication non verbale, Montreal, Canada

³ University of Portsmouth, Portsmouth, United Kingdom

⁴ Université Blaise Pascal – Clermont Université, Clermont-Ferrand, France

⁵ Université Toulouse – Jean Jaurès, Toulouse, France

⁶ Université de Paris VIII, Saint-Denis, France

Acknowledgments:

The authors would like to thank Hugues Delmas, Ian Fraser, Barry Morrison, Jim Stiff and Éric Raymond for their constructive comments on an earlier version of this manuscript.

Manuscript accepted for publication in

Psychiatry, Psychology and Law

The final publication is available at Taylor and Francis via

<http://dx.doi.org/10.1080/13218719.2017.1260619>

Corresponding author:

Vincent Denault
Département de communication
Université de Montréal
Pavillon Marie-Victorin
PO Box 6128 Centre-ville STN
Montréal (Québec)
H3C 3J7
Canada

E-mail: vincent.denault@umontreal.ca

Abstract

For the past 40 years, lie detection has predominantly been studied in the context of police-suspect and investigative interviews. In their paper, Leach et al. (2016) examined whether niqabs or hijabs interfere with the trial judges' ability to detect deception and concluded that veiling enhanced trial judges' ability to make accurate veracity judgments. In this comment, we argue that the conclusions made by Leach et al. are based upon an inaccurate experimental court paradigm and suffer from methodological and analytical issues. It is our opinion that the applicability of their research findings to real-life court proceedings alongside potential changes to court practices and policies based on Leach et al. should be regarded as naïve and misinformed.

Keywords: courtroom; credibility assessment; cross-examination; deception detection; lie detection; niqab; trial.

To Veil or Not to Veil, Detecting Lies in The Courtroom: A Comment on Leach et al. (2016)

During trials, despite the oath to tell the truth, the whole truth and nothing but the truth, witnesses can provide false testimony. While some sincerely believe that their memory of an event is accurate, although it is factually erroneous (Laney & Loftus, 2013; Schacter & Loftus, 2013), other witnesses can deliberately attempt to mislead trial judges. However, unlike mistaken eyewitness identification, very few studies have been made about deception in courtrooms (Fawcett, 2014). Over the past four decades, lie detection has predominantly been studied in the context of police-suspect and investigative interviews (Granhag & Strömwall, 2004; Vrij, 2008).

In their paper, Leach et al. (2016) aimed to study the largely uncharted territory of deception in courtrooms, specifically where a witness wears a niqab or hijab. According to Leach et al., different jurisdictions assume that wearing a veil that covers the wearer's face, apart from the eyes, hampers the trial judges' ability to detect deception:

Judges in the United States, the United Kingdom, and Canada have ruled that witnesses may not wear the niqab – a type of face veil – when testifying, in part because they believed that it was necessary to see a person's face to detect deception. (Leach et al., 2016, p. 401)

Leach et al. (2016) conducted two empirical studies to examine whether the niqab or hijab interferes with trial judges' ability to detect deception. Their conclusions are clear and unequivocal:

Contrary to the assumptions underlying the court decisions cited earlier, lie detection was not hampered by veiling across two studies. In fact, observers were more accurate at detecting deception in witnesses who wore niqabs or hijabs than those who did not veil. (Leach et al., 2016, p. 407)

In this comment, we argue that the conclusions by Leach et al. (2016) are based upon an inaccurate experimental court paradigm and suffer from methodological and analytical issues. In addition, their conclusion that ‘the data consistently suggested that minimizing visual information actually improved participants’ lie detection performance’ (p. 407) appears unwarranted and has the potential to influence the outcome of trials adversely. Based on the foregoing, we discuss other potential implications of the research findings by Leach et al. in real-life court proceedings.

The Inaccurate Experimental Court Paradigm

In order to examine whether niqabs or hijabs interfere with trial judges’ ability to detect deception, Leach et al. (2016) stressed the importance of recreating an experimental court paradigm consistent with real-life court proceedings. A resounding feature of their experimental court paradigm is initiated by the witnesses’ preparation:

Witnesses were given 2 minutes to prepare their testimony and, as *in real trials* [emphasis added], they were provided with the questions that would be asked by the defense lawyer. Once they were prepared, witnesses were randomly assigned to don a black niqab, a black hijab, or remain unveiled. In addition, they were asked to wear an opaque black shawl to conceal and control for clothing. Veils and shawls were placed on the witnesses by a trained research assistant. (Leach et al., 2016, p. 403)

The study then continues with the hearing, namely the direct examination (examination-in-chief) and cross-examination:

Witnesses were interviewed by two female experimenters. To *simulate courtroom procedures* [emphasis added], one experimenter played the role of the sympathetic defense lawyer and asked 16 information gathering questions (e.g., ‘Please describe everything that you saw the woman do’). The other experimenter conducted a challenging cross-examination as the prosecutor and asked seven unanticipated questions (e.g., ‘The

police found the man's laptop. The defendant's fingerprints were on it. How do you explain that?"). (Leach et al., 2016, p. 403)

However, although Leach et al. (2016) set up a scenario to replicate a witness's preparation as 'in real trials' (Leach et al., 2016, p. 403) and to 'simulate courtroom procedures' (Leach et al., 2016, p. 403) for a hearing, their experimental court paradigm bears no relation to real-life court proceedings.

The Witness's Preparation

After being randomly assigned to watch a video of a woman observing a stranger's bag or stealing items from the bag, 'witnesses were informed that the woman had been accused of theft and they were being called to testify on her behalf (i.e., they were to state that they did not see her steal anything)' (Leach et al., 2016, p. 403). Then, witnesses received the defense lawyers' list of questions and had two minutes to prepare their testimony. However, *in real trials*, such a modus operandi is completely unrealistic. In fact, several months or years can pass between an event and the moment a witness testifies about it. Such delays can weaken emotional reactions and displays (Wilson, Gilbert, & Centerbar, 2003) and thereby affect the witness's credibility (Heath, 2009; Heath & Grannemann, 2015). Furthermore, having a two-minute witness preparation without the assistance of a defense lawyer lacks consistency with real-life court proceedings in the United States and Canada (i.e., extensive witness preparation) as well as in the United Kingdom (i.e., forbidden witness preparation).

In contrast to the United Kingdom where barristers 'must not rehearse, practise with or coach a witness in respect of their evidence' (Bar Standard Board, 2015, p. 27), lawyers in the United States and Canada can extensively prepare their witnesses. Thus, lawyers not only review the evidence with their witnesses, but also prepare and rehearse the direct examination and the cross-examination. The process may take several hours, or even days. Lawyers go over the

questions and practice the direct examination with their witnesses. They also review the topics on which the trial judge or the opposing counsel might ask questions and practice the cross-examination with their witnesses (Mauet, 2013; Posey & Wrightsman, 2005). However, preparations and rehearsals can reduce the cognitive cost of lying (Van Bockstaele, Verschuere, Moens, Suchotzki, Debey, & Spruyt, 2012) and thereby hamper the trial judge's ability to detect deception (Vrij, Granhag, Mann, & Leal, 2011; Vrij, Granhag, & Porter, 2010). In other words, the experimental witness's preparation is inconsistent with a real-life witness's preparation and thus questions the ecological validity of the conclusions made by Leach et al. (2016), notably because variations in emotion and cognitive load associated with real-life court proceedings will affect the trial judges' decision as to whether a witness is lying or telling the truth: a significant limitation not addressed by Leach et al.

Before the hearing, witnesses were also randomly assigned to wear a black niqab or a black hijab, or remain unveiled. However, not only did Leach et al. (2016) neglect to mention whether the witnesses were accustomed to the wearing of a niqab on a day-to-day basis, but one can assume that witnesses randomly assigned to wear a niqab did not wear it because of a sincere religious belief. Such an experimental court paradigm creates a totally fictitious situation, especially so with regards to Canada. In *R. v. N.S.* (2012), the Supreme Court of Canada called into question the Ontario Court of Appeal decision *R. v. N.S.* (2010) and developed a framework on the issue of when, if ever, a witness who wears a niqab can be required to remove it:

... where a niqab is worn because of a sincerely held religious belief, a judge should order it removed if the witness wearing the niqab poses a serious risk to trial fairness, there is no way to accommodate both rights, and the salutary effects of requiring the witness to remove the niqab outweigh the deleterious effects of doing so. (*R. v. N.S.*, 2012, p. 751).

Thus, in accordance to *R. v. N.S.* (2012), if the wish to wear a niqab is not based on a sincere religious belief, the witness will be required to remove the niqab while testifying, and not because it is necessary to see a person's face to detect deception as implied by Leach et al. (2016). In other words, *in real trials*, the issue will only arise with witnesses who believe that their religion requires them to wear a niqab while testifying in court. However, in the experimental witness's preparation, the novelty of the situation will likely affect the witnesses' emotional reactions and displays (Lewis, Haviland-Jones & Barrett, 2008) and impression management processes (Leary, 1996). This will in turn affect the witnesses' non-verbal behaviour, an overlooked limitation considering that Leach et al. examine the assumption that wearing a veil that covers the wearer's face, apart from the eyes, hampers the trial judges' ability to detect deception. This characteristic of the experimental witnesses' preparation further calls into question the ecological validity of the conclusions made by Leach et al.

The Hearing

After the witness's preparation, one experimenter asked the witness 16 information-gathering questions (e.g., 'Please describe everything that you saw the woman do.')

and another experimenter asked the witness seven unanticipated questions (e.g., 'The police found the man's laptop. The defendant's fingerprints were on it. How do you explain that?'). However, Leach et al. (2016) failed to *simulate courtroom procedures* for a hearing, especially so with regards to the cross-examination. In real-life court proceedings, direct examination is not always followed by cross-examination. However, when a cross-examination is conducted, the goal is generally to tell the lawyers' side of the story:

To begin with, why are you cross-examining the witness in the first place? *The point of cross is not to get information from the witness* [emphasis added]. The point of cross is not to get the witness to change his story. And most of the time, the point of cross is not to

destroy the witness with contradictions and clever impeachment. The point of cross is to let you – the lawyer – tell your side of the case so the witness has to agree that what you say is true. (McElhaney, 1997, p. 82)

To conduct successful cross-examinations – while their approach may vary depending on their goal, their style, and the communication skills of the witnesses – lawyers will generally use short leading questions that call for a yes or no answer:

Not only are leading questions perfectly permissible on cross-examination, but on cross-examination we must ask *nothing but* [emphasis added] leading questions. Our whole aim on cross-examination is to grab the witness by the collar and take him where we want him to go. (Younger, 1976, p. 294)

Thus, to *simulate courtroom procedures*, the experimenter should have asked short leading questions that call for a yes or no answer and not open-ended questions to get information from the witness (e.g., ‘The police found the man’s laptop. The defendant’s fingerprints were on it. How do you explain that?’). This is an error within any courtroom scenario where there is a cross-examination. Even if generally accepted principles promote the use of open-ended questions over leading questions in investigative interview settings (Lamb, Hershkowitz, Orbach, & Esplin, 2008; Milne & Bull, 2003; Read, Powell, Kebbell, & Milne, 2009), this is not the kind of question generally asked in cross-examination. Furthermore, questions asked in direct examination have many goals (e.g., to introduce witnesses and ask background questions, set the scene and recreate the action, and address the weaknesses of a witness before the opponent is able to), but to gather information is generally not one of them (Baldwin, 2013). Thus, the aim by Leach et al. (2016) to study deception in courtrooms within an experimental court paradigm bears more resemblance to investigative interview settings. This is a significant limitation not

addressed by Leach et al., the more so considering that open-ended questions improve deception detection accuracy in investigative interview settings (Vrij & Granhag, 2012).

Considering that the unanticipated nature of open-ended questions also improves deception detection accuracy in investigative interviews settings (Vrij, Fisher & Blank, 2015), the overall accuracy rates reported by Leach et al. (2016) are highly unlikely to be representative of what happens in real-life court proceedings. In fact, in real-life court proceedings, if cross-examination is prepared and rehearsed, as it often is in the United States and Canada, defendants will know the topics on which the judge or the opposing counsel might ask questions. Unanticipated open-ended questions are even more improbable because of the disclosure of evidence, a basic principle in criminal proceedings that helps to prepare and rehearse cross-examinations of defense lawyers' witnesses.

Furthermore, in Leach et al. (2016), participants asked to decide whether the witnesses in the video footages were lying or telling the truth could not ask questions directly to them. As such, this experimental court paradigm creates a fictitious situation, especially so with regards to Canada where administrative, civil, and criminal trials are mostly adjudicated by trial judges without juries and where, according to the Supreme Court of Canada, trial judges should be actively involved in trials:

... it is clear that judges are no longer required to be as passive as they once were; to be what I call sphinx judges. We now not only accept that a judge may intervene in the adversarial debate, but also believe that it is sometimes essential for him to do so for justice in fact to be done. Thus a judge may and sometimes must ask witnesses questions, interrupt them in their testimony and if necessary call them to order. (Brouillard Also Known As Chatel v. The Queen, 1985, p. 44)

Considering that a significant increase in overall accuracy rates is observed when questioning is used strategically (Hartwig, Granhag, Strömwall, & Vrij, 2005; Hartwig, Granhag, Strömwall, & Kronkvist, 2006; Levine, Shaw, & Shulman, 2010), the overall accuracy rates reported by Leach et al. (2016) are unlikely to be an actual representation of what happens in actual court proceedings, the more so considering the base rate of liars versus truth tellers in Leach et al., which is also likely to be largely unrepresentative of the base rate of liars versus truth tellers in actual court proceedings.

In the experimental court paradigm by Leach et al. (2016), witnesses were not plaintiffs in a civil case as in *Muhammad v. Enterprise Rent-A-Car* (2006), complainants in a criminal case as in *R. v. N.S.* (2010), or defendants in a criminal case as in *The Queen v. D.(R.)* (2013); three judicial decisions presented as pivotal in the assumption that Leach et al. applied to the experimental court paradigm (i.e., that niqabs interfere with the trial judges' ability to detect deception). They were, in fact, selfless witnesses devoid of any relationship with the complainants or the defendants. In other words, in real-life court proceedings, selfless witnesses generally have nothing to win and nothing to lose if they tell the truth. On the contrary, if they lie, they may face a criminal charge of perjury associated with a fine, a prison sentence, or both (Farmer & Hancock, 2014). Therefore, it is safe to hypothesize that selfless witnesses will be more truthful, and trial judges will show a stronger truth bias against that category of witnesses than against defendants facing prison in criminal cases or plaintiffs facing bankruptcy in civil cases, thus affecting the base rate of liars versus truth tellers and the overall accuracy rates (Burgoon & Levine, 2010; Levine, Clare, Green, Serota & Park, 2014), which can also be affected by the oath that witnesses take before testifying (Jacquemet, Luchini, Rosaz & Shogren, 2015; Talwar & Crossman, 2012), another key component of real-life court proceedings missing in the experimental court paradigm by Leach et al.

To conclude, video footage generated and used in Study 1 and Study 2 have no relationship to real-life court proceedings. Leach et al. (2016) did not set up a scenario that replicated a witness's preparation as 'in real trials' (Leach et al., 2016, p. 403) nor did they 'simulate courtroom procedures' (Leach et al., 2016, p. 403) for a hearing. This is a significant limitation considering that Leach et al. assert that the 'two studies reported here provide unique tests of the behavioral assumptions underlying important courts decisions in the United States, United Kingdom, and Canada' (Leach et al., 2016, p. 408). Therefore, the authors' research findings are based upon an inaccurate experimental court paradigm that undermines their claim that 'In addition to the potential policy implications concerning the wearing of a niqab or hijab on the stand, the studies reinforce the value that behavioral science data have for informing judiciaries' (Leach et al., 2016, p. 408).

The Methodological Issues

Not only is the experimental court paradigm inconsistent with real-life court proceedings, the conclusions drawn from the results by Leach et al. (2016) suffer from methodological and analytical issues, which hinder their claim that 'niqabs do not interfere with – and may, in fact, improve – the ability to detect deception' (Leach et al., 2016, p. 401) of trial judges – namely, issues regarding the use of 'forced choice' of cues to deception and their use of signal detection theory.

'Forced choice' of cues to deception.

In Study 1, the participants were asked to decide whether the witnesses in the video footages were lying or telling the truth, to indicate the confidence of each of their decisions with a scale from 0% (not at all confident) to 100% (extremely confident) and to choose the cues they used to make their decision from a list of diagnostic and non-diagnostic verbal and non-verbal cues to deception. The 'forced choice' of cues to deception may have hindered the reporting of

the actual cues used by the participants to detect deception in courtrooms. That is, participants may have been using cues not provided within the list determined by the experimenters. In addition, such an experimental court paradigm does not reflect the way in which trial judges (at least) make veracity judgments – that is, without a predetermined list of diagnostic and non-diagnostic verbal and non-verbal cues to deception. However, by allowing open-ended responses, the participants would have been able to provide cues that are not only specific to that found within previous literature (Bogaard, Meijer, Vrij, & Merckelbach, 2016). This is an overlooked methodological issue considering that the study of deception in courtrooms is largely uncharted territory, even more in the specific context where a witness wears a niqab or hijab. In other words, before attempts to test an assumption with inferential statistics, researchers should have a thorough understanding of the phenomenon under study (Park, Levine, McCornack, Morrison & Ferrara, 2002; Rozin, 2001).

Moreover, Leach et al. (2016) neglect to mention whether the experimenters who determined the list of diagnostic and non-diagnostic verbal and non-verbal cues to deception were blinded to the hypothesis of their study. In addition, it is unclear whether the experimenters who singled out the 60 video clips evenly distributed into the six conditions (liars vs. truth tellers, and niqab vs. hijab vs. no veil) were blinded to the liars versus truth tellers condition. If not blinded, a bias might have affected the confection of the list and the selection of the 60 video clips (e.g., selecting unconvincing liars in the niqab and hijab conditions and convincing liars in the no veil condition). Given the above, their conclusions that ‘witnesses in niqabs revealed significantly more verbal than nonverbal cues’ (Leach et al., 2016, p. 408) and that ‘participants were more likely to base their decisions on verbal cues than nonverbal cues when viewing witnesses from this group’ (Leach et al., 2016, p. 408) are unjustified, not only because of the use of ‘forced choice’ cues to deception, but perhaps because of the biased list of diagnostic and non-diagnostic

verbal and non-verbal cues to deception and 60 video clips. In other words, the actual cues displayed by the witnesses and used by the participants to detect deception in courtrooms are unknown.

Finally, it is worth noting that, in Study 1, only the diagnostic verbal and non-verbal cues to deception were analysed. The authors disregarded the non-diagnostic cues to deception and justified the decision on the rationale that ‘including known non-diagnostic cues in the analysis would have unnecessarily impeded the likelihood of uncovering significant effects’ (Leach et al., 2016, p. 405). However, recent research has demonstrated that some individuals show higher accuracy in veracity judgments without citing diagnostic cues to deception, while actually citing more non-diagnostic or assumed ‘invalid’ cues to deception (Jupe, Akehurst, Vernham, & Allen, 2016). Therefore, the exclusion of non-diagnostic verbal and non-verbal cues to deception appears to be an arbitrary decision in relation to the potential research findings. The fact that the actual cues displayed by the witnesses and used by the participants to detect deception in courtrooms are unknown is a limitation downplayed by Leach et al. (2016).

It is also worth noting that, in Study 2, participants from the United Kingdom and the Netherlands were not asked to provide the verbal and non-verbal cues to deception that they used to decide whether witnesses in the video footage were lying or telling the truth. However, if Study 2 was to serve as a direct replication of Study 1, as Leach et al. (2016) presented it, the design of Study 1 should have also asked participants from the United Kingdom and the Netherlands for the verbal and non-verbal cues to deception that they used when making their veracity judgements. Therefore, Study 2 is not a direct replication of Study 1, and its design downplays the already unwarranted conclusions of Study 1 related to cues displayed by the witnesses and used by the participants to detect deception.

Signal Detection Theory

In Study 1 and Study 2, the results are furthered with analysis drawn from signal detection theory (SDT; MacMillan & Creelman, 1991; Meissner & Kassin, 2002). An SDT analysis calculates the response bias (β) and discrimination accuracy (d'), which is a measure of pure sensitivity achieved by measuring the signal and the noise means in standard deviation units. Such calculations correct for response bias and guessing. Positive values suggest an ability to discriminate liars from truth tellers, and negative values suggest response confusion.

Discrimination accuracy analysis includes the calculation of hits, misses, false alarms, and correct rejections. A hit is a lie correctly identified as a lie, a miss is a lie incorrectly identified as a truth, a false alarm is a truth incorrectly identified as a lie, and a correct rejection is a truth correctly identified as a truth. Accurate decisions are composed of hits and correct rejections, and inaccurate decisions are composed of misses and false alarms. However, Leach et al. (2016) state that only hits and false alarms were calculated, which only draw partial conclusions based on an analysis from decisions made where they correctly identified a lie as a lie and incorrectly identified a truth as a lie. Given the above, the authors' reported discrimination accuracy (d') analysis appears only to be partially completed in relation to the potential research findings.

In addition, irrespective of the veiling condition, the research findings by Leach et al. (2016) highlighted truth biases with participants demonstrating an accuracy rate of .72 for truths and .38 for lies ($d = 1.66$) in Study 1 and .71 for truths and .39 for lies ($d = 1.52$) in Study 2. The truth biases are further highlighted by the reported response bias (β). The participants in Study 1 exhibited a truth bias in the hijab condition but not in the niqab or no veil condition. In Study 2, the participants exhibited a truth bias in the hijab and niqab condition but not in the no veil condition. Considering that the truth bias is related to the veracity effect, which can impede overall accuracy rates (Levine, Park & McCornack, 1999), it is possible that part of the difference

in Study 1 and Study 2 between overall accuracy rates in the veil conditions and in the no veil condition may in fact be down to a truth bias, a limitation downplayed by Leach et al.

To conclude, the research findings and inferences drawn by Leach et al. (2016) suffer from methodological and analytical issues that further undermine their claim that ‘in addition to the potential policy implications of their research concerning the wearing of a niqab or hijab on the stand, the studies reinforce the value that behavioural science data have for informing judiciaries’ (Leach et al., 2016, p. 408). Since their conclusions are also based on an inaccurate experimental court paradigm, it is our opinion that the applicability of their research findings to real-life court proceedings alongside potential changes to court practices and policies based on Leach et al. should be regarded as naïve and misinformed.

Discussion

In their paper, Leach et al. (2016) examined whether the niqab or hijab interfere with the trial judges’ ability to detect deception. However, while we applaud the authors for opening a discussion about deceptive courtroom interactions, we argue that the conclusions by Leach et al. are based upon an inaccurate experimental court paradigm and suffer from methodological and analytical issues. Such weaknesses raise serious concerns with regards to the clear and unequivocal assertion that ‘the two studies reported here provide unique tests of the behavioral assumptions underlying important courts decisions in the United States, United Kingdom, and Canada’ (Leach et al., 2016, p. 408). Furthermore, their conclusion that ‘the data consistently suggested that minimizing visual information actually improved participants’ lie detection performance’ (Leach et al., 2016, p. 407) appears unwarranted and has the potential to influence the outcome of trials adversely.

Psychology and law scholars play a crucial role in the development of evidence-based practices and policies. Racial biases (Meissner & Brigham, 2001; Michel, Caldara, & Roission,

2006), witness identification (Erickson, Lampinen, & Moore, 2016; Wixted, Read, & Lindsay, 2016), eyewitness memory (Brainerd, 2013; Loftus, 2005) and false confession (Leo, 2009; Shaw & Porter, 2015) are a few examples of where psychological science has triggered changes in the judicial system. However, while some psychology and law research domains sometimes take years to have an impact upon institutionalized practices and policies – sometimes even with only limited success – a single peer-reviewed paper on credibility assessment or deception detection, such as Leach et al. (2016), has a potential to influence the outcome of trials adversely that should not be overlooked. In Canada, for example, the way trial judges assess witnesses' credibility provides considerable insights into such potential impact.

According to the Supreme Court of Canada, the highest court in Canada, credibility assessment is 'a matter that must be left to the common sense of the trier of fact' (R. v. François, 1994, p. 839). Credibility assessment 'must always be the product of the judge or jury's view of the diverse ingredients it has perceived at trial, combined with experience, logic and an intuitive sense of the matter' (R. v. Marquard, 1993, p. 248). Demeanour is one of the diverse elements that trial judges can take into account:

The trier of fact will observe the witness's demeanour and the way she answers the questions. The result may be that the trier of fact does not accept the witness's evidence, accepts only part of her evidence, or reduces the weight accorded to her evidence. (R. v. D.A.I., 2012, para 72)

The way trial judges assess credibility has been met with severe criticism. For example, Porter and Ten Brinke (2009) described how credibility assessments can be biased by non-diagnostic or assumed 'invalid' cues to deception (e.g., babyfacedness and gaze aversion). In other words, trials' outcome can be biased by common sense and stereotypes, a phenomenon well documented since the 1960's (Garfinkel, 1967) and still highly relevant in psychology and law

research (Baker, Porter, Ten Brinke & Mundy, 2016; Beety, 2013; Dumas & Testé, 2006; Eberhardt, Davies, Purdie-Vaughns, & Johnson, 2006; Porter & Ten Brinke, 2009; Seelau & Seelau, 2005; Zebrowitz & McDonald, 1991). More recently, Denault (2015) highlighted real case law where non-diagnostic or assumed ‘invalid’ cues to deception (e.g., gaze, shrug, and flush) affected the assessment of witnesses’ credibility. Considering that ‘credibility is an issue that pervades most trials, and at its broadest may amount to a decision on guilt or innocence’ (R. v. Handy, 2002, para 115), Denault (2015) advocated that particular attention should be paid to notions about credibility assessment and deception detection that trial judges come across during training and are subjected to in the media. It is especially relevant as those notions can affect – consciously or unconsciously – the way they assess credibility (Porter et Ten Brinke, 2009) and may be prevalent throughout their careers, without knowing they are non-diagnostic or assumed ‘invalid’ cues to deception.

According to the Supreme Court of Canada, credibility assessment ‘does not always lend itself to precise and complete verbalization’ (R. v. R.E.M., 2008, para 49). In other words, trial judges do not have the obligation to explicitly mention in their judgments all the factors they take into consideration when assessing credibility. Furthermore, even if they cite non-diagnostic or assumed ‘invalid’ cues to deception in their judgments, chances are their decisions will be upheld ‘in the absence of a palpable and overriding error’ (R. v. Gagnon, 2006, para 20) because, according to the Supreme Court of Canada, ‘the trial judge has the advantage, denied to the appellate court, of seeing and hearing the evidence of witnesses’ (R. v. W. (R.), 1992, p. 131). In other words, trial judges receive very little, if any, feedback on their veracity judgments. What is more detrimental is when a trial judge makes a personal decision with regards to an individual’s truthfulness (in particular when the trial judge believes the individual to be a liar), and the individual is subsequently convicted of a crime, the cues used by the trial judge are likely to be

reinforced. This reinforcement of non-diagnostic or assumed ‘invalid’ cues to deception may span decades of service.

Furthermore, since trial judges are limited in their understanding of lie detection research (Strömwall & Granhag, 2003) and science in general (Kozinski, 2015; Chin & Dallen, 2016), and some even attended pseudoscientific training (Denault, 2015; Denault, Larivée, Plouffe, & Plusquellec, 2015), we see no reason why trial judges would not, in good faith, accept the conclusion drawn by Leach et al. (2016) that ‘the data consistently suggested that minimizing visual information actually improved participants’ lie detection performance’ (Leach et al., 2016, p. 407) and start paying less attention to witnesses’ and lawyers’ facial expressions. However, accepting the conclusion without an understanding of the limitations addressed by Leach et al., the other shortcomings we highlighted in this comment as well as pivotal concepts about lie detection and social interactions would be detrimental. The benefits of paying less attention to witnesses’ and lawyers’ facial expressions are neither theoretical nor empirically grounded arguments.

The function of witnesses’ and lawyers’ facial expressions goes well beyond the issue of lie detection. Being the major vector of communication, the face plays a central role in social interactions (Madal & Awasthi, 2015; Russell & Fernandez-Dols, 1997) and offers meaningful information about others such as intentions (Morris & Keltner, 2000; Pietroni, Van Kleef, De Dreu, & Pagliaro, 2008) and emotions (Du, Tao, & Martinez, 2014; Smith, Cottrell, Gosselin, & Schyns, 2005). The face, therefore, plays a major role in shaping trial interactions (i.e., direct examinations, cross-examinations, or arguments) just as it does with any other social interaction (Frith, 2009; Schmidt & Cohn, 2001). In other words, since lie detection is far from systematically being at the heart of every administrative, civil, and criminal trial, paying less

attention to witnesses' and lawyers' facial expressions could result in consequences that go far beyond the initial problem it is intended to solve.

Furthermore, while trial judges can use non-diagnostic or assumed 'invalid' cues to deception throughout their careers, the explicit mention of those cues in their judgments can create legal precedents for other trial judges to follow, even if those cues are contrary to scientific consensus. For example, Denault, Delmas, and Rochat (2016) highlighted how a set of dubious credibility assessment criteria put forward in a 1986 arbitration ruling have been used up until the current day by other arbitrators, including one very unusual criterion initially presented by the Supreme Court of Canada in 1897:

Witness who testifies to an affirmative is to be credited in preference to one who testifies to a negative, *magis creditur duobus testibus affirmantibus quam mille negantibus*, because he who testifies to a negative may have forgotten a thing that did happen, but it is not possible to remember a thing that never existed. (Lefeunteum v Beaudoin, 1897, p. 93)

Thus, while some psychology and law research subjects might take years to have an impact upon institutionalized practices and policies in court, it is important not to underestimate how easy and quickly notions about credibility assessment and deception detection that trial judges come across during training and are subjected to in the media, or in a single peer-review paper on credibility assessment or deception detection, can adversely influence the outcome of trials. To set a legal precedent, it only takes one lawyer who uses the research findings and conclusions drawn by Leach et al. (2016) to argue that niqabs improve the trial judges' ability to detect deception and one trial judge from the United States, the United Kingdom, or Canada who does not understand the extent of the limitations within Leach et al., the other shortcomings we highlighted in this comment, and pivotal concepts about lie detection and social interactions.

For example, Leach et al. (2016) concluded that: ‘Discrimination between lie- and truth-tellers was no better than guessing in the latter group, replicating previous research findings (Bond & DePaulo, 2006). It was only when witnesses wore veils (i.e., hijabs or niqabs) that observers performed above chance levels’ (Leach et al., 2016, p. 407). However, even if the overall accuracy rates reported by the authors in the niqab (.55 in Study 1, .57 in Study 2) and hijab conditions (.58 in Study 1, .59 in Study 2) are above chance, they are very weak. The average overall accuracy rates in deception detection research are approximately .54 (Aamodt & Mitchell, 2006; Bond & DePaulo, 2006), where .50 is the overall accuracy rate of chance, and, compared to the niqab and hijab conditions, similar or better overall accuracy rates have been achieved with training (Driskell, 2012; Hauch, Sporer, Michael & Meissner, 2014; Shaw, Porter, & Ten Brinke, 2013). A significant increase in overall accuracy rates is also observed with the use of different techniques such as using contextual information (Blair, Levine, & Shaw, 2010) and strategic interrogative questioning (Levine et al., 2010) as well as the strategic use of evidence (Hartwig et al., 2006).

Although they are not uncommon, the overall accuracy rates reported by Leach et al. (2016) raise intriguing questions about the accuracy of their conclusion that ‘the data consistently suggested that minimizing visual information actually improved participants’ lie detection performance’ (Leach et al., 2016, p. 407). In other words, according to Leach et al., the overall accuracy rates should be higher in the niqab condition than in the hijab condition and be higher in the hijab condition than in the no veil condition. However, while they are higher in the hijab condition (.58 in Study 1, .59 in Study 2) than in the no veil condition (.52 in Study 1, .51 in Study 2), the overall accuracy rates are higher in the hijab condition (.58 in Study 1, .59 in Study 2) than in the niqab condition (.55 in Study 1, .57 in Study 2). The accuracy rates when there are more facial cues (i.e., in the hijab condition) are higher than when there are fewer facial cues (i.e.,

in the niqab condition). This is inconsistent with the conclusion made by Leach et al.

Furthermore, since the hijab covers the neck, ears, and hairs, the no veil condition primarily has, according to the Facial Action Coding System (Ekman & Friesen, 1978), two more facial cue than the hijab condition – that is, the neck tightener (AU21) and swallowing (AU80). Therefore, it would be more accurate to say that data consistently suggested that minimizing the view of the neck improved participants' lie detection performance, not to mention that, contrary to other lie detection studies (Leal, Vrij, Mann, & Fisher, 2010; Suchotzki, Verschuere, Crombez, & De Houwer, 2013), the accuracy rates of each participant were averaged to calculate the overall accuracy rates, resulting in a decrease of the within-subject variance and, in turn, of the mean square error (Abelson, 1995; Pinheiro & Bates, 2000). This methodological choice resulted in an overestimation of the effect sizes reported by Leach et al.

Based on the aforementioned, the standard procedure would be to adapt the experimental court paradigm and to complete a reanalysis of the data. However, whereas research goals are to inform members of the judiciary, we see no purpose of doing so at this point because, contrary to what is implied by Leach et al. (2016), the assumption they want to test (i.e., that niqabs interfere with the trial judges' ability to detect deception) is not consistent with actual courtroom assumptions, especially so with regards to Canada.

Consider *R. v. N.S.* (2012) where the Supreme Court of Canada developed a framework on the issue of when, if ever, a witness who wears a niqab can be required to remove it. When the part of the framework on the right to a fair trial was addressed, the Supreme Court of Canada principally concluded that the niqab may impede the cross-examination by the lawyer and the credibility assessment by the trial judge: two concepts distinct from the concept of deception detection as it is referred to in Leach et al. (2016):

On the record before us, I conclude that there is a strong connection between the ability to see the face of a witness and a fair trial. Being able to see the face of a witness is not the only – or indeed perhaps the most important – factor in cross-examination or accurate credibility assessment. But its importance is too deeply rooted in our criminal justice system to be set aside absent compelling evidence. (*R. v. N.S.*, 2012, pp. 744-745)

More specifically, in *R. v. N.S.* (2012) and *R. v. N.S.* (2010) – the two Canadian decisions cited by Leach et al. (2016) – the word ‘lie’ (and ‘lie detection’) is not mentioned, and the word ‘deception’ (and ‘deception detection’) is not presented as it is referred to in Leach et al. *R. v. N.S.* (2012) mentions it only once when referring to the fact that the niqab may impede cross-examination by the lawyer: ‘Non-verbal communication can provide the cross-examiner with valuable insights that may uncover uncertainty or deception, and assist in getting at the truth’ (para. 24). In other words, the niqab may impede the lawyer’s cross-examination and the trial judge’s credibility assessment, two actual courtroom assumptions not in line with the assumption that Leach et al. wanted to test (i.e., that niqabs interfere with the trial judges’ ability to detect deception).

Thus, with regards to Canada, the authors’ claim that ‘banning the niqab because it interferes with one’s ability to determine whether the speaker is lying or telling the truth is not supported by science’ (Leach et al., 2016, p. 408) is tendentious. The Supreme Court of Canada did not ban the niqab. Irrespective of the conclusions drawn regarding the issue of when, if ever, a witness who wears a niqab can be required to remove it is the *equivalent* of a ban, the Supreme Court of Canada did not conclude that the niqab interferes with the trial judges’ ability to detect deception as it is referred to in Leach et al. (2016). The Supreme Court of Canada concluded that the niqab may impede trial judges’ ability to assess credibility. Even if lie detection scholars regularly use them interchangeably, deception detection is not a synonym of credibility

assessment. By definition, deception is ‘a successful or unsuccessful deliberate attempt, without forewarning, to create in another a belief which the communicator considers to be untrue’ (Vrij, 2008, p. 15). Credibility ‘incorporates competence, or the perceived ability to accurately recall and event, and trustworthiness, and the perceived ability to tell the truth’ (Tabak & Klettke, 2014, p. 47). This confusion is not unexpected considering that Leach et al. uses ‘lie’ and ‘deception’ interchangeably, two other terms that are not synonyms (Galasinski, 2000).

Furthermore, such an adaptation of the experimental court paradigm and the completion of a re-analysis of the data would take several months or years, a period during which Leach et al. (2016) has the potential to cause irremediable harm to the judicial system. Therefore, the objective of this comment was to immediately expose the flaws of a peer-reviewed paper that will likely be brought to the attention of trial judges. This clarification is all the more important considering that Leach et al. received media attention, including that of major Canadian online newspapers (Hopper, 2016; Mercier, 2016), including one where a citation attributed to Leach was that ‘There’s concrete data from over 500 people showing that, in fact, the courts were incorrect’ (Hopper, 2016), a citation that has the power to mislead the general public as well as judicial professionals. However, Leach et al. do not offer compelling evidence to set aside the ‘deeply rooted presumption in our legal system that seeing a witness’s face is important to a fair trial, by enabling effective cross-examination and credibility assessment’ (*R. v. N.S.*, 2012, p. 728).

Conclusively, since deception in courtrooms is largely uncharted territory, future research should establish research hypotheses relevant to real-life court proceedings, highlighting the importance to establish an effective dialogue between scholars and judicial professionals. However, before any attempts to test an assumption with inferential statistics, descriptive research to understand the phenomenon of deception in courtrooms should be conducted (Park et

al., 2002; Rozin, 2001). While this comment is indeed a very critical appraisal of the research by Leach et al. (2016), it is done with the intention of maintaining the integrity of evidence-based practices in courtrooms. To echo Feeley (2013), simple solutions often fail, so, while always allowing the wearing of a niqab in the courtroom, or, for that matter, always requiring the niqab to be removed, may simplify the debate around an already sensitive subject within particular jurisdictions, it is important for psychology and law scholars to acknowledge that consequences of evidence based practice that they recommend can exceed their area of expertise. Such research findings may lessen political, sociological, and cultural implications of evidence based practice and may in fact be detrimental to justice itself.

[Word count: 7042]

References

- Aamodt, M. G., & Mitchell, H. (2006). Who can best catch a liar? A meta-analysis of individual differences in detecting deception. *Forensic Examiner, 15*, 6-11.
- Abelson, R. P. (1995). *Statistics as principal argument*. New York: Psychology Press.
- Baldwin, S. (2013). *Art of Advocacy: Direct Examination*. Dayton: Lexis Nexis.
- Baker, A., Porter, S., Ten Brinke, L., & Mundy, C. (2016). Seeing is believing: observer perceptions of trait trustworthiness predict perceptions of honesty in high-stakes emotional appeals. *Psychology, Crime & Law, 22*(9), 817-831.
<http://dx.doi.org/10.1080/1068316X.2016.1190844>
- Bar Standard Board. (2015). *The BSB handbook*. London: Bar Standard Board.
- Beety, V. E. (2013). Criminality and corpulence: Weight bias in the courtroom. *Seattle Journal for Social Justice, 11*(2), 523-554.
- Blair, J. P., Levine, T. R., & Shaw, A. S. (2010). Content in context improves deception detection accuracy. *Human Communication Research, 36*, 423-442.
<http://dx.doi.org/10.1111/j.1468-2958.2010.01382.x>
- Bogaard, G., Meijer, E. H., Vrij, A., & Merckelbach, H. (2016). Scientific content analysis (SCAN) cannot distinguish between truthful and fabricated accounts of a negative event. *Frontiers in Psychology Psychol, 7*. <http://dx.doi.org/10.3389/fpsyg.2016.00243>
- Bond, C. F., Jr., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10*, 214-234. http://dx.doi.org/10.1207/s15327957pspr1003_2
- Brainerd, C. J. (2013). Developmental reversals in false memory: A new look at the reliability of children's evidence. *Current Directions in Psychological Sciences, 22*, 335-341.
<http://dx.doi.org/10.1177/0963721413484468>
- Brouillard Also Known As Chatel v. The Queen, [1985] 1 SCR 39, 1985 SCC 56.

- Burgoon, J.K., & Levine, T.R. (2010). Advances in deception detection. In S. Smith & S. Wilson (Eds.). *New directions in interpersonal communication* (pp. 201-220). Thousand Oaks: Sage Publications.
- Chin, J. M., & Dallen, S. (2016). R v Awer and the dangers of science in sheep's clothing. *Criminal Law Quarterly*, Forthcoming.
- Denault, V. (2015). *Communication non verbale et crédibilité des témoins* [Nonverbal communication and witness credibility]. Cowansville: Éditions Yvon Blais.
- Denault, V., Delmas, H., & Rochat, N. (2016, July). Credibility assessment of witnesses: Dubious criteria and pseudoscience. 26th Annual Conference of the European Association of Psychology and Law, Université Toulouse - Jean Jaurès, Toulouse, France.
- Denault, V., Larivée, S., Plouffe, D., et Plusquellec, P. (2015). La synergologie, une lecture pseudoscientifique du langage corporel [Synergology, a pseudoscientific reading of body language]. *Revue de Psychoéducation*, 43, 425-455.
- Driskell, J. E. (2012). Effectiveness of deception detection training: A meta-analysis. *Psychology, Crime & Law*, 18, 713-731. <http://dx.doi.org/10.1080/1068316X.2010.535820>
- Du, S., Tao, Y., & Martinez, A. M. (2014). Compound facial expressions of emotion. *Proceedings of the National Academy of Sciences of the United States of America*, 111, E1454–E1462. <http://dx.doi.org/10.1073/pnas.1322355111>
- Dumas, R., & Testé, B. (2006). The influence of criminal facial stereotypes on juridic judgments. *Swiss Journal of Psychology*, 65(4), 237-244.
- Eberhardt, J. L., Davies, P. G., Purdie-Vaughns, V. J., & Johnson, S. L. (2006). Looking deathworthy perceived stereotypicality of black defendants predicts capital-sentencing outcomes. *Psychological Science*, 17(5), 383-386. <http://dx.doi.org/10.1111/j.1467-9280.2006.01716.x>

- Ekman, P., & Friesen, W. V. (1978). *Facial Action Coding System: Investigator's Guide*. Palo Alto : Consulting Psychologists.
- Erickson, W. B., Lampinen, J. M., & Moore, K. N. (2016). Eyewitness identifications by older and younger adults : A meta-analysis and discussion. *Journal of Police and Criminal Psychology, 31*, 108-121. <http://dx.doi.org/10.1007/s11896-015-9176-3>
- Farmer, C., & Hancock, J. (2014). Perjury. In T. Levine (Ed.), *Encyclopedia of Deception* (Vol. 2, pp. 753-756). Los Angeles : Sage.
- Fawcett, H. E. (2014). Witness, False testimony of. In T. Levine (Ed.), *Encyclopedia of Deception* (Vol. 2, pp. 937-940). Los Angeles : Sage.
- Feeley, M. M. (2013). *Court reform on trial: Why simple solutions fail*. New Orleans: Quid Pro Books.
- Frith, C. (2009). Role of facial expressions in social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences, 371*, 3453-3458.
<http://dx.doi.org/10.1098/rstb.2009.0142>
- Galasinski, D. (2000). *The Language of deception: A discourse analytical study*. Thousand Oaks: Sage Publications.
- Garfinkel, H. (1967). *Studies in ethnomethodology*. Cambridge: Polity.
- Granhag, P. A., & Strömwall, L. A. (2004). *The detection of deception in forensic contexts*. Cambridge: Cambridge University Press.
- Hauch, V., Sporer, S. L., Michael, S. W., & Meissner, C.A. (2014). Does training improve the detection of deception? A meta-analysis. *Communication Research, 43*, 283-343.
<http://dx.doi.org/10.1177/0093650214534974>

- Hartwig, M., Granhag, P. A., Strömwall, L. A., & Vrij, A. (2005). Detecting deception via strategic disclosure of evidence. *Law and Human Behavior, 29*, 469-484.
<http://dx.doi.org/10.1007/s10979-005-5521-x>
- Hartwig, M., Granhag, P. A., Strömwall, L., & Kronkvist, O. (2006). Strategic use of evidence during police interviews: When training to detect deception works. *Law and Human Behavior, 30*, 603-619. <http://dx.doi.org/10.1007/s10979-006-9053-9>
- Heath, W. P. (2009). Arresting and convicting the innocent: the potential role of an “inappropriate” emotional display in the accused. *Behavioral Sciences and the law, 27*(3), 313-332. <http://dx.doi.org/10.1002/bsl.864>
- Heath, W. P., & Grannemann, B. D. (2015). Expectations for defendant emotion. *Applied psychology in criminal justice, 11*(2), 126-146.
- Hopper, T. (2016, July 1). Easier to spot a liar in a niqab, says study challenging Canada’s courtroom ban on Muslim veils. *National Post*. Retrieved from <http://news.nationalpost.com/news/canada/easier-to-spot-a-liar-in-a-niqab-says-study-challenging-canadas-courtroom-ban-on-muslim-veils>
- Jacquemet, N., Luchini, S., Rosaz, J., & Shogren, J. (2015). Truth-telling under Oath. *Documents de travail du Centre d'Economie de la Sorbonne*. Retrieved from <https://halshs.archives-ouvertes.fr/halshs-01224135/>
- Jupe, L., Akehurst, L., Vernham, Z., & Allan, J. (2016). Teenage offenders' ability to detect deception in their peers. *Applied Cognitive Psychology, 30*, 401-408.
<http://dx.doi.org/10.1002/acp.3214>
- Kozinski, A. (2015). Criminal law 2.0. *Georgetown Law Review, 44*, iii-xliv.
- Lamb, M. E., Hershkowitz, I., Orbach, Y., & Esplin, P. W. (2008). *Tell me what happened: Structured investigative interviews of child victims and witnesses*. Chichester, UK: Wiley.

- Laney, C., & Loftus, E. F. (2013). Recent advances in false memory research. *South African Journal of Psychology, 43*(2), 137-146. <http://dx.doi.org/10.1177/008124631348423>
- Leach, A. M., Ammar, N. England, D. N., Remigio, L. M., Kleinberg, B., & Verschuere, B. J. (2016). Less is more? Detecting lies in veiled witnesses. *Law and Human Behavior, 40*, 401-410. <http://dx.doi.org/10.1037/lhb0000189>
- Leal, S., Vrij, A., Mann, S., & Fisher, R. P. (2010). Detecting true and false opinions: The devil's advocate approach as a lie detection aid. *Acta psychologica, 134*, 323-329. <http://dx.doi.org/10.1016/j.actpsy.2010.03.005>
- Leary, M. R. (1996). *Self-presentation: Impression management and interpersonal behavior*. Boulder: Westview Press.
- Lefeunteum v. Beaudoin, 28 SCR 89, 1897 SCC 51.
- Leo, R. A. (2009). False confessions: Causes, consequences, and implications. *Journal of the American Academy of Psychiatry and the Law, 37*, 332-343.
- Levine, T. R., Clare, D. D., Green, T., Serota, K. B., & Park, H. S. (2014). The effects of truth-lie base rate on interactive deception detection accuracy. *Human Communication Research, 40*, 350-372. <http://dx.doi.org/10.1111/hcre.12027>
- Levine, T. R., Park, H. S., & McCornack, S. A. (1999). Accuracy in detecting truths and lies: Documenting the "Veracity effect". *Communication Monographs, 66*, 125-144. <http://dx.doi.org/10.1080/03637759909376468>
- Levine, T. R., Shaw, A., & Shulman, H. C. (2010). Increasing deception detection accuracy with strategic questioning. *Human Communication Research, 36*, 216-231. <http://dx.doi.org/10.1111/j.1468-2958.2010.01374.x>
- Lewis, M., Haviland-Jones, J. M., & Barrett, L. F. (2008). *Handbook of emotions*. New York: The Guilford Press.

- Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning and Memory, 12*, 361-366.
<http://dx.doi.org/10.1101/lm.94705>
- MacMillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. New York: Cambridge University Press.
- Mandal, M. K., & Awasthi, A. (2015). *Understanding facial expressions in communication: Cross-cultural and multidisciplinary perspectives*. New Delhi: Springer.
- Mauet, T. A. (2013). *Trial techniques and trials*. New York: Wolters Kluwer.
- McElhaney, J. M. (1997). Evasive witnesses: Use cross-examination to tell your side of the story. *ABA Journal, 83*, 82-83.
- Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law, 7*, 3-35.
<http://dx.doi.org/10.1037/1076-8971.7.1.3>
- Meissner, C. A., & Kassin, S. M. (2002). "He's guilty!": Investigator bias in judgments of truth and deception. *Law and Human Behavior, 26*, 469-480.
<http://dx.doi.org/10.1023/A:1020278620751>
- Mercier, N. (2016, September 30). Le niqab au tribunal: le verdict de la science [The niqab in court : The verdict of science]. *Actualité*. Retrieved from
<http://www.lactualite.com/societe/le-niqab-au-tribunal-le-verdict-de-la-science-2/>
- Michel, C., Caldara, R., & Rossion, B. (2006). Same-race faces are perceived more holistically than other-race faces. *Visual Cognition, 14*, 55-73.
<http://dx.doi.org/10.1080/13506280500158761>
- Milne, R., & Bull, R. (2003). *Investigative interviewing: Psychology and practice*. Chichester, UK: Wiley.

- Morris, M. W., & Keltner, D. (2000). How emotions work: The social functions of emotional expression in negotiations. *Research in Organizational Behavior*, 22, 1-50.
[http://dx.doi.org/10.1016/S0191-3085\(00\)22002-9](http://dx.doi.org/10.1016/S0191-3085(00)22002-9)
- Muhammad v. Enterprise Rent-A-Car*, No. 06–41896-GC (31st D. Mich, Oct. 11, 2006).
- Park, H. S., Levine, T. R., McCornack, S. A., Morrison, K., et Ferrerra, M. (2002). How people really detect lies. *Communication Monographs*, 69, 144-157.
<http://dx.doi.org/10.1080/714041710>
- Pietroni, D., Van Kleef, G. A., De Dreu, C. K. W., & Pagliaro, S. (2008). Emotions as strategic information: Effects of other's emotional expressions on fixed-pie perception, demands, and integrative behavior in negotiation. *Journal of Experimental Social Psychology*, 44, 1444-1454. <http://dx.doi.org/10.1016/j.jesp.2008.06.007>
- Pinheiro, J., & Bates, D. (2000). *Mixed-effects models in S and S-PLUS*. Berlin: Springer Science + Business Media.
- Porter, S., & ten Brinke, L. (2009). Dangerous decisions: A theoretical framework for understanding how judges assess credibility in the courtroom. *Legal and Criminological Psychology*, 14(1), 119-134. <http://dx.doi.org/10.1348/135532508X281520>
- Posey, A. J., & Wrightsman, L. S. (2005). *Trial consulting*. New York: Oxford University Press.
- Read, J. M., Powell, M. B., Kebbell, M. R., & Milne, R. (2009). Investigative interviewing of suspected sex offenders: A review of what constitutes best practice. *International Journal of Police Science and Management*, 11, 442-459.
<http://dx.doi.org/10.1350/ijps.2009.11.4.143>
- Rozin, P. (2001). Social psychology and science: Some lessons from Solomon Asch. *Personality and Social Psychology Review*, 5, 2-14. http://dx.doi.org/10.1207/S15327957PSPR0501_1

Russell, J. A., & Fernandez-Dols, J. M. (1997). *The psychology of facial expression*. Cambridge: Cambridge University Press.

R. v. D.A.I., [2012] 1 SCR 149, 2012 SCC 5.

R. v. François, [1994] 2 SCR 827, 1994 SCC 52.

R. v. Gagnon, [2006] 1 SCR 621, 2006 SCC 17

R. v. Handy, [2002] 2 SCR 908, 2002 SCC 56.

R. v. Marquard, [1993] 4 SCR 223, 1993 SCC 37.

R. v. N.S., [2012] 3 SCR 726, 2012 SCC 72.

R. v. N.S., 2010 ONCA 670.

R. v. R.E.M., [2008] 3 SCR 3, 2008 SCC 51.

R. v. W. (R.), [1992] 2 SCR 122, 1992 SCC 56

Schacter, D. L., & Loftus, E. F. (2013). Memory and law: what can cognitive neuroscience contribute?. *Nature neuroscience*, 16(2), 119-123. <http://dx.doi.org/10.1038/nn.3294>

Schmidt, K. L., & Cohn, J. F. (2001). Human facial expressions as adaptations: Evolutionary questions in facial expression research. *Yearbook of Physical Anthropology*, 44, 3-24. <http://dx.doi.org/10.1002/ajpa.20001>

Seelau, S. M. & Seelau, E. P. (2005). Gender-role stereotypes and perceptions of heterosexual, gay and lesbian domestic violence. *Journal of Family Violence*, 20, 363-371.

Shaw, J., & Porter, S. (2015). Constructing rich false memories of committing crime.

Psychological Science, 26, 291-301. <http://dx.doi.org/10.1177/0956797614562862>

Shaw, J., Porter, S., & Ten Brinke, L. (2013). Catching liars: Training mental health and legal professionals to detect high-stakes lies. *Journal of Forensic Psychiatry & Psychology*, 24, 145-159. <http://dx.doi.org/10.1080/14789949.2012.752025>

- Smith, M. L., Cottrell, G. W., Gosselin, F., & Schyns, P. G. (2005). Transmitting and decoding facial expressions. *Psychological Science, 16*, 184-189. <http://dx.doi.org/10.1111/j.0956-7976.2005.00801.x>
- Strömwall, L. A., & Granhag, P. A. (2003). How to detect deception? Arresting the beliefs of police officers, prosecutors and judges. *Psychology, Crime & Law, 9*, 19-36. <http://dx.doi.org/10.1080/10683160308138>
- Suchotzki, K., Verschuere, B., Crombez, G., & De Houwer, J. (2013). Reaction time measures in deception research: Comparing the effects of irrelevant and relevant stimulus-response compatibility. *Acta psychologica, 144*, 224-231. <http://dx.doi.org/10.1016/j.actpsy.2013.06.014>
- Tabak, S. J., & Klettke, B. (2014). Mock jury attitudes towards credibility, age, and guilt in a fictional child sexual assault scenario. *Australian Journal of Psychology, 66*, 47-55. <http://dx.doi.org/10.1111/ajpy.12035>
- Talwar, V., & Crossman, A. (2012). Children's lies and their detection: Implications for child witness testimony. *Developmental Review, 32*(4), 337-359. <http://dx.doi.org/10.1016/j.dr.2012.06.004>
- The Queen v. D.(R.)* [2013] EW Misc 13 (CC).
- Van Bockstaele, B., Verschuere, B., Moens, T., Suchotzki, K., Debey, E., & Spruyt, A. (2012). Learning to lie: effects of practice on the cognitive cost of lying. *Frontiers in Psychology, 3*, 526-552. <http://dx.doi.org/10.3389/fpsyg.2012.00526>
- Vrij, A. (2008). *Detecting lies and deceit*. Chichester: John Wiley & Sons.
- Vrij, A., Fisher, R. P., & Blank, H. (2015). A cognitive approach to lie detection: A meta-analysis. *Legal and Criminological Psychology*, First publication online. <http://dx.doi.org/10.1111/lcrp.12088>

Vrij, A., & Granhag, P. A. (2012). Eliciting cues to deception and truth: What matters are the questions asked. *Journal of Applied Research in Memory and Cognition*, 1(2), 110-117.

<http://dx.doi.org/10.1016/j.jarmac.2012.02.004>

Vrij, A., Granhag, P. A., Mann, S., & Leal, S. (2011). Outsmarting the liars: Towards a cognitive lie detection approach. *Current Directions in Psychological Science*, 20, 28-32.

<http://dx.doi.org/10.1177/0963721410391245>

Vrij, A., Granhag, P. A., & Porter, S. B. (2010). Pitfalls and opportunities in nonverbal and verbal lie detection. *Psychological Science in the Public Interest*, 11, 89-121.

<http://dx.doi.org/10.1177/1529100610390861>

Wilson, T. D., Gilbert, D. T., & Centerbar, D. B., (2003). Making sense: The causes of emotional evanescence. In I. Brocas & J. Carrillo (Eds.). *The Psychology of Economic Decisions* (pp. 209-233). New York: Oxford University Press.

Wixted, J. T., Read, J. D., & Lindsay, D. S. (2016). The effect of retention interval on the eyewitness confidence-accuracy relationship for eyewitness identification. *Journal of Applied Research in Memory and Cognition*, 5, 192-203. [http://dx.doi.org/10.1007/s10979-](http://dx.doi.org/10.1007/s10979-009-9192-x)

[009-9192-x](http://dx.doi.org/10.1007/s10979-009-9192-x)

Younger, I. (1976). *The art of cross-examination*. Chicago: American Bar Association.

Zebrowitz, L. A., & McDonald, S. M. (1991). The impact of litigants' babyfacedness and attractiveness on adjudications in small claims courts. *Law and Human Behavior*, 15(6),

603-623.