Less is more? Detecting lies in veiled witnesses

Leach, A.-M.; Ammar, N.; England, D.N.; Remigio, L.M.; Kleinberg, B.; Verschuere, B. J.

DOI
10.1037/lhb0000189

Publication date
2016

Document Version
Final published version

Published in
Law and Human Behavior

Law and Human Behavior
2016, Vol. 40, No. 4, 401–410

0147-7307/16/$12.00 http://dx.doi.org/10.1037/lhb0000189

Less Is More? Detecting Lies in Veiled Witnesses

Amy-May Leach, Nawal Ammar,
D. Nicole England, and Laura M. Remigio
University of Ontario Institute of Technology

Bennett Kleinberg and Bruno J. Verschuere
University of Amsterdam

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 (*Muhammad v. Enterprise Rent-A-Car*, 2006; *R. v. N. S.*, 2010; *The Queen v. D*(R), 2013). In two studies, we used conventional research methods and safeguards to empirically examine the assumption that niqabs interfere with lie detection. Female witnesses were randomly assigned to lie or tell the truth while remaining unveiled or while wearing a hijab (i.e., a head veil) or a niqab (i.e., a face veil). In Study 1, laypersons in Canada ($N = 232$) were more accurate at detecting deception in witnesses who wore niqabs or hijabs than in those who did not wear veils. Concealing portions of witnesses' faces led laypersons to change their decision-making strategies without eliciting negative biases. Lie detection results were partially replicated in Study 2, with laypersons in Canada, the United Kingdom, and the Netherlands ($N = 291$): observers' performance was better when witnesses wore either niqabs or hijabs than when witnesses did not wear veils. These findings suggest that, contrary to judicial opinion, niqabs do not interfere with—and may, in fact, improve—the ability to detect deception.

Keywords: lie detection, Muslims, witnesses, veiling, minimal information

Wearing a niqab—a veil that covers the wearer's face, except for her eyes—is increasingly prevalent, but contentious. In the 1970s, 1% of the Muslim population wore face veils; currently, approximately one third of female Muslims engage in the practice (see Figure 1 for sample niqab; al-Ghazali, 2008). There has been considerable debate about the appropriateness of wearing a niqab (e.g., Khiabany & Williamson, 2008; Mistry, Bhugra, Chaleby, Khan, & Sauer, 2009; Vakulenko, 2007). In fact, the wearing of a niqab has been officially banned from all public places in several countries, such as Belgium, Egypt, France, and Turkey (Loi n° 2010–1192 du 11 octobre 2010 interdisant la dissimulation du visage dans l'espace public, 2010; Loi visant à interdire le port de tout vêtement cachant totalement ou de manière principale le visage, 2011; Syed, 2010).

The permissibility of the niqab has also been called into question by the courts. For example, a judge in an American small claims court dismissed a plaintiff's complaint when she refused to remove her veil in order to testify (*Muhammad v. Enterprise Rent-A-Car*, 2006). Similarly, in Canada, an alleged victim of childhood sexual assault was ordered to testify at a preliminary inquiry without her niqab (*R. v. N. S.,* 2010). Most recently, a defendant who had been charged with witness intimidation was directed to remove her niqab while presenting evidence in the United Kingdom (*The Queen v. D(R)*, 2013). In banning the wearing of a niqab while testifying, the various courts attempted to balance the need to establish a witness's identity, the strength of the women's religious beliefs, and the right to freedom of religion. Ultimately, however, the right to a fair trial—and the threat to that right posed by allowing a witness to wear the niqab while testifying—appeared to override the witnesses' right to veil. The presiding judges opined that, in an adversarial trial, a judge must be able to see a witness's face to assess her truthfulness.

Considerable psychology–law research has been devoted to testing assumptions underlying legal decisions and laws. For example, Wells and Quinlivan (2009) found that beliefs about human cognition, which formed the basis of the U.S. Supreme Court's decision on how to evaluate claims of suggestiveness of police lineups in *Manson v. Braithwaite* (1977), were inconsistent with contemporary research findings in the eyewitness identification literature. In the current studies, we examined the notion embodied in important court decisions in the United States, the United Kingdom, and Canada (e.g., *N. S. v. Her Majesty the Queen et al.*, 2012): that a fact-finder's ability to detect deception among witnesses is compromised by the niqab.

Typically, observers' lie detection performance is poor. Average accuracy for laypersons and justice officials is very close to 50%, or chance levels (Aamodt & Custer, 2006; Bond & DePaulo, 2006). In the majority of lie detection studies, however, lie-tellers'

This article was published Online First June 27, 2016.

Amy-May Leach, Nawal Ammar, D. Nicole England, and Laura M. Remigio, Faculty of Social Science and Humanities, University of Ontario Institute of Technology; Bennett Kleinberg and Bruno J. Verschuere, Clinical Psychology, University of Amsterdam.

This research was supported by Social Sciences and Humanities Research Council of Canada Grant 430-2011-0407 to Amy-May Leach and Nawal Ammar. We also thank Brian Cutler for his comments on the manuscript.

Correspondence concerning this article should be addressed to Amy-May Leach, University of Ontario Institute of Technology, Faculty of Social Science and Humanities, 2000 Simcoe Street North, Oshawa, Ontario, Canada. E-mail: amy.leach@uoit.ca

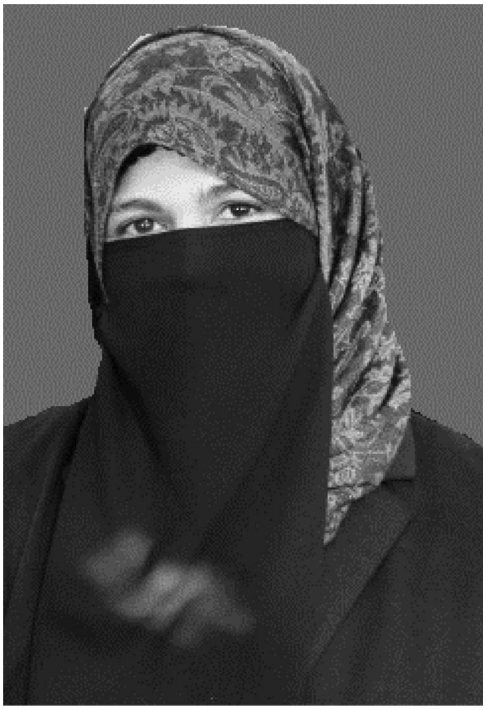

Figure 1. One type of veil—the niqab—that is the focus of this research project.

and truth-tellers' faces were visible. Our literature review uncovered no previously published research on the effects of religious garments on lie detection.

A few studies were indirectly relevant to peoples' abilities to detect deception among veiled witnesses. Research on cross-cultural lie detection was informative, for example. It is highly unlikely that every single defendant, plaintiff, witness, and decision-maker (e.g., juror, judge) involved in a case would wear a niqab. Although viewing veiled witnesses would not constitute cross-cultural lie detection, observers' inexperience with veiling might be analogous to a lack of familiarity between cultures. In two studies, laypersons were slightly better at detecting deceit within their own cultures than across cultures (Bond & Atoum, 2000; Bond, Omar, Mahmoud, & Bonser, 1990). Yet, Vrij and Winkel (1991) failed to find significant differences between Dutch and Surinamese observers' cross-cultural lie detection. Given the limited number of studies and mixed findings (see Taylor, Larner, Conchie, & van der Zee, 2014, for a full discussion), it remains unclear from this literature whether there is a meaningful disadvantage to detecting deception in witnesses in niqabs.

Instead, it might be important to focus on the more limited information that is afforded by niqabs as compared to bare faces. People are able to make inferences based on minimal information. For example, point-light displays of biological motion and static images of facial features can be used to discern others' attributes (Baron-Cohen, Wheelwright, & Jolliffe, 1997; Blais, Roy, Fiset, Arguin, & Gosselin, 2012; Troje, 2002). Similarly, another form of minimization, thin slices (i.e., exposure to less than 5 minutes of behavior) reveal performance, interpersonal relationships, and individual differences (e.g., Ambady, Bernieri, & Richeson, 2000; Ambady & Rosenthal, 1992). Focusing on general impressions might discourage the use of irrelevant details and increase efficient, intuitive processing without taxing cognitive resources (Ambady, 2010; Murphy & Balzer, 1986).

Minimization of information principles have been applied to lie detection. In one direct test, observers who were afforded brief glimpses of behavior were more accurate than observers who viewed lie- and truth-tellers' entire accounts (Albrechtsen, Meissner, & Susa, 2009). It is not necessarily that this approach encourages unconscious decision-making, but rather that it focuses observers on a limited number of diagnostic cues (Street & Richardson, 2014). These minimization effects have also been examined in terms of the medium of presentation (i.e., audio vs. visual vs. audiovisual; Burgoon, Blair, & Strom, 2008; Zuckerman, Koestner, & Colella, 1985). This research might be the most relevant to the study of face veiling because it involves restricting the nonverbal and verbal cues that are available for decision-making. A meta-analysis of 50 studies (Bond & DePaulo, 2006) revealed that overall lie detection accuracy was similar, whether observers received audio (i.e., more restricted) or audiovisual (i.e., less restricted) information. To date, there has been no research on how exposure to the full range of verbal cues, but only a subset of nonverbal cues, affects performance.

We found no empirical evidence in the lie detection literature suggesting that a niqab should impair lie detection because it conceals portions of the wearers' face; rather, existing research suggests that the opposite could occur. Niqabs should minimize the amount of information that is available to observers and prevent them from basing their lie detection decisions on misleading facial cues (e.g., smiling; DePaulo et al., 2003). In turn, the veiling of the witness might force observers to attend to sources of information that are more diagnostic of deception, such as verbal content (Vrij, 2008). Niqabs also explicitly highlight a specific subset of nonverbal cues (i.e., the witnesses' eyes). It is widely believed that a person's eyes reveal deception (The Global Deception Research Team, 2006), and the eye region can be used to identify complex mental states (e.g., guilt; Baron-Cohen et al., 1997). In particular, blinking, and pupil dilation are effective cues to deceit in certain contexts (e.g., DePaulo et al., 2003; Leal & Vrij, 2008). By encouraging the use of verbal cues and/or eye region cues, niqabs could actually facilitate the detection of deception.

Although niqabs may lead to improved lie detection performance, they might also elicit response biases. Veiled Muslim women report being stared at, insulted, and assaulted (Clarke, 2013). Being Muslim is associated with an aggressive stereotype (Fischer, Greitemeyer, & Kastenmuller, 2007), and niqabs are particularly threatening (Behiery, 2013). It is possible that people attribute a range of other negative behaviors, such as lying, to women who wear veils (see Hoodfar, 1997 for a similar argument about head veiling). The typical dark color of niqabs could also invoke the black clothing stereotype, in which individuals who wear dark (vs. light) colors are less likely to be judged as credible (Akehurst, Kohnken, Vrij, & Bull, 1996). In turn, there might be a tendency to label women who wear niqabs as lie-tellers, regardless of the underlying veracity of those witnesses' accounts. Maeder, Dempsey, and Pozzulo (2012) examined whether an alleged sexual assault victim's veiling influenced the perceived culpability of the defendant. In their study, the victim was described as wearing a burqa (i.e., a garment in a solid color that, in addition to covering the hair and face, conceals the entire body), a hijab, or no veil. Mock jurors, who read transcripts of the victim's testimony, were more confident in the defendant's guilt when the victim was described as wearing a burqa or hijab than when she did not veil. In sum, various sources directly and indirectly associated with veiling led us to examine the possibility that niqabs produced a response bias during lie detection. No research, to date, has examined the effects of actively minimizing only a subset of nonverbal cues—while highlighting others—on response bias.

Study 1

In Study 1, we examined participants' lie detection accuracy, response biases, and decision strategies when evaluating the testimony of eyewitnesses in three veiling conditions: niqab, hijab, and no veil. We hypothesized that lie detection accuracy would be higher in the niqab condition than in the hijab or no-veil conditions because it would minimize the availability of misleading cues to deception and facilitate the use of more effective strategies (e.g., Albrechtsen et al., 2009). In addition, we predicted that veils would activate stereotypically negative views of Muslim women (e.g., Fischer et al., 2007); therefore, we expected a lie bias (i.e., tendency to indicate that witnesses were lying) in the niqab and hijab conditions but not in the no-veil condition. Given that niqabs are portrayed less positively than hijabs (Behiery, 2013), we hypothesized that the lie bias would be stronger in the niqab condition than in the hijab condition. Finally, we conducted exploratory analyses to determine whether expected lie detection effects could be accounted for by participants in the niqab condition attending to witnesses' eyes and the content of their accounts (i.e., verbal cues) to a greater extent than participants who were able to see witnesses' entire faces (i.e., the hijab and no-veil conditions) and/or the witnesses' actual nonverbal and verbal behaviors.

Method

Participants. Two hundred and 32 students at a Canadian university (138 females, 94 males; M age = 20.09 years, SD = 3.83) completed the study in exchange for course credit. Participants self-identified as belonging to the following ethnic groups: Arab/West Asian (n = 22), Black (n = 25), Chinese (n = 8), White (n = 74), Hispanic (n = 1), Korean (n = 1), Latin American (n = 3), South Asian (n = 79), South East Asian (n = 10), other (n = 9). The majority of participants (n = 223) did not wear hijabs or niqabs and self-identified as Christians (n = 95).

Study design. We employed a 2 (veracity: lie-tellers vs. truth-tellers) × 3 (veiling condition: niqab vs. hijab vs. no veil) mixed-factors design. We manipulated veiling condition between participants to decrease the potential impact of demand characteristics.

Materials.

Video footage. In individual sessions, female witnesses (N = 80, M age = 20.23 years, SD = 5.74) were shown a video of a woman who was watching a stranger's bag. As determined by random assignment, half of the women also observed her stealing items from the bag. Then, all of the 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). Thus, half of the witnesses were lying and half of the witnesses were telling the truth. Witnesses were given 2 minutes to prepare their testimony and, as in real trials, 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.

Witnesses were interviewed by two female experimenters. To simulate courtroom procedures, 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?"). Role assignment was counterbalanced, and both experimenters were blind to the veracity of the witness's testimony. To increase the stakes associated with deception, witnesses were told that they might receive $50 if they convinced both experimenters that they were telling the truth. In fact, all of the witnesses were given the opportunity to win the incentive in a draw. At the end of each session, witnesses rated their perceptions of their interviews. These data are available from the corresponding author.

The interview was the only portion of the session that was videotaped. We excluded data from 19 witnesses because the quality of the video footage was poor or their garments (i.e., veils or shawls) were askew. In addition, one witness in the hijab condition confessed to having seen the woman steal items (i.e., she did not follow our instruction to lie about the theft). Clips were not selected based on the cues revealed by witnesses. The full range of lie-telling and truth-telling proficiency is present in the justice system (i.e., there is a continuum from poor to proficient truth-telling, and poor to proficient lie-telling). Generalizability was ensured by randomly assigning witnesses to condition and presenting all of their responses to observers, regardless of their quality. In total, we compiled clips of 10 lie-tellers and 10 truth-tellers in each veiling condition (M length per interview = 2.06 min, SD = 0.37). Demographic characteristics of the witnesses were similar across conditions. Clip order was randomly assigned and counterbalanced within each condition, producing two versions of each set of 20 videos.

Coding. All of the videos were coded for the onsets and offsets of nonverbal and verbal cues using Datavyu (i.e., a video coding and data visualization tool; Datavyu Team, 2014). In addition, research assistants coded the degree to which certain cues occurred (see Appendix>). One research assistant coded all of the footage, whereas another research assistant coded 25% of each video, as recommended by Datavyu. Interrater reliability, calculated using ICCs, was high (M = .86, SD = .16).

Lie detection measure. Participants were asked to indicate whether the witness in the video clip was lying or telling the truth about having seen the woman steal the items. Participants were awarded a "1" for each correct response and a "0" for each incorrect response. Then, all scores were averaged to determine overall accuracy, resulting in a score between 0 (*no lie detection ability*) and 1 (*perfect lie detection ability*).

Confidence measure. Using a scale, from 0% (*not at all confident*) to 100% (*extremely confident*), participants indicated how confident they were in each lie detection judgment. Ratings were averaged across witnesses to yield overall confidence scores. Because confidence analyses were exploratory in nature, and they did not reveal significant effects, we will not report them here. Interested readers can obtain the data from the corresponding author.

Cue use measure. Participants were asked to indicate which verbal cues (e.g., amount of detail) and nonverbal cues (e.g., eye contact) they used to make their decisions from the same list containing empirically verified actual and perceived cues to deception that was coded by research assistants. For each cue, participants were given a "1" if they indicated that they had used that cue to make their decisions, and a "0" if they had not used the cue. Each variable was classified as a nonverbal or verbal cue. Then, we calculated a difference score (i.e., subtracted overall verbal from nonverbal cue use).

Experience measure. We asked participants to report their experiences with lie detection on a scale from 1 (*not at all*) to 5 (*extremely*) and describe any relevant additional experience in the area. Participants also indicated whether they had ever worked in law enforcement and had taken any courses related to lie detection.

Procedure. Given that the two variables were manipulated in the videotaped stimuli, the procedure was identical for all participants. Individually, following random assignment to one of the three veiling conditions, we showed each participant 20 video clips of lying and truth-telling witnesses using a computer program (Jarvis, 2008). After each clip, the participant rendered a lie detection and confidence rating. At the end of the session, the participant indicated the cues that he or she had used to make decisions and provided demographic information. Each session lasted approximately one hour.

Results

Preliminary analyses revealed nonsignificant effects of participant gender, race, veiling, religious affiliation, and lie detection experience. All reported analyses are collapsed across those factors.

Participants' accuracy. We conducted a Veracity × Veiling Condition analysis of variance (ANOVA) on accuracy scores. The cell means are shown in Table 1. There was a significant main effect of veiling condition, $F(2, 229) = 9.07$, $p < .001$, $\eta_p^2 = .07$. Post hoc tests, using Tukey's honest significant difference, revealed that participants were more accurate when viewing witnesses who wore hijabs or niqabs than those who did not wear veils, $p < .001$, $d = 0.63$, 95% confidence interval (CI) [0.30, 0.96], and $p = .038$, $d = 0.33$, 95% CI [0.01, 0.65], respectively. There was no significant difference, in terms of overall accuracy, between participants in the hijab and niqab conditions, $p = .175$, $d = 0.32$, 95% CI [−0.00, 0.63]. Regardless of veiling condition, participants were more accurate when judging truth-tellers ($M = .72$, $SD = .20$) than lie-tellers ($M = .38$, $SD = .21$), $F(1, 229) = 210.24$, $p < .001$, $\eta_p^2 = .48$, $d = 1.66$, 95% CI [1.45, 1.87]. However, there was no significant interaction between veracity and veiling condition, $F(2, 229) = 1.19$, $p = .306$, $\eta_p^2 = .01$.

Participants' signal detection. As noted by Meissner and Kassin (2002), focusing solely on accuracy can obscure the distinction between discrimination (i.e., the ability to identify lie- and truth-tellers) and response bias (i.e., the tendency to choose a particular response). Given our interest in both of these factors, we followed Meissner and Kassin's example and conducted a signal detection analysis. All calculations were based on Wixted and Lee's (2014) formulas. Specifically, we calculated "hits" (i.e., the percentage of correct classifications of lie-tellers) and "false alarms" (i.e., the percentage of truth-tellers incorrectly classified as lie-tellers).

Discrimination. We conducted a one-way ANOVA, with veiling condition as the independent variable, on discrimination (i.e., d'). Echoing the overall accuracy analysis, there was a significant effect of veiling, $F(2, 229) = 8.18$, $p < .001$, $\eta_p^2 = .07$ (see Table 1). Post hoc tests revealed that participants were better able to discriminate between lie-tellers and truth-tellers in hijabs than those who did not wear veils, $p < .001$, $d = 0.63$, 95% CI [0.30, 0.95]. The difference between participants in the niqab and no-veil conditions approached significance, $p = .056$, $d = 0.38$, 95% CI [0.06, 0.70]. Performance was similar when participants viewed witnesses wearing hijabs or niqabs, $p = .196$, $d = 0.26$, 95% CI [−0.05, 0.58].

We also compared participants' d' scores to zero (i.e., no sensitivity). Participants could discriminate between lie- and truth-telling witnesses who wore niqabs, $t(77) = 5.18$, $p < .001$, $d = 0.59$, 95% CI [0.34, 0.83], or hijabs, $t(76) = 6.84$, $p < .001$, $d =$

Table 1
Lie Detection Performance

Veiling condition	Accuracy *M* (*SD*)	Discrimination (d') *M* (*SD*)	Discrimination (d') 0 (no sensitivity)	Response bias (β) *M* (*SD*)	Response bias (β) 1[a] (no bias)
Study 1					
Niqab	$.55_a$ (.09)	$.23_{ab}$ (.39)	>	$.94_a$ (.31)	—
Hijab	$.58_a$ (.10)	$.34_a$ (.44)	>	$.86_a$ (.37)	<
No veil	$.52_b$ (.09)	$.08_b$ (.39)	—	$.99_a$ (.38)	—
Study 2					
Niqab	$.57_a$ (.12)	$.30_a$ (.52)	>	$.93_{ab}$ (.31)	<
Hijab	$.59_a$ (.11)	$.43_a$ (.47)	>	$.86_a$ (.26)	<
No veil	$.51_b$ (.13)	$.03_b$ (.39)	—	1.01_b (.41)	—

Note. Means sharing a common subscript are not statistically different at α = .05 according to Tukey's honest significant difference tests. Dashes in column 4 indicate that there was no difference from 0 and dashes in column 6 indicate that there was no difference from 1.
[a] Means significantly less than 1 indicate a truth bias.

0.78, 95% CI [0.39, 1.17]; however, they were unable to do so in the no-veil condition, $t(76) = 1.70$, $p = .094$, $d = 0.19$, 95% CI [−0.97, 0.29].

Response bias. A one-way ANOVA revealed that participants' response biases (i.e., β) were not affected by veiling condition, $F(2, 229) = 2.10$, $p = .126$, $\eta_p^2 = .02$ (see Table 1). By comparing β scores to one (i.e., no response bias), we could examine participants' tendencies to label witnesses as lie-tellers or truth-tellers. Participants exhibited a truth bias toward witnesses who were wearing hijabs, $t(76) = -3.23$, $p = .002$, $d = -0.37$, 95% CI [−0.60, −0.14]. There was no evidence of bias in the niqab condition, $t(77) = -1.66$, $p = .101$, $d = -0.19$, 95% CI [−0.28, −0.09] or in the no-veil condition, $t(76) = -0.30$, $p = .767$, $d = -0.03$, 95% CI [−0.05, −0.02].

Participants' cue use. Although not a primary research question, we were interested in whether participants' lie detection performance could be explained by the cues that they reported using to render decisions. We conducted a multivariate analysis of variance (MANOVA) on participants' self-reported reliance on the eye region (i.e., blinking, eye contact, and pupil dilation) to detect deception. There was only a significant effect of veiling condition on the combined dependent variables, $F(6, 456) = 2.56$, $p = .019$; Pillai's trace = .07 $\eta_p^2 = .03$. Examining the univariate effects revealed that veiling did not affect the use of eye contact, $F(2, 229) = 0.17$, $p = .848$, $\eta_p^2 = .00$, or pupil dilation, $F(2, 229) = 0.99$, $p = .373$, $\eta_p^2 = .01$, as cues to deceit. In fact, eye contact was frequently cited as a cue to deception across all conditions ($M = .92$, $SD = .27$). However, blinking use did vary with veiling condition, $F(2, 229) = 1.44$, $p = .003$, $\eta_p^2 = .05$. Post hoc tests indicated that participants were equally likely to report that they used blinking to detect deception when the witnesses wore niqabs ($M = .60$, $SD = .49$), or did not veil ($M = .60$, $SD = .49$), $p = .998$, $d = 0.00$, 95% CI [−0.32, 0.32]. Participants stated that they relied less on witnesses' blinking in the hijab condition ($M = .36$, $SD = .48$) than in the niqab, $p = .008$, $d = -0.49$, 95% CI [−0.82, −0.17], or no-veil conditions, $p = .010$, $d = -0.49$, 95% CI [−0.82, −0.17].

Veiling also affected overall reported cue use, $F(2, 229) = 14.75$, $p < .001$, $\eta_p^2 = .11$. Participants were more likely to state that they based their decisions on verbal cues than nonverbal cues when witnesses wore niqabs ($M = -.17$, $SD = .10$) than when they wore hijabs ($M = -.10$, $SD = .08$), $d = -0.77$, 95% CI [−1.10, −0.44], or did not veil ($M = -.10$, $SD = .08$), $d = -0.77$, 95% CI [−1.10, −0.44], all $ps < .001$. There were no differences, in terms of overall cue use, when participants viewed witnesses who wore niqabs or did not veil, $p = 1.000$, $d = 0.00$, 95% CI [−0.31, 0.32].

Coded cues. We performed the same analyses on the eye region as above to allow for a comparison between participants' self-reports and the actual presence of cues to deception. Only the effect of veracity was statistically significant, $F(2, 53) = 3.84$, $p = .028$; Pillai's trace = .13 $\eta_p^2 = .67$. Veracity did not affect blinking, $F(1, 54) = 0.06$, $p = .812$, $\eta_p^2 = .06$. However, lie-tellers ($M = 26.83$, $SD = 10.81$) made eye contact less frequently than truth-tellers ($M = 34.20$, $SD = 10.72$), $F(1, 54) = 7.67$, $p = .008$, $\eta_p^2 = .78$, $d = -0.68$, 95% CI [−1.21, −0.15].

To determine whether one type of cue was more likely to occur, the data were transformed into *z* scores and each variable was classified as a nonverbal or verbal cue. Then, we calculated a difference score (i.e., subtracted overall verbal from nonverbal cue use). A Veracity × Veiling Condition ANOVA on the difference score data revealed a significant main effect of veiling, $F(1, 54) = 8.15$, $p = .001$, $\eta_p^2 = .23$. Witnesses who wore niqabs ($M = .32$, $SD = .36$) were more likely to reveal verbal (vs. nonverbal) information than witnesses who wore hijabs ($M = -.26$, $SD = .60$), $p = .001$, $d = 1.18$, 95% CI [0.49, 1.87], or did not veil ($M = -.06$, $SD = .36$), $p = .031$, $d = 1.06$, 95% CI [0.38, 1.75]. There were no differences between witnesses who wore hijabs or did not veil, $p = 1.000$, $d = 0.00$, 95% CI [−0.31, 0.32]. There were no other significant effects.

We also conducted an exploratory Veracity × Veiling Condition MANOVA on the overall presence of empirically verified cues to deception. Participants rated both diagnostic and nondiagnostic cues to reduce any effects of demand characteristics; only the former were analyzed here. Including known nondiagnostic cues in the analysis would have unnecessarily impeded the likelihood of uncovering significant effects. Interested readers can obtain these analyses from the corresponding author, however. In total, we analyzed 15 cues (i.e., fidgeting, inconsistencies, admitted lack of memory, length of response, negative statements, spontaneous corrections, unfriendly facial expressions, word/phrase repetitions, vocal tension, coherence/plausibility, vagueness, cooperativeness, nervousness, amount of detail, and pitch). There was no significant interaction between veracity and veiling condition, $F(28, 84) = 0.88$, $p = .639$; Pillai's trace = .45 $\eta_p^2 = .23$, nor a significant main effect of veiling on the combined dependent variables, $F(28, 84) = 1.02$, $p = .449$; Pillai's trace = .51 $\eta_p^2 = .25$. However, there was a statistically significant difference between lie-tellers and truth-tellers, $F(14, 41) = 2.46$, $p = .013$; Pillai's trace = .46 $\eta_p^2 = .46$. A closer examination of the univariate effects revealed that lie-tellers spoke in a higher pitch, were less cooperative, and provided accounts that were less coherent than truth-tellers (see Table 2).

Discussion

As predicted, participants were more accurate when witnesses wore niqabs than when witnesses did not wear veils. They did not, however, exhibit different response biases toward the former group. Our sample consisted of participants from an ethnically and religiously diverse student population at a Canadian university. Perhaps the characteristics of the student body, exposure to a cross-cultural curriculum, and/or social desirability concerns could account for the lack of response bias toward witnesses wearing veils. Indeed, the demographic composition and diversity of religious life in Canada, as well as the country's historical endorsement of a cultural mosaic approach to multiculturalism, might have made biased decision-making unlikely. It may be more prevalent in geographical regions where the niqab has been more publically opposed and expressing a negative response bias would be more socially acceptable.

The lack of bias in the current study could also be attributed to the context in which data collection took place. *R. v. N. S.* (2010), the Canadian case in which a witness was asked to remove her veil in court, was tried within the university's catchment area. There were numerous appeals and a Supreme Court trial that took place in the midst of data collection (see *N. S. v. Her Majesty the Queen et al.*, 2012). Simultaneously, a neighboring provincial govern-

Table 2
Mean Nonverbal and Verbal Behaviors by Veracity

Behaviors	Lie-tellers *M* (*SD*)	Truth-tellers *M* (*SD*)	*d* [CI]	*p*
Unfriendly facial expressions	0.90 (1.24)	0.70 (1.75)	.13 [−.38, .65]	.600
Fidgeting	1.47 (2.64)	0.60 (1.25)	.42 [.07, .52]	.111
Overall nervousness	3.87 (1.01)	3.73 (0.83)	.15 [.06, .52]	.557
Word or phrase repetitions	0.67 (1.12)	.53 (1.33)	.11 [.06, .52]	.680
Pitch	3.13 (0.35)	2.77 (0.54)	.79 [.25, 1.33]	.001
Vocal tension	2.07 (1.02)	1.93 (0.87)	.15 [−.37, .67]	.559
Length of responses	23.60 (0.77)	23.43 (0.94)	.20 [−.32, .71]	.467
Coherence	4.73 (0.52)	5.00 (0.00)	−.73 [−1.27, −.20]	.008
Amount of detail	2.60 (1.10)	2.70 (0.92)	−.09 [−.61, .41]	.706
Spontaneous corrections	1.40 (1.38)	1.23 (1.33)	.13 [−.39, .64]	.633
Admitted lack of memory	0.27 (0.83)	0.07 (0.37)	.31 [−.21, .83]	.236
Inconsistencies	0.07 (0.25)	0.00 (0.00)	.40 [−.12, .92]	.139
Vagueness	2.63 (1.27)	2.90 (1.35)	−.21 [−.72, .31]	.437
Negative statements	0.00 (0.00)	0.00 (0.00)		
Cooperativeness	4.87 (0.35)	5.00 (0.00)	−.53 [−1.05, −.00]	.044

Note. CI = confidence interval.

ment drafted Bill 60 (2013), which limited State employees' abilities to wear overt religious symbols and conceal their faces; in essence, it would have severely restricted wearing of the niqab. In response, the university's local hospital mounted a recruitment campaign—including ads and signs on city streets—depicting a medical professional wearing a hijab next to the slogan, "We don't care what's on your head. We care what's in it" (Mok, 2013). Thus, participants were exposed to explicit messages from local authorities that veiling should not bias decision-making, in addition to significant public debate surrounding the permissibility of wearing niqabs in court. That exposure might have affected any preexisting response biases.

Study 2

Study 2 served as both a direct replication (in Canada) as well as an extension to two other countries (i.e., the Netherlands and the United Kingdom). We chose the Netherlands because its government recently came very close to banning the niqab (Government of the Netherlands, 2012). We also sought to replicate findings in the United Kingdom because a ruling on the permissibility of wearing a niqab in British courts was imminent. Indeed, shortly after data collection began, a judge ruled that a woman must unveil in court (*The Queen v. D(R)*, 2013). A comparison between these locations and Canada would further test the generalizability of results.

As in Study 1, we expected that participants would be better able to detect deception in witnesses who wore niqabs than in witnesses who did not wear veils. We hypothesized that response bias would vary by region. Specifically, participants in Canada were expected to exhibit similar response tendencies, regardless of veiling condition, replicating the findings from the first study. We posited that the participants in the Netherlands would exhibit the originally hypothesized pattern of response due to the government's stance on veiling. Dutch participants were expected to be more likely to indicate that women were lying when they were wearing niqabs than when they did not wear veils. We did not have set hypotheses about the nature of response bias in the U.K. sample.

Method

Participants. Two hundred and 91 students at universities in Canada, the United Kingdom, and the Netherlands (201 females, 90 males; *M* age = 21.11 years, *SD* = 4.33) completed the study in exchange for extra credit or a small honorarium. Participants self-identified as belonging to the following ethnic groups: Arab/West Asian (n = 19), Black (n = 20), Chinese (n = 7), White (n = 194), Hispanic (n = 1), Korean (n = 1), Latin American (n = 1), South Asian (n = 26), South East Asian (n = 13), Other (n = 9). The majority of participants (n = 286) did not report wearing any type of veil or having a religious affiliation (n = 171).

Study design. We employed a 2 (veracity: lie-tellers vs. truth-tellers) × 3 (country: Canada vs. the Netherlands vs. the United Kingdom) × 3 (veiling condition: niqab vs. hijab vs. no veil) mixed-factors design. As in Study 1, veiling condition was a between-participants factor, whereas veracity was a within-participants factor.

Materials and procedure. The procedure was similar to Study 1, with a few key differences. Participants in the United Kingdom and the Netherlands were not asked to provide the cues that they used to render their lie detection decisions. In addition, we assessed the English proficiency of participants in the Netherlands, using the criteria established by the Centre for Canadian Language Benchmarks (2010), to ensure that they could understand the witnesses' accounts. Dutch participants were asked to self-report their overall English proficiency on a 12-point scale (*Basic* = 1–3; *Intermediate* = 4–8; *Advanced* = 9–12). Average proficiency was on the boundary between Intermediate and Advanced (*M* = 8.87, *SD* = 1.88). At the conclusion of each session, participants listened to two messages that were read aloud in English. After each message, they were asked three multiple-choice questions about its content. Each correct answer was awarded a "1," whereas each incorrect answer was awarded as "0." Thus, the highest possible score was 6 (out of 6 questions). Participants' objective language comprehension was extremely high (*M* = 5.02, *SD* = 0.95) and would be considered "Advanced" according to the Canadian Language Benchmarks. Our ANOVA revealed that there was a similar distribution of English comprehension scores across veiling conditions, $F(1, 96) = 1.94$, $p = .150$, $\eta_p^2 = .04$.

Results

There were nonsignificant effects of race, gender, veiling, religious affiliation, and lie detection experience. Thus, we collapsed across those variables when conducting the following analyses.

Participants' accuracy. A Veracity × Country × Veiling Condition ANOVA indicated that there was a significant main effect of veiling condition, $F(2, 281) = 13.28$, $p < .001$, $\eta_p^2 = .09$ (see Table 1). Post hoc tests revealed that participants were better able to detect the deception of women who wore niqabs or hijabs than of those who did not veil, $p = .001$, $d = 0.48$, 95% CI [0.19, 0.77] and $p = .001$, $d = 0.66$, 95% CI [0.37, 0.95], respectively. Performance in the niqab and hijab conditions was similar, $p = .392$, $d = -0.17$, 95% CI [−0.46, 0.11]. In addition, participants were more accurate when judging truth-tellers ($M = .71$, $SD = .22$) than lie-tellers ($M = .39$, $SD = .20$), $F(1, 281) = 225.14$, $p < .001$, $\eta_p^2 = .45$, $d = 1.52$, 95% CI [0.18, 1.34]. There was no significant main effect of country, $F(2, 281) = 1.44$, $p = .240$, $\eta_p^2 = .01$. Interactions between veracity and veiling condition, $F(2, 281) = 0.13$, $p = .878$, $\eta_p^2 = .00$, veracity and country, $F(2, 281) = 2.70$, $p = .069$, $\eta_p^2 = .02$, country and veiling condition, $F(2, 281) = 0.88$, $p = .475$, $\eta_p^2 = .01$, and all three variables, $F(4, 281) = 1.06$, $p = .376$, $\eta_p^2 = .02$ were also nonsignificant.

Participants' signal detection. As in Study 1, we used a signal detection analysis to examine the independent contributions of discrimination and bias.

Discrimination. We performed a Country × Veiling Condition ANOVA on discrimination (i.e., d'). Again, there was a significant effect of veiling condition, $F(2, 281) = 14.37$, $p < .001$, $\eta_p^2 = .09$ (see Table 1). Post hoc tests indicated that participants were better able to discriminate between lie-tellers and truth-tellers in niqabs and hijabs than those who did not wear veils, $p = .001$, $d = 0.59$, 95% CI [0.29, 0.88] and $p < .001$, $d = 0.96$, 95% CI [0.63, 1.22], respectively. Participants performed similarly when viewing witnesses who were wearing hijabs or niqabs, $p = .232$, $d = -0.26$, 95% CI [−0.54, −0.02]. There was no significant main effect of country, $F(2, 281) = 0.86$, $p = .424$, $\eta_p^2 = .01$, or interaction between the variables, $F(4, 281) = 0.93$, $p = .444$, $\eta_p^2 = .01$.

One-sample t tests, comparing discrimination scores to zero (i.e., no sensitivity), revealed that participants could discriminate between lie- and truth-telling witnesses who wore niqabs, $t(95) = 5.69$, $p < .001$, $d = 0.58$, 95% CI [0.36, 0.80] or hijabs, $t(96) = 8.98$, $p < .001$, $d = 0.91$, 95% CI [0.46, 1.37]. Participants could not discriminate between lie- and truth-tellers who did not wear veils beyond chance levels, however, $t(96) = 0.45$, $p = .652$, $d = 0.46$, 95% CI [0.02, 0.07].

Response bias. According to a Country × Veiling Condition ANOVA, participants' biases (i.e., β) were affected by veiling condition, $F(2, 281) = 5.03$, $p = .007$, $\eta_p^2 = .04$ (see Table 1). Post hoc tests indicated that participants who viewed witnesses in hijabs displayed a different pattern of response bias than those who saw witnesses who did not veil, $p = .005$, $d = -0.44$, 95% CI [−0.72, −0.15]. Participants who viewed witnesses in niqabs did not differ from those who saw witnesses in hijabs, $p = .335$, $d = 0.24$, 95% CI [−0.04, 0.53], or without veils, $p = .198$, $d = -0.22$, 95% CI [−0.50, 0.06]. There was no significant main effect of country, $F(2, 281) = 0.97$, $p = .382$, $\eta_p^2 = .01$, or interaction between the variables, $F(4, 281) = 0.44$, $p = .778$, $\eta_p^2 = .01$.

We compared participants' β scores to one (i.e., no bias) to examine their tendencies to label witnesses as lie-tellers or truth-tellers within each veiling condition. Participants exhibited a truth bias toward witnesses in niqabs, $t(95) = -2.27$, $p = .025$, $d = -0.23$, 95% CI [-[−0.43, −0.02] and hijabs, $t(97) = -5.21$, $p < .001$, $d = -0.53$, 95% CI [−0.79, −0.26]. They did not exhibit response biases when witnesses did not wear veils, $t(96) = 0.29$, $p = .775$, $d = 0.03$, 95% CI [0.01, 0.04].

Discussion

We partially replicated Study 1's primary findings. Participants were more accurate at detecting the deception of witnesses who wore niqabs or hijabs than that of witnesses who did not wear veils. There was no evidence of a negative response bias toward women who veiled in any country. Rather, participants exhibited a tendency to indicate that women who wore niqabs or hijabs were telling the truth.

General Discussion

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. Discrimination between lie- and truth-tellers was no better than guessing in the latter group, replicating previous findings (Bond & DePaulo, 2006). It was only when witnesses wore veils (i.e., hijabs or niqabs) that observers performed above chance levels. Thus, veiling actually improved lie detection (see Table 1).

It is unlikely that these findings were simply false positives. Simmons, Nelson, and Simonsohn (2011) have identified four researcher degrees of freedom that can increase Type I error: disclosing only certain subsets of conditions or dependent variables, employing covariates, and altering the sample size. We did not engage in any of those practices. All conditions and dependent variables were reported, and covariates were not used. The sample sizes differed between the two studies, but the difference was not due to an attempt to manipulate significance. Rather, because this work was the first of its kind, we had no basis upon which to predict effect sizes for use in an a priori power analysis for Study 1. We set a healthy sample size (i.e., 75 participants per veiling condition) and ceased data collection when our target was reached. Due to the nature of our university's participant pool (i.e., testing sessions were posted online at least one week in advance and participants could modify appointments up until the beginning of each session), our final sample size was slightly above what was specified. A post hoc power analysis of the discrimination findings, using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007), revealed that the study was adequately powered (power = .97). By using the effect size from the discrimination findings, we were able to estimate the required sample to produce statistical power at the same level in Study 2 (i.e., $N = 290$); we terminated data collection when it was reached. Thus, there is no reason to believe that "*p*-hacking" was responsible for our significant lie detection results.

Increases in lie detection accuracy associated with veiling might be attributed to the added emphasis on witnesses' eyes. Participants reported that they were no more likely to use the eye region to detect deceit when witnesses wore niqabs than when they did not veil. Eye-tracking data suggest that, when forming social impressions, people spend more time looking at the eyes than any other feature (Janik, Wellends, Goldberg, & Dell'Osso, 1978). People's eyes, and their perceived link to deception, might be so salient that highlighting them with a niqab was superfluous. Indeed, over 90% of the participants in our study reported using eye contact as a cue to deceit whether the witnesses veiled or not. However, self-report should be treated with a degree of caution (e.g., Nisbett & Wilson, 1977). In our study, lie-tellers were more likely to avert their gaze than truth-tellers; veils should have highlighted this difference. Improvements in lie detection performance suggest that participants might have attended to, or interpreted, eye gaze information more accurately in the veiling conditions.

Deception detection strategies were also affected by the amount of visual information that was available. Compared to the other conditions, witnesses in niqabs revealed significantly more verbal than nonverbal cues. Appropriately, participants were more likely to base their decisions on verbal cues than nonverbal cues when viewing witnesses from this group. During several testing sessions, participants did not watch all of the videos (i.e., they turned away from the screens and listened to the testimony). However, this practice only seemed to occur when witnesses wore niqabs. Future research should examine the frequency of self-selected minimization of information (e.g., using eye tracking). Establishing that observers watched the witnesses would then allow researchers to explore the specific mechanisms underlying decision-making (e.g., correlate deception cues with deception judgments using a lens model analysis; see Hartwig & Bond, 2011).

Despite not being explicitly discussed by the courts in the aforementioned cases, we considered whether response bias affected decisions related to veiled witnesses. The decisions of judges and other members of the justice system are typically guided by principles related to fair treatment, such as those laid out in the Equal Treatment Bench Book in the United Kingdom (Judicial College, 2013). The same might not be true of jurors. Tending to (dis)believe a veiled witness due to preexisting stereotypes would severely undermine court proceedings. It was, thus, encouraging that participants were not negatively biased against witnesses who wore niqabs, even in the absence of explicit instruction. These findings replicated previous work, in which mock jurors were similarly unaffected when a witness was described as having worn a burqa (Maeder et al., 2012).

We cannot completely discount the possibility that findings were due to social desirability, however, because participants were not blind to veiling condition. Of course, if participants altered their responses systematically, that could not explain the above-chance discrimination between lie- and truth-tellers in the veiling conditions (i.e., response bias and discrimination are independent; Green & Swets, 1966). Only response biases should have been affected. Yet, were our findings merely a reflection of socially acceptable norms, then we might have expected differences in response biases between the countries; participants in the Netherlands—a country that had considered banning veils (e.g., Government of the Netherlands, 2012)—might have been less positive toward witnesses who wore niqabs, for example. Instead, participants in the Netherlands, Canada and the United Kingdom viewed veiled witnesses similarly. Judges and jurors always know whether a witness is wearing a niqab while testifying and, presumably, would exhibit the same tendencies as participants in our study. Indeed, meta-analyses have failed to find consistent differences in lie detection performance between students, community members, and justice officials (Aamodt & Custer, 2006; Bond & DePaulo, 2006).

An additional limitation of this work is that we randomly assigned our witnesses to lie and/or wear a veil. This practice was important from a scientific standpoint because it helped to ensure initial equivalence between the groups. Experimentally manipulating lying (vs. inducing volitional, naturalistic lies) should not have significantly affected the results and is in keeping with previous research on lie detection (see Vrij, 2008 for a review). Witnesses thought that the study was involving, and they were motivated to be believed: the deception paradigm invoked experimental realism. However, because we randomly assigned witnesses to veiling condition, we might also have obscured natural differences between the groups (Ammar & Leach, 2013). For example, in Ammar and Leach's (2013) study, the women who wore niqabs were less likely to be native English speakers than women who did not veil. Emerging work suggests that laypersons and police officers are not only less able to discriminate between lie- and truth-tellers who are speaking in a non-native language, but also view them less positively than native speakers (Leach & Da Silva, 2013). It is unknown how natural variations in veiled witnesses' language proficiencies would have mitigated our findings. In the future, researchers might wish to examine people's assessments of actual niqab-wearers to address this issue.

The two studies reported here provide unique tests of the behavioral assumptions underlying important courts decisions in the United States, United Kingdom, and Canada. The essence of these decisions is that women must remove their niqabs while testifying to ensure the fairness of court proceedings (e.g., *The Queen v. D(R)*, 2013). Although preliminary, in the sense that we have reported only two empirical studies addressing these assumptions, the data consistently suggested that minimizing visual information actually improved participants' lie detection performance. It is noteworthy that witnesses themselves believed that they would be more accurately judged when wearing niqabs. Thus, seeing a person's entire face does not appear to be necessary for lie detection; 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 scientific evidence. 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.

References

Aamodt, M. G., & Custer, H. (2006). Who can best catch a liar? A meta-analysis of individual differences in detecting deception. *Forensic Examiner, 15,* 6–11.

Akehurst, L., Kohnken, G., Vrij, A., & Bull, R. (1996). Lay persons' and police officers' beliefs regarding deceptive behavior. *Applied*

Cognitive Psychology, 10, 461–471. http://dx.doi.org/10.1002/(SICI)1099-0720(199612)10:6<461::AID-ACP413>3.0.CO;2-2

Albrechtsen, J. S., Meissner, C. A., & Susa, K. J. (2009). Can intuition improve deception detection performance? *Journal of Experimental Social Psychology, 45,* 1052–1055. http://dx.doi.org/10.1016/j.jesp.2009.05.017

al-Ghazali, M. (2008). *Lisat min al-islam* [Not from Islam]. Cairo, Egypt: Dar al-shorouq.

Ambady, N. (2010). The perils of pondering: Intuition and thin slice judgments. *Psychological Inquiry, 21,* 271–278. http://dx.doi.org/10.1080/1047840X.2010.524882

Ambady, N., Bernieri, F. J., & Richeson, J. A. (2000). Toward a histology of social behavior: Judgmental accuracy from thin slices of the behavioral stream. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 32, pp. 201–251). San Diego, CA: Academic Press.

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin, 111,* 256–274. http://dx.doi.org/10.1037/0033-2909.111.2.256

Ammar, N., & Leach, A.-M. (2013, November). *Veiling in the courtroom: Muslim women's credibility and perceptions*. Paper presented at the Qatar Foundation Annual Research Conference, Doha, Qatar.

Baron-Cohen, S., Wheelwright, S., & Jolliffe, T. (1997). Is there a "language of the eyes"? Evidence from normal adults, adults with autism or Asperger syndrome. *Visual Cognition, 4,* 311–331. http://dx.doi.org/10.1080/713756761

Behiery, V. (2013). Bans on Muslim facial veiling in Europe and Canada: A cultural history of vision perspective. *Social Identities: Journal for the Study of Race*. Nature and Culture, *19,* 775–793.

Bill 60: Charter Affirming the Values of State Secularism and Religious Neutrality and of Equality between Women and Men, and Providing a Framework for Accommodation Requests, Quebec National Assembly, 40th Legislature. (2013).

Blais, C., Roy, C., Fiset, D., Arguin, M., & Gosselin, F. (2012). The eyes are not the window to basic emotions. *Neuropsychologia, 50,* 2830–2838. http://dx.doi.org/10.1016/j.neuropsychologia.2012.08.010

Bond, C. F., Jr., & Atoum, A. O. (2000). International deception. *Personality and Social Psychology Bulletin, 26,* 385–395. http://dx.doi.org/10.1177/0146167200265010

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

Bond, C. F., Jr., Omar, A., Mahmoud, A., & Bonser, R. N. (1990). Lie detection across cultures. *Journal of Nonverbal Behavior, 14,* 189–204. http://dx.doi.org/10.1007/BF00996226

Burgoon, J. K., Blair, J. P., & Strom, R. E. (2008). Cognitive biases and nonverbal cue availability in detecting deception. *Human Communication Research, 34,* 572–599. http://dx.doi.org/10.1111/j.1468-2958.2008.00333.x

Centre for Canadian Language Benchmarks. (2010). *Canadian language benchmarks: English as a second language for adults*. Retrieved on August 25th, 2011, from http://www.cic.gc.ca/english/pdf/pub/language-benchmarks.pdf

Clarke, L. (2013). *Women in niqab speak: A study of niqab in Canada*. Toronto, Ontario, Canada: Canadian Council of Muslim Women.

Datavyu Team. (2014). *Datavyu: A video coding tool*. New York, NY: Databrary Project, New York University. Retrieved from http://datavyu.org

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129,* 74–118. http://dx.doi.org/10.1037/0033-2909.129.1.74

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175–191. http://dx.doi.org/10.3758/BF03193146

Fischer, P., Greitemeyer, T., & Kastenmuller, A. (2007). What do we think about Muslims? The validity of Westerners' implicit theories about the associations between Muslims' religiosity, religious identity, aggression potential, and attitudes toward terrorism. *Group Processes & Intergroup Relations, 10,* 373–382. http://dx.doi.org/10.1177/1368430207078697

The Global Deception Research Team. (2006). A world of lies. *Journal of Cross-Cultural Psychology, 37,* 60–74. http://dx.doi.org/10.1177/0022022105282295

Government of the Netherlands. (2012, January 27). *Government approves ban on clothing that covers the face*. Retrieved from http://www.government.nl/news/2012/01/27/government-approves-ban-on-clothing-that-covers-the-face.html

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York, NY: Wiley.

Hartwig, M., & Bond, C. F., Jr. (2011). Why do lie-catchers fail? A lens model meta-analysis of human lie judgments. *Psychological Bulletin, 137,* 643–659. http://dx.doi.org/10.1037/a0023589

Hoodfar, H. (1997). The veil in their minds and on our heads: Veiling practices and Muslim women. In L. Lowe & D. Lloyd (Eds.), *The politics of culture in the shadow of capital* (pp. 248–279). Durham, NC: Duke University Press.

Janik, S. W., Wellens, A. R., Goldberg, M. L., & Dell'Osso, L. F. (1978). Eyes as the center of focus in the visual examination of human faces. *Perceptual and Motor Skills, 47,* 857–858. http://dx.doi.org/10.2466/pms.1978.47.3.857

Jarvis, B. G. (2008). *MediaLab (Version 2008.1.33e)* [Computer software and manual]. New York, NY: Empirisoft Corporation.

Judicial College. (2013). *Equal treatment bench book*. London, England: Judicial College.

Khiabany, G., & Williamson, M. (2008). Veiled bodies—Naked racism: Culture, politics and race in the sun. *Race & Class, 50,* 69–88. http://dx.doi.org/10.1177/0306396808096394

Leach, A. M., & Da Silva, C. S. (2013). Language proficiency and police officers' lie detection performance. *Journal of Police and Criminal Psychology, 28,* 48–53. http://dx.doi.org/10.1007/s11896-012-9109-3

Leal, S., & Vrij, A. (2008). Blinking during and after lying. *Journal of Nonverbal Behavior, 21,* 87–102.

Loi n° 2010–1192 du 11 octobre 2010 interdisant la dissimulation du visage dans l'espace public [Act No 2010-1192 of 11 October 2010 prohibiting the concealing of the face in public], J. O. 0237, October 12, 18344 (2010).

Loi visant à interdire le port de tout vêtement cachant totalement ou de manière principale le visage, Moniteur, July 13 (2011).

Maeder, E. M., Dempsey, J., & Pozzulo, J. (2012). Behind the veil of juror decision-making: Testing the effect of Muslim veils and defendant race in the courtroom. *Criminal Justice and Behavior, 39,* 666–678. http://dx.doi.org/10.1177/0093854812436478

Manson v. Brathwaite, 432 U.S. 98 (1977).

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

Mistry, H., Bhugra, D., Chaleby, K., Khan, F., & Sauer, J. (2009). Veiled communication: Is uncovering necessary for psychiatric assessment? *Transcultural Psychiatry, 46,* 642–650. http://dx.doi.org/10.1177/1363461509351366

Mok, T. (2013, September 12). "We don't care what's on your head": Ontario hospital launches ad aimed at Quebec medical students, values charter. *The National Post*. Retrieved from http://news.nationalpost.com/2013/09/12/we-dont-care-whats-on-your-head-toronto-area-hospital-ad-aims-at-quebec-medical-students-values-charter/

Muhammad v. Enterprise Rent-A-Car, No. 06–41896-GC (31st D. Mich, Oct. 11, 2006).

Murphy, K. R., & Balzer, W. K. (1986). Systematic distortions in memory-based behavior ratings and performance evaluations: Consequences for rating accuracy. *Journal of Applied Psychology, 71,* 39–44. http://dx.doi.org/10.1037/0021-9010.71.1.39

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84,* 231–259. http://dx.doi.org/10.1037/0033-295X.84.3.231

N. S. v. Her Majesty the Queen, et al., No. 33989 (SCC 2012).

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

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22,* 1359–1366. http://dx.doi.org/10.1177/0956797611417632

Street, C. N. H., & Richardson, D. C. (2015, December). The focal account: Indirect lie detection need not access unconscious, implicit knowledge. *Journal of Experimental Psychology: Applied, 21,* 342–355. http://dx.doi.org/10.1037/xap0000058

Syed, Z. (2010, June 21) Banning the niqab. *Islamic Insights*. Retrieved from http://www.islamicinsights.com

Taylor, P., Larner, S., Conchie, S., & Van der Zee, S. (2014). Cross-cultural deception detection. In P. A. Granhag, A. Vrij, & B. Verschuere (Eds.), *Deception detection: Current challenges and cognitive approaches* (pp. 175–202). Chichester, England: Wiley Blackwell.

The Queen v. D(R) [2013] EW Misc 13 (CC).

Troje, N. F. (2002). Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. *Journal of Vision, 2,* 371–387. http://dx.doi.org/10.1167/2.5.2

Vakulenko, A. (2007). "Islamic headscarves" and the European Convention on Human Rights: An intersectional perspective. *Social & Legal Studies, 16,* 183–199. http://dx.doi.org/10.1177/0964663907076527

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities* (2nd ed.). New York, NY: Wiley.

Vrij, A., & Winkel, F. W. (1991). Cultural patterns in Dutch and Surinam nonverbal behavior: An analysis of simulated police/citizen encounters. *Journal of Nonverbal Behavior, 15,* 169–184. http://dx.doi.org/10.1007/BF01672219

Wells, G. L., & Quinlivan, D. S. (2009). Suggestive eyewitness identification procedures and the Supreme Court's reliability test in light of eyewitness science: 30 years later. *Law and Human Behavior, 33,* 1–24. http://dx.doi.org/10.1007/s10979-008-9130-3

Wixted, J., & Lee, K. (2014). *Signal detection theory*. http://dx.doi.org/10.1002/9781118445112.stat06743

Zuckerman, M., Koestner, R., & Colella, M. J. (1985). Learning to detect deception from three communication channels. *Journal of Nonverbal Behavior, 9,* 188–194. http://dx.doi.org/10.1007/BF01000739

Appendix

Nonverbal and Verbal Cues

The nonverbal cues were eye contact, blinking, pupil dilation, smiling, covering mouth and eyes, facial expressiveness, unfriendly facial expressions, shifts in posture, self-manipulations (e.g., self-touching or scratching), leg and foot movements, fidgeting, and use of hand gestures to illustrate speech. Vocal cues included stuttering, grammatical errors, repetitions of words or phrases, voice pitch, vocal tension, rate of speech, speech hesitations, number of pauses, length of pauses, coherence of account, length of answers, amount of detail, inclusion of unusual details, spontaneous corrections, admitting lack of memory, inconsistent information, generalizations, vagueness, complaints, cooperativeness, and overall nervousness.

When coding cues, research assistants counted the frequency of the majority of the nonverbal and verbal behaviors listed above. Cues that were more difficult to quantify in that manner—vocal tension, coherence, vagueness, cooperativeness, nervousness, facial expressiveness, generalizations, rate of speech, and amount of detail—were rated on a scale from 1 to 5. In addition, pupil dilation was not coded by research assistants because it was not sufficiently visible in all videos.

All of the nonverbal and verbal cues were presented to participants as part of the *Cue Use Measure*. Participants indicated that they had used the cue by selecting the box next to the word or phrase.

Received March 16, 2015
Revision received January 12, 2016
Accepted February 22, 2016 ■