Unveiling the Truth: The Effect of Muslim Garments and Face Covering on the Perceived Credibility of a Victim's Court Testimony

Weyam Fahmy and Brent Snook

Memorial University of Newfoundland

Kirk Luther

Lancaster University

Meagan I. McCardle

Memorial University of Newfoundland

Authors Note

Weyam Fahmy, Brent Snook, and Meagan McCardle, Department of Psychology, Memorial University of Newfoundland, St. John's, NL, Canada.

Kirk Luther, Department of Psychology, Lancaster University, Bailrigg, Lancaster, United Kingdom.

Corresponding author: Brent Snook, Department of Psychology, Memorial University of Newfoundland, 230 Elizabeth Ave, St. John's, NL, A1B 3X9, Canada. E-mail: bsnook@mun.ca

Abstract

The perceived credibility of a sexual assault victim's court testimony was examined. A 2 (Face Covered: No, Yes) x 2 (Muslim Garment: No, Yes) between-participant design was used. Participants (N = 120) were assigned to watch one of four videos of a sexual assault victim providing testimony, and asked to rate her credibility. The effect of Muslim Garment on victim credibility ratings was significant; the victim was perceived as more credible when she wore a niqab or hijab compared to when she did not wear either of these garments. The effect of Face Covering on credibility ratings was non-significant, and the interaction was non-significant. The implications for women who wear Muslim garments while testifying about sexual assault are discussed.

Keywords: credibility assessment, attitudes, Muslim veil, facial cues, sexual assault

Public Significance Statement

It has been suggested that women may be perceived negatively by triers of fact if they wear a Muslim garment (i.e., a niqab or hijab) while testifying in court. The current experiment examined this notion by showing people a video of a sexual assault victim's court testimony and asking them to rate her credibility. The results showed that the victim was perceived to be more credible when she was wearing a Muslim garment (i.e., hijab or niqab) than when she was not wearing either garment.

Unveiling the Truth: The Effect of Muslim Garments and Face Covering on the Perceived Credibility of a Victim's Court Testimony

In 2008, a sexual assault complainant was ordered to remove her niqab, a full-face Muslim veil, while testifying in court against two defendants (*R. v. N.S.*, 2009). The defendants' lawyers argued that sections 7 and 11(d) of the Canadian Charter of Rights and Freedoms would be violated if N.S. was permitted to testify with her face covered because it would prevent effective cross-examination and the trier of fact would be unable to assess credibility. Essentially, it was argued that the defendants' right to a fair trial would be compromised by allowing N.S. to remain veiled. The issue of whether N.S. was permitted to wear her niqab while testifying was challenged in the Supreme Court of Canada (SCC). In 2012, the SCC ruled that the requirement to remove the niqab should be made on a case-by-case basis. N.S. subsequently removed her niqab to provide her testimony.

The SCC ruling ignited reactions from various groups. Several civil rights organizations (e.g., The Canadian Council on American-Islamic Relations) objected to the ruling because it was thought that it infringed upon N.S.'s right to freedom of religion. Legal scholars and academics have also raised objections because the ruling is at odds with a robust body of empirical evidence showing that facial expressions are unrelated to deception, that professionals in legal disciplines (e.g., judges, lawyers) are unable to detect deception beyond chance levels, and that it is not possible to improve deception detection with training programs that focus exclusively on non-verbal cues (see Aamodt & Custer, 2006; Bond & DePaulo, 2008; DePaulo, Lindsay, Malone, Muhlenbruck, Charlton, & Cooper, 2003; Hauch, Sporer, Michael, & Meissner, 2014; Sporer & Schwandt, 2007; also see Snook, McCardle, Fahmy, & House, 2017). The aforementioned research suggests, as indicated by Qureshi (2014), that wearing a niqab while testifying would

not compromise trial fairness. However, Qureshi also indicated that there may be unintended consequences of permitting niqab-wearing women to testify; she argued that wearing a niqab may cause victims to be perceived as less credible because it evokes negative stereotypes about Muslims. The goal of the current research is to test that potential unintended consequence by examining how people perceive the credibility of niqab-wearing victims in court.

The Case of *N.S.*

In 2007, a niqab-wearing woman (N.S.) in Ontario accused two men of sexually assaulting her repeatedly from 1982 to 1987, and was called to testify in the preliminary hearing. The defendants' lawyers argued that N.S. should not be allowed to wear the niqab while testifying because it would prevent the court from evaluating the credibility of her statements and prevent them from conducting a proper cross-examination. N.S. objected to this request and cited her right to freedom of religion. The preliminary inquiry judge, after holding a *voir dire* on the matter, concluded that N.S.'s religious beliefs were not sincere because she removed her niqab a few years prior when obtaining a driver's license. Although the judge ordered her to remove her niqab, the preliminary inquiry was adjourned because N.S. disobeyed the court ruling.

N.S. appealed the ruling from the preliminary hearing. Judge Marroco of the Superior Court of Justice agreed that N.S. should be allowed to wear a niqab while testifying, but added the caveat that the preliminary inquiry judge would be allowed to exclude her testimony if wearing the niqab compromised the accused's right to a fair trial. Eventually, the Court of Appeal ruled that a failure to resolve the conflict between the accused's right to trial fairness and N.S.'s freedom of religion rights would result in N.S. having to remove her niqab (*R. v. N.S.*, 2009). The case was then returned to the preliminary inquiry, but N.S. appealed to the SCC.

In 2012, the SCC ruled, based on the majority opinion of Justices McLachlin, Deschamps, Fish, and Cromwell, that the decision of whether a witness would be allowed to testify while wearing a niqab would have to be made on a case-by-case basis by using a specific framework. This framework allows judges to make this decision by answering the following four questions: (1) Would requiring the witness to remove her niqab while testifying interfere with her religious freedom? (2) Would permitting the witness to wear her niqab while testifying create a serious risk to trial fairness? (3) If a conflict between religious freedom and trial fairness exists, is there a way to accommodate both rights and avoid the conflict between them? (4) Do the salutary effects of requiring the witness to remove her niqab outweigh the deleterious effects of doing so? Although the aforementioned framework is well-intended, it essentially a Hobson's choice because of the deeply intertwined nature of trial fairness and the availability of facial evidence. Most, if not all, lawyers would argue that trial fairness hinges on the ability to assess demeanour evidence; part of demeanour evidence is the information gathered from the face/head area (see *R v. Lifchus*, 1997 for arguments on the weight of demeanour evidence in rendering credibility judgements; also see Wellman, 1997 for evidence that lawyers are trained to use demeanour evidence in their examinations). This venerated belief is also evident in the comments made throughout the *R. v. N.S.* (2012) ruling by McLachlin, Deschamps, Fish, and Cromwell. For instance, the Chief Justice wrote that "non-verbal communication can provide the cross-examiner with valuable insights that may uncover uncertainty or deception, and assist in getting at the truth" (para, 24), and that "covering a witness's face may also impede credibility assessment by the trier of fact, be it judge or jury" (para, 25). Furthermore, she stated "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", that seeing the face "...is too deeply rooted in the criminal justice system to be set

aside absent compelling evidence" (para, 27), and that "in the absence of evidence showing that these beliefs, backed by centuries of practice, are unsubstantiated "myths" that should be excised from the law, we should not take such a radical step..." of allowing the face to be covered (para, 49). However, she stated that covering the face would be allowed "where evidence is uncontested" and thus, "credibility assessment and cross-examination are not in issue" (para, 28). That being said, it is highly unlikely that evidence provided by a victim of sexual assault, or any other key witness to a crime, would not be subject to a cross-examination. Such comments by the majority of justices and the aforementioned decision framework highlight the concern that it will be very difficult to convince judges that trial fairness can be achieved without seeing a witnesses' face.

Further, Justices LeBel and Rothstein argued that witnesses should not be allowed to wear the niqab while testifying because the niqab hinders communication and prevents the witness from being able to interact fully with the triers of fact (para, 77). Although Justice Abella held that observing demeanor cues is important in credibility assessment, she argued that N.S. should be allowed to wear her niqab. She reasoned that prohibiting women from testifying while wearing the niqab could discourage these women from participating in the justice system (*R. v. N.S.*, 2012). The case was returned to the preliminary inquiry judge and N.S. removed her niqab to testify.

The Muslim Garment: Religion or Demeanour Evidence?

In a legal analysis of this ruling, Qureshi (2014) argued that covering the face (i.e., hiding demeanor cues) does not harm a defendants' right to a fair trial because there is a wealth of research showing that people, including professionals such as judges, are unable to detect deceit accurately using demeanor cues. However, an equally central, and under-investigated aspect of

Qureshi's argument was that allowing Muslim women to wear a niqab while testifying would not benefit them because the religious garment may trigger negative biases. A closer look at this argument reveals that there are at least two confounding factors that could lead to negative biases. The first factor is that a Muslim garment may evoke negative stereotypes about Muslims. The second factor is that covering the face may result in negative perceptions because people are unable to access facial cues that they believe provide valuable information about an individual's credibility (e.g., emotionality).

Religion. N.S. and other niqab-wearing complainants are arguably likely to be viewed as less credible than non-niqab wearing complainants because judgments of their credibility will be influenced by stereotypes pertaining to their religion (Qureshi, 2014). For instance, Bakht (2009, 2014) argued that niqab-wearing witnesses could be perceived as less credible because of the subconscious biases that jurors may hold against Muslims as out-group members. Niqab-wearing women are especially prone to marginalization because the niqab declares their religion and out-group membership publicly (Chakraborti & Zempi, 2012).

Park, Felix, and Lee (2007) conducted a series of experiments to evaluate both the implicit and explicit attitudes that individuals have of Arab-Muslims. In the study of interest, participants performed an implicit association test (IAT), where they were presented with typical Arab-Muslim and White names, and pleasant and unpleasant words. Participants also completed several explicit attitude measures [i.e., the anti-Muslim prejudice scale, an open-ended question about what they knew or heard about Muslims, and a feeling thermometer (i.e., rate individuals on a 99-point scale, where 1 = *cold and unfavorable* and 99 = *warm and favorable*)]. The IAT revealed a faster association between White + pleasant pairs than Arab-Muslim + pleasant pairs,

and the open-ended question revealed that participants most commonly associated Arab-Muslims with negative attributes (e.g., terrorism, deep religiosity, discrimination against women). Research has demonstrated that negative attitudes toward Muslims have been increasing in Western countries (e.g., Sheridan, 2006). Recent public opinion polls in Canada have also shown that there is an aversion toward the niqab and what individuals believe it represents. For example, a recent report found that 68% of Canadians support Bill-62 which prohibits niqab-wearing women from receiving government services, attending educational institutions, and utilising public transit (IPSOS, 2017). More generally, a 2016 report found that 57% of Canadians hold negative views of Muslims – a number that has increased by 7% since the last report in 2013 (Association for Canadian Studies, 2016). Similar findings have been reported in other Western countries such as the United Kingdom and the United States of America (ComRes, 2016; Telhami, 2016).

However, there is some evidence to suggest that Muslim garments do not evoke negative stereotypes or biased decisions in legal settings. In a study by Maeder, Dempsey, and Pozzulo (2012), participants read a mock-trial transcript of a sexual assault case that varied by victim's race (White or Middle Eastern) and type of Muslim garment worn by the victim (burqa, hijab, or no Muslim garment). University students were then asked to answer a series of questions (e.g., rate the credibility of the victim, render a verdict, rate the defendant's guilt, indicate their level of confidence in the verdict, decide on a sentence length for the defendant). With the exception of confidence, their results showed that neither Muslim garments nor race had any significant effect on the participants' answers to any of the questions asked. With regards to confidence, participants were more confident in the defendant's guilt when the victim was described as wearing a Muslim garment (i.e., burqa or hijab). These findings are counterintuitive to the data

from opinion polls suggesting that Canadians may have negative perceptions of women wearing Muslim garments.

Face Covering. Much insight on how people perceive victims that cover their face can be drawn from research on how people use facial cues to make judgments. Research has shown that we rely on facial cues from an early age; children examine faces to navigate the social world, and associate certain facial expressions with specific emotions, situations, and roles (Crain, 2016). Studies have found that children also use facial cues to infer emotions (e.g., Gnepp, 1983), and that their accuracy in doing so improves as they become older (Camras & Allison, 1985). Adults continue to use facial cues to make a range of judgements about people such as their political orientation (e.g., Snook, Grant, & Button, 2009), personality (e.g., Penton-Voak, Pound, Little, & Perrett, 2006), trustworthiness (e.g., Zebrowitz, Voinescu, & Collins, 1996), and truthfulness (e.g., Akehurst, Köhnken, Vrij, & Bull, 1996).

Of importance to the current study is the use of facial cues in judging credibility. With specific reference to the legal system, extant research suggests that facial cues play a role in witness credibility assessment. For example, research has shown that perceptions of credibility are affected by the extent to which people perceive a witness to be competent, trustworthy, physically attractive, friendly, confident, articulate, and likeable (Brodsky, Griffin, & Cramer, 2010). In addition, the witness's emotional state (e.g., nervousness), gestures, fluency of speech, facial expressions, and eye contact impact perceptions of credibility (e.g., Brodsky et al., 2010; Miller & Burgoon, 1982). As mentioned, it has also been made clear from case law (e.g., *R. v. N.S.*, 2012) that accessing facial cues is a necessary requisite for determining credibility.

The Current Study

In *R. v. N.S.* (2012), N.S. was forced to take off her niqab to testify. Qureshi (2014) suggested that wearing the niqab may affect the way N.S, and other niqab-wearing women, are perceived because of negative stereotypes associated with Muslim garments or because covering the face violates the norm of accessing facial cues to render judgements within the criminal justice system (as well as the norm of using the face to make judgements about others in everyday interactions). The goal of the current study was to evaluate whether credibility judgments of an alleged victim (henceforth referred to as victim) testifying in court would be affected by whether or not (a) she covers her face, and (b) the face covering is religious, specifically Islamic, in nature.

The current study extends existing research in a number of ways. For example, whereas Maeder and colleagues (2012) asked participants to judge a victim's credibility based on a written mock-trial transcript and a description of the victim's manner of dress, we asked participants to make such judgements after viewing a video presentation of a victim's testimony. The video presentation of the testimony aimed to increase ecological validity as it more closely reflects the stimuli that would be presented within court cases. In addition, we added a measure of social desirability to assess, and control for, participants' likelihood of responding in a socially acceptable manner (e.g., by giving Muslim victims high ratings of credibility to hide any anti-Muslim beliefs they may hold), and used the Witness Credibility Scale, which has been demonstrated to be a reliable and valid means of measuring the perceived credibility of a witness (Brodsky et al., 2010). Furthermore, and perhaps most importantly, our study provides a conceptual replication of Maeder and colleagues' research regarding people's attitudes towards victims who wear a Muslim garments while testifying.

Given the documented prevalence of negative Muslim stereotypes, we hypothesized that sexual assault victims who wear a Muslim garment (i.e., niqab or hijab) will be rated as less credible than victims not wearing such garments when providing testimony. As covering the face appears to violate criminal justice and societal norms of interpersonal interactions, we also hypothesized that victims who cover their face will be rated as less credible than victims who do not cover their face. In addition, we predicted an ordinal interaction whereby covering the face with a niqab (Face Covered/Muslim Garment) will result in significantly lower credibility ratings than when covering the face with a balaclava (Face Covered/No Muslim Garment).

Methods

Participants

Participants (*N* = 135) were undergraduate students at Memorial University. The data from 12 participants were excluded because they did not pass the manipulation check (i.e., they did not perceive the niqab or hijab as a religious garment, or they perceived the balaclava as a religious garment). Data from one participant were excluded because s/he did not follow instructions, and data from two participants were excluded because they chose not to have their data included in the analysis (as per IRB policy). A *post-hoc* power analysis was conducted and revealed that with a sample size of 120 participants (i.e., 30 in each condition) the probability of detecting a medium effect size was 78%, which is within the accepted range (Cohen, 1992). Of the remaining 120 participants, 98 (81.67%) were women, and the mean age of participants was 20.62 years (*SD* = 3.67; *Range*: 17 – 39). Of the 117 participants who reported their religion, 61 (52.14%) indicated that they were not religious, 52 (44.44%) indicated they were Christian, three (2.56%) indicated that they practiced a religion other than those listed, and one (0.85%) indicated s/he was Muslim. Six participants (5.00%) indicated that they had been accused of a

crime, 35 (28.33%) indicated they had been the victim of a crime, 12 (10.00%) indicated that they had some experience working with victims of sexual assault, and 19 (15.83%) indicated that they had some experience working with victims of other crimes. There was a significant difference between the four conditions in terms of the distribution of participants who reported being a victim of a crime ($\chi^2 = 9.36$, $df = 3$, $p = 0.025$); relative to the other three conditions, twice as many participants in the balaclava face covered condition (described below) indicated that they were victims of a crime. There were no significant differences between the four conditions with respect to any of the remaining demographic variables or scores on the Marlowe-Crowne Social Desirability Scale (MC-SDS; Crowne & Marlowe, 1960).

Design

A 2 (Face Covered: No, Yes) x 2 (Muslim Garment: No, Yes) between-participant design was used. The victim wore a niqab in the Face-Covered/Muslim Garment condition and wore a balaclava in the Face-Covered/No Muslim Garment condition. The victim wore a hijab in the No Face-Covered/Muslim Garment condition and did not wear any face or hair covering in the No Face-Covered/No Muslim Garment condition (see Figure 1 for a photograph of the victim in each condition). The dependent variable was the mean score from the Witness Credibility Scale (Brodsky et al., 2010).

Materials

The following materials were used in the study: (1) an informed consent form, (2) video testimony for the four conditions, (3) Witness Credibility Scale, (4) demographic questionnaire, and (5) the Marlowe-Crowne Social Desirability Scale.

Videos. Four videos were recorded at the Supreme Court, Trial Division in St. John's, Newfoundland and Labrador, Canada (see Supplemental Material for videos). In each video, a

Caucasian woman was videotaped on the witness stand providing her testimony about a sexual assault she allegedly experienced. The script used in the video was extracted from a transcript of an actual court case where a woman was allegedly sexually assaulted by a friend; all details that could potentially identify the individuals involved were anonymized. The victim and event script remained constant between conditions. The total talking time (no covering: 6:06, balaclava: 5:50, niqab: 5:43, hijab: 5:37), and the average rate of speech (no covering: 208, balaclava: 199, niqab: 222, hijab: 225 words/minute) were not significantly different between conditions. In all four videos, the actress wore a black long-sleeved dress. In the Face Covered/Muslim Garment (i.e., niqab) and Face Covered/No Muslim Garment (i.e., balaclava) conditions, her entire face and hair were covered with black garments leaving only her eyes visible. In the No Face Covering/Muslim Garment (i.e., hijab) condition, her hair was covered with a black garment and her face was visible. In the No Face Covering/No Muslim Garment condition (i.e., control) her hair and face were exposed.

The Witness Credibility Scale. The Witness Credibility Scale was used as a measure of participants' perceived credibility of the victim (Brodsky et al., 2010). The scale consisted of 20 items (e.g., believability, truthfulness, reliability) that were rated on a 10-point scale (*1 = low rating; 10 = highest rating*). This scale has demonstrated reliability ($\alpha = 0.95$) along with concurrent, divergent, and construct validity (Brodsky et al., 2010).

Demographic questionnaire. The demographic form contained questions about the participants' age, gender, and religion. Participants were also asked if they had ever been accused of a crime or had ever been the victim of a crime. They were also asked if they had any experience working with victims of sexual assault, and if they had any experience working with victims of crimes

other than sexual assault. As a manipulation check, participants were asked if the actress in the video they had viewed had her face covered, and if she was wearing a religious garment.

Marlowe-Crowne Social Desirability Scale. Social desirability is the tendency of individuals to respond in ways that are socially favorable or appropriate, rather than responding truthfully (Mick, 1996). The Marlowe-Crowne Social Desirability Scale (M-C SDS) includes 33 statements and respondents must decide whether each statement is true or false as it applies to them (e.g., I like to gossip sometimes; Crowne & Marlowe, 1960). The M-C SDS is one of the most widely used social desirability scales (Vu, Tran, Pham, & Ahmed, 2011) and was initially found to have an internal consistency of $\alpha = 0.88$ (Crowne & Marlowe, 1960), and a subsequent replication reported an internal consistency of 0.72 (Loo & Thorpe, 2000). The M-C SDS has also been found to be reliable across diverse samples and experimental contexts (e.g., Sârbescu, Costea, & Rusu, 2012; Vu et al., 2011).

Procedure

Participants were recruited through the university's participant recruitment system. Upon arrival at the research laboratory, each participant was greeted at the door and directed to one of four testing cubicles. Participants were assigned randomly to one of the four conditions and seated at a testing cubicle which contained a computer, 20-inch monitor, and headphones. Participants were shown the video that corresponded with their assigned condition. After viewing the video, participants were asked to complete the Witness Credibility Scale, the M-C SDS, and the demographic form. The experiment took approximately 30 minutes to complete. Participants were thanked for participating, debriefed about the purpose of the study, and compensated with a bonus mark on their final grade in a psychology course.

Results

The Cronbach's Alpha values for the Witness Credibility Scale in each condition were excellent (M = 0.89; *Range*: 0.87 – 0.92; see Table 1), indicating that all 20 items on the scale were correlated and measured the same construct for each condition (Cronbach, 1951; DeVellis, 2012). Participants' responses on the 20 items were summed to create an overall score that was used for all subsequent analyses. Scores could range from 20 (lowest possible credibility score) to 200 (highest possible credibility score), with a midpoint of 110.

Mean credibility ratings for each of the 20 items on the credibility scale are shown in Table 1. The mean credibility score for the victim in the niqab condition was 153.23 (SD = 19.92), 147.93 (SD = 19.61) for the hijab condition, 141.80 (SD = 23.14) for the balaclava condition, and 140.53 (SD = 19.30) for the control condition.

Correlations were computed between the witness credibility scores and all demographic variables, and between the witness credibility scores and the social desirability scores. Except for a significant positive correlation between the demographic variable "experience working with victims of crimes other than sexual assault" and witness credibility, all correlations were non-significant. Credibility ratings were analyzed using a two-factor analysis of covariance (ANCOVA) with Face Covered and Muslim Garment as the two independent variables, and experience working with victims of other crimes as a covariate. The covariate, experience working with victims of crimes other than sexual assault, was significant $F(1, 115) = 5.14$, $p = .025$. Participants who had experience working with victims of crimes other than sexual assault rated the victim as significantly more credible (M = 150.92, SD = 14.06) than participants who did not have such experience (M = 138.73, SD = 22.00); the size of the effect was medium, d = 0.66.

An ANCOVA showed that the effect of Muslim Garment on perceived credibility ratings was significant, $F(1, 115) = 7.82$, $p = .006$. Participants rated victims who wore a Muslim garment (i.e., niqab, hijab; $M = 150.58$, $SD = 19.78$) as significantly more credible than victims who did not wear a Muslim garment ($M = 141.17$, $SD = 21.14$); the size of the effect was small-to-medium, $d = 0.46$. A follow-up test showed that there was no significant difference between participant's credibility ratings of victims who wore the niqab or hijab, $t(1,58) = 1.04$, $p = .303$, $d = 0.27$. The effect of Face Covering on credibility ratings was non-significant, $F(1, 115) = 0.54$, $p = .466$. The mean credibility rating of the victim when her face was covered was 147.52 ($SD = 22.17$) and was 144.23 ($SD = 19.65$) when her face was exposed ($d = 0.16$). The interaction between Face Covering and Muslim Garment on credibility ratings was non-significant, $F(1, 115) = 0.15$, $p = .698$.

Discussion

We examined the effects of wearing a Muslim garment and covering the face on a victim's perceived credibility while testifying in court. Contrary to our prediction, participants rated victims wearing a Muslim garment as more credible than those who did not wear a Muslim garment. Also contrary to our prediction was the finding that covering the face fully did not have a significant effect on credibility ratings; nor was there a significant interaction between the Muslim Garment and Face Covered main effects. When considering Qureshi's (2014) concerns that wearing the niqab may hinder a victim's perceived credibility, our results suggest that such concerns may not be warranted in legal settings. In fact, our findings raise an interesting question about how trial fairness may be impacted by the greater levels of credibility afforded to victims who wear Muslim garments while testifying.

There are at least three plausible explanations for the finding that sexual assault victims wearing a Muslim garment received higher credibility ratings than compared to those who did not wear a Muslim garment. First, some research has shown that religious individuals are perceived by their peers to be more altruistic, empathetic, and honest than non-religious individuals (Saroglou, Pichon, Trompette, Verschueren, & Dernelle, 2005). It is therefore possible that the overtly religious symbol cued participants to use a 'what is religious is good' heuristic when rendering credibility judgements, and therefore, perceived the Muslim victim as more honest and credible. Second, it may be possible that the Muslim garment dispels the common rape myth that the sexual assault victim was "asking for it". The Muslim garment is a conservative manner of dress and may be indicative of a person's commitment to religious values, and so it may be perceived as representing sexually conservative attitudes that do not condone pre-marital or casual sexual encounters. As a result, participants may have ascribed higher credibility ratings to victims wearing a Muslim garment because they did not believe that a sexually conservative person would have consented to this sexual encounter; thus, rendering any alleged sexual contact non-consensual. Conversely, a third possible explanation is that participants drew upon a stereotypical belief commonly held in Western countries; that Muslim women are oppressed (especially those who wear a niqab or hijab) and therefore are more vulnerable to sexual abuse (Bullock & Jafri, 2000).

In contrast to our second hypothesis, we found that covering the face did not lead to lower credibility ratings. Covering the face runs counter to the norms of the justice system, and thus it was expected that covering the face would be seen as an anomaly that would negatively influence credibility ratings of the victims. Our findings lead to the provisional conclusion that whether or not a sexual assault victim chooses to cover her face while testifying in court does not

seem to have any effect on credibility ratings. It is unclear why similar credibility ratings were obtained from participants who did and did not have access to facial cues. It is possible that participants focused on the content of the testimony, rather than the facial cues, to determine credibility. Assuming that being unable to see a person's face violates social norms, it may be the case that participants developed plausible and acceptable explanations for this norm violation (e.g., wearing the niqab is a religious right, the balaclava may be safeguarding the victim's identity or covering an injury).

The main limitation of the current study is the inability to generalize our findings to Canadian jurors (i.e., those rendering credibility judgments about victims in court), all Muslim women (i.e., those being judged), or all crime types. The use of a relatively homogenous sample (i.e., predominantly female university students) raises the concern that our findings may not generalize to Canadian jurors. However, this concern is somewhat mitigated by findings from a recent meta-analysis which revealed that there was little or no difference between student and community samples when rendering legal decisions (Bornstein et al., 2017). In the current study, we were interested primarily in individual attitudes towards, and perceptions of, Muslim sexual assault victims, and so we asked participants to make independent judgements of credibility. Consequently, our results do not generalize to how a jury deliberation impacts a jury's verdict (see Bornstein, 2017 for a review on jury research methodology). It is important that future research addresses this limitation by conducting a mock trial, allowing participants to deliberate, render verdicts, and make sentencing decisions. Another generalizability issue pertains to the race of the Muslim victim. A Caucasian woman was chosen to portray the victim in this study to prevent racial cues from confounding the results. It may be that a Caucasian victim did not match participants' expectations of what a Muslim woman ought to look like (i.e., she did not look

Middle Eastern/Arab; Bullock & Jafri, 2000). Participants' credibility ratings may have been different had the victim's appearance matched those expectations. Future explorations of this topic should examine the interaction between race and Muslim garments on credibility judgments. Our results are also limited in terms of what can be concluded about victim credibility when testifying about different types of crime. We used sexual assault to match the crime type from *R. v. N.S.* (2012). However, it would be of interest to see whether or not our findings would be replicated for various crime types.

The effect of Muslim garments on the perceived credibility of a victim giving court testimony is an under-studied topic. Beyond addressing the aforementioned generalizability issues, there are other areas that researchers could extend the study of this issue. In particular, one area pertains to the method used to cover the face in the Face Covered/No Muslim Garment (i.e., balaclava) condition. In using the balaclava, we aimed to isolate face covering from religion. Consequently, it is important that future research explore alternative strategies for covering the face in a non-religious manner (e.g., a medical mask). It would also be important to study the effect of evoking other aspects of the victim's out-group membership (e.g., foreign accent) on credibility ratings. Future research should also examine the effect of other overt religious symbols on judgements of victim credibility (e.g., wearing a cross).

The primary goal of the present research was to examine the effect of wearing a niqab on the perceived credibility of a sexual assault victim's testimony. Contrary to Qureshi's (2014) concern and our hypothesis, we did not find that wearing the niqab resulted in lower ratings of credibility. Surprisingly, we found that wearing a Muslim garment (i.e., niqab or hijab) resulted in higher ratings of credibility. This counterintuitive finding raises many important questions about the extent to which Muslim garments evoke positive biases, and the extent to which these

laboratory findings translate to actual biases in behaviours and decisions by triers of fact. It is important to note that research in this area is in its infancy. Any decisions on policies or recommendations on the presence of the Muslim garments in court must be undergirded by a robust body of empirical data. Until such a substantive body of research emerges on this issue, the consequences of wearing a Muslim garment while testifying remain opaque.

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., Köhnken, G., Vrij, A., & Bull, R. (1996). Lay persons' and police officers' beliefs regarding deceptive behaviour. *Applied Cognitive Psychology*, *10*, 461-471.

Association for Canadian Studies (2016). State of relations, March 2016: Relations, Contact and perceptions of religious, aboriginal and racial groups in Canada. Retrieved from: https://acs-aec.ca/en/social-research/racism-discrimination-human-rights/

Bakht, N. (2009). Objection your honour! Accommodating niqab wearing women in court. In R. Grillo, R. Ballard, A. Ferrari, A. J. Hoekema, M. Maussen, & P. Shah (Eds.), *Legal practice and cultural diversity* (pp. 115-133). Oxford, UK: Routledge.

Bakht, N. (2014). In your face: Piercing the veil of ignorance about niqab-wearing women. *Social and Legal Studies*, *24*, 419-441. doi: 10.1177/0964663914552214

Bond, C. F., & Depaulo, B. M. (2008). Individual differences in judging deception: accuracy and bias. *Psychological Bulletin*, *134*, 477–492. doi: 10.1037/0033-2909.134.4.477

Bornstein, B. H. (2017). Jury simulation research: Pros, cons, trends and alternatives. In M. B. Kovera (Ed.), *The psychology of juries* (pp. 207-226). Washington, DC: American Psychological Association

Bornstein, B. H., Golding, J. M., Neuschatz, J., Kimbrough, C., Reed, K., Magyarics, C., & Luecht, K. (2017). Mock juror sampling issues in jury simulation research: A meta-analysis. *Law and Human Behavior, 41*, 13-28. doi: 10.1037/lhb0000223

Brodsky, S. L., Griffin, M. P., & Cramer, R. J. (2010). The Witness Credibility Scale. *Behavioral Sciences and the Law*, *28*, 211–223. doi:10.1002/bsl.917

Bullock, K. H., & Jafri, G. J. (2000). Media (mis)representations: Muslim women in the

Canadian nation. *Canadian Woman Studies/Les Cahiers De La Femme, 20*, 35-40.

Camras, L. A., & Allison, K. (1985). Children's understanding of emotional facial expressions and verbal labels. *Journal of Nonverbal Behavior*, *9*, 84–94. doi:10.1007/BF00987140

Chakraborti, N., & Zempi, I. (2012). The veil under attack: Gendered dimensions of Islamophobic victimization. *International Review of Victimology*, *18*, 269–284. doi:10.1177/0269758012446983

Cohen, J. (1992). Statistical power analysis. *Current directions in psychological science*, *1*, 98-101. doi: 10.1111/1467-8721.ep10768783

ComRes (2016). Ahmadi Muslims – Perceptions of the Calphite. Retrieved from http://www.comresglobal.com/polls/ahmadiyya-perceptions-of-the-caliphate-polling/

Crain, W. (2016). *Theories of development: Concepts and applications*. New York, NY: Routledge.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297–334. doi:10.1007/bf02310555

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*, 349-354. doi:10.1037/h0047358.

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

DeVellis, R.F. (2012). *Scale Development: Theory and Applications* (third edition). Sage: Thousand Oaks, CA.

Gnepp, J. (1983). Children's social sensitivity: Inferring emotions from conflicting cues.

Developmental Psychology, *19*, 805–814. doi:10.1037//0012-1649.19.6.805

Hauch, V., Sporer, S. L., Michael, S. W., & Meissner, C. A. (2016). Does training improve the detection of deception? A meta-analysis. *Communication Research*, *43*, 283- 343. doi: 10.1177/0093650214534974

IPSOS (2017). Majority (68%) of Canadians support Bill 62-style law in their province. Retrieved from: https://www.ipsos.com/en-ca/news-polls/quebec-bill-62

Loo, R., & Thorpe, K. (2000). Confirmatory factor analyses of the full and short versions of the Marlowe-Crowne Social Desirability Scale. *The Journal of Social Psychology, 140*, 628–635. doi:10.1080/00224540009600503

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

Mick, D. G. (1996). Are studies of dark side variables confounded by socially desirable responding? The case of materialism. *The Journal of Consumer Research, 43*, 106- 119. doi:10.1086/209470

Miller, G. R., & Burgoon, J. K. (1982). Factors affecting witness credibility. In N. L. Kerr, & R. M. Bray (Eds.), *The psychology of the courtroom* (pp. 169–194). New York: Academic Press.

Park, J., Felix, K., & Lee, G. (2007). Implicit attitudes toward Arab-Muslims and the moderating effects of social information. *Basic and Applied Social Psychology, 29*, 35–45. doi:10.1080/01973530701330942

Penton- Voak, I. S., Pound, N., Little, A. C., & Perrett, D. I. (2006). Personality judgements from natural and composite facial images: more evidence for a "kernel of truth" in social

perception. *Personality and Individual Differences*, *41*, 1107–1118. doi:10.1016/j.paid.2006.04.015

Qureshi, A. M. (2014). Relying on demeanour evidence to assess credibility during trial: A critical examination. *Criminal Law Quarterly*, *61*, 269-301. doi:10.2139/ssrn.2384966

R. v. Lifchus (1997). 3 SCR 320, SCC 319

R. v. N.S. (2009). 21203 (ON SC)

R. v. N.S. (2012). 3 SCR 726, SCC 72

Sârbescu, P., Costea, I., & Rusu, S. (2012). Psychometric properties of the Marlowe-Crowne Social Desirability Scale in a Romanian sample. *Procedia - Social and Behavioral Sciences*, *33*, 707–711. doi:10.1016/j.sbspro.2012.01.213

Saroglou, V., Pichon, I., Trompette, L., Verschueren, M., & Dernelle, R. (2005). Prosocial behavior and religion: New evidence based on projective measures and peer ratings. *Journal for the Scientific Study of Religion, 44*, 323-348. doi: 10.1111/j.1468-5906.2005.00289.x

Sheridan, L. P. (2006). Islamophobia pre- and post-September 11th, 2001. *Journal of Interpersonal Violence*, *21*, 317–336. doi:10.1177/0886260505282885

Snook, B., Grant, M. J., & Button, C. M. (2009). A judgement analysis of social perceptions of attitudes and ability. *Thinking and Reasoning*, *15*, 319–336. doi:10.1080/13546780903135722

Snook, B., McCardle, M. I., Fahmy, W., & House, J. C. (2017). Assessing truthfulness on the witness stand: Eradicating deeply rooted pseudoscientific beliefs about credibility assessment by triers of fact. *Canadian Criminal Law Review, 22*, 305–313.

Sporer, S. L., & Schwandt, B. (2007). Moderators of nonverbal indicators of deception - A meta-

analytic synthesis. *Psychology Public Policy and Law*, *13*, 1–34. doi: 10.1037/1076-8971.1.13.1.1

Telhami, S. (2016). American attitudes toward Muslims and Islam. Retrieved from https://www.brookings.edu/research/american-attitudes-toward-muslims-and-islam/

Vu, A., Tran, N., Pham, K., & Ahmed, S. (2011). Reliability of the Marlowe-Crowne social desirability scale in Ethiopia, Kenya, Mozambique, and Uganda. *BMC Medical Research Methodology*, *11*, 162. doi:10.1186/1471-2288-11-162

Wellman, F. (1997). *The art of cross-examination* (4th Ed.). New York, NY: Touchstone.

Zebrowitz, L. A., Voinescu, L., & Collins, M. A. (1996). "Wide-eyed" and "crooked-face": Determinants of perceived and real honesty across the life span. *Personality and Social Psychology Bulletin, 22*, 1258-1269. doi:10.1177/0146167296221200

Photograph 1. Face Covered/Muslim Garment Condition (i.e., niqab)

Photograph 2. No Face Covered/ Muslim Garment Condition (i.e., hijab)

Photograph 3. Face Covered/ No Muslim Garment Condition (i.e., balaclava)

Photograph 4. No Face Covered/ No Muslim Garment Condition (i.e., control)

Figure 1. Photographs of the actress who played the role of the victim in all conditions

Table 1. *Mean (and Standard Deviation) Credibility Ratings for Each Item on the Witness Credibility Scale*

	Experimental Condition			
Item	Face Covered/ Muslim Garment (Niqab)	No Face Covered/ Muslim Garment (Hijab)	Face Covered/ No Muslim Garment (Balaclava)	No Face Covered/ No Muslim Garment (Control)
1. Unfriendly/Friendly	7.83 (1.58)	8.00 (1.53)	7.70 (1.66)	7.13 (1.89)
2. Disrespectful/Respectful	8.67 (1.79)	8.73 (1.23)	8.50 (1.53)	8.20 (2.19)
3. Unkind/Kind	8.23 (1.79)	8.20 (1.37)	7.73 (2.00)	7.90 (1.24)
4. Ill-mannered/Well-mannered	9.13 (1.14)	9.00 (1.02)	8.33 (1.90)	8.40 (1.65)
5. Unpleasant/Pleasant	7.60 (1.75)	7.60 (1.65)	7.33 (2.17)	7.13 (1.59)
6. Untrustworthy/Trustworthy	8.43 (0.97)	7.60 (1.59)	7.53 (1.93)	7.80 (1.73)
7. Untruthful/Truthful	8.60 (1.10)	7.63 (1.89)	7.50 (2.10)	8.03 (1.65)
8. Undependable/Dependable	8.37 (1.16)	7.63 (1.63)	7.37 (1.90)	7.57 (1.70)

9. Dishonest/Honest	8.60 (1.13)	7.57 (1.99)	7.47 (1.83)	8.00 (1.58)
10. Unreliable/Reliable	8.30 (1.15)	7.73 (1.95)	7.50 (2.00)	7.63 (1.65)
11. Not confident/Confident	7.50 (2.24)	6.93 (1.84)	6.60 (2.44)	6.50 (2.40)
12. Inarticulate/Well-spoken	8.30 (1.66)	8.10 (1.52)	7.63 (2.24)	7.87 (1.70)
13. Tense/Relaxed	4.23 (1.83)	4.57 (2.13)	4.70 (2.39)	3.70 (1.70)
14. Shaken/Poised	4.87 (2.05)	5.07 (2.30)	5.00 (2.49)	3.50 (2.06)
15. Not self-assured/Self-assured	6.43 (1.85)	6.50 (1.83)	6.73 (1.91)	5.73 (2.23)
16. Uninformed/Informed	7.57 (1.52)	7.77 (1.57)	7.00 (1.80)	7.13 (1.57)
17. Illogical/Logical	8.10 (1.60)	7.63 (1.61)	7.17 (1.84)	7.60 (1.77)
18. Uneducated/Educated	8.47 (1.31)	8.33 (1.45)	7.83 (1.53)	8.10 (1.35)
19. Unwise/Wise	7.67 (1.73)	7.20 (1.67)	6.50 (2.03)	6.60 (1.45)
20. Unscientific/Scientific	6.33 (2.01)	6.13 (1.93)	5.40 (1.73)	5.87 (1.81)
Cronbach's Alpha	0.92	0.89	0.89	0.87

Note. Each item on the Witness Credibility Scale ranged from 1 (the person does not possess this trait at all) to 10 (the person definitely possesses this trait).