

THE EFFECTS OF RELIGION AND STEREOTYPE CONTENT ON VERDICTS AND SENTENCE SEVERITY WHEN DEFENDING TERROR CHARGES.

Daniel Frings, Kirsty Rice, and Ian P. Albery.

School of Applied Sciences, Division of Psychology, London South Bank University

Little evidence exists to test if a defendant's religion affects their verdict outcome or sentencing. The current study addresses this question and also tests the role of stereotype content as an explanatory variable. Participants ($n=141$) were presented with crimes which were either stereotypical of Muslims or not. Participants viewed details of a case resulting in either a terror or a theft charge, with a Muslim, Christian or Atheist suspect. Both being a Muslim and defending terror crimes led to more frequent guilty verdicts and more severe sentences. Muslims were perceived as more cold and competent. The colder and more competent suspects were perceived, the more likely they were to be found guilty and the more severe the sentence. Warm/cold evaluations mediated the effect of religion. These findings suggest that Muslim terror defendants may be affected by systematic bias in trials, and that this may be driven by the stereotypes content.

Keywords: Jury; religion; terrorism; stereotypes; Muslim

Race and religion bias in juries

A basic principle of many justice systems is that all individuals are treated equally in the eyes of the law. Under such systems, decisions of guilt and innocence should be based on an evaluation of the facts of the case rather than race, gender, religion or other demographics. However, experimental research shows that jurors find members of minority ethnic defendants guilty more often than non-minority defendants, even when presented with an identical case (e.g., Abshire & Bornstein, 2003). This pattern has been observed amongst group such as Mexican-Americans (Esqueda, Espinoza & Culhane, 2008) and African-Americans (Sommers & Ellsworth, 2000). Ethnic identity has also been linked to the severity of sentences handed down once a guilty verdict has been reached (e.g., Demuth & Steffensmeier, 2004). In a recent experimental study, for example, the death sentence

Author Notes: Conflict of Interest: Frings declares he has no conflict of interest. Rice declares she has no conflict of interest. Albery declares he has on conflict of interest.

All correspondence should be addressed to the first author care of: Division of Psychology, London South Bank University, London, SE1 0AA. Phone: +442078155888, email: fringsd@lsbu.ac.uk

Funding: This study received no specific funding.

Ethical approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

was deployed by white jurors more frequently when defendants were black (Lynch & Haney, 2009). Such observations have been supported by meta-analysis in which a small but significant effect of race on judgments of guilt and sentencing severity was found, with jurors being more likely to deliver guilty verdicts and give longer sentences to other race defendants (Mitchell, Haw, Pfeifer & Meissner, 2005).

Although the effects of ethnicity on trial outcomes have been well studied, relatively little research has addressed another key demographic – religion. The current study addresses this by looking at the effects of being identified as being religious on verdict and sentencing decisions. Whilst a number of studies have discussed anti-Islamic prejudice in the UK (e.g., Dwyer, Shah & Sanghera, 2008; Mythen, Walklate & Khan, 2009; Pantazis & Pemberton, 2009), how jury decision making may be influenced by religious focused stereotyping has received much less attention (see Sheridan, 2006). Evidence suggests that not only are Europeans more prejudiced against Muslims than against any other religious group, but that this prejudice was evident even before the terrorist attacks of September 11th 2001 (Strabac & Listhaug, 2008). In addition, this level of anti-Muslim prejudice does not increase with the size of the Muslim population - suggesting that the prejudice is pre-determined and is not set solely by a specific community's presence or size. In a sample of 222 British Muslims, Sheridan (2006) showed that when questioned about discrimination following September 11, 2001 (in which a UK terrorist attack perpetrated by Muslim terrorists took place), participants reported an increase in both direct (76.3%) and indirect (82.6%) discrimination. This suggests that world events affect both prejudice and stereotypes towards the minority groups involved and that religious group membership may, in some contexts, be a stronger predictor of prejudice than race or ethnicity. What is not clear, however, is the extent to which jury decision-making is affected by a members of the trial's apparent or actual religion. Amongst Muslim jurors, those identified as more 'devoted' to their religion were more likely to employ not guilty verdicts in mock trials compared with jurors identified as more 'fundamental' (Miller, Maskaly, Peoples & Sigillo, 2014). Amongst victims, some evidence suggests seeing a female Muslim victim testifying whilst wearing a niqab (veil) led to more frequent guilty verdicts than observing unveiled victims testifying (Maeder, Dempsey & Pozzulo, 2012). Existing work on the effects of defendants' religion presents mixed results. Johnson (1985) showed that when a Christian defendant included their religious affiliation as part of the defense, more frequent guilty verdicts and longer sentences resulted. In contrast, other work showed that highlighting conversion to Christianity (but not being a lifelong Christian) led to less punitive sentences when judged by a sample of mostly Christian mock jurors. Similarly, Kerr, Hymes, Anderson and Weathers, (1995) observed that perceiving similarity in religious affiliation between mock jurors and defendants led to leniency. Miller, Maskaly, Green and Peoples (2011) observed a similar effect on a generic similarity measures between jurors participants and hypothetical defendants. The current study aimed to add to this body of research by manipulating the religion of a suspect in a case presented to participants who then subsequently made an individual decision on the suspect's guilt or innocence as well a judgment on sentencing. It also aimed to extend the extant literature by looking at the role of stereotypes as an underlying mechanism.

The role of stereotypes

The activation of accessible and relevant stereotypes may be one mechanism through which bias against Muslim defendants may operate. Fiske, Cuddy, Glick and Xu (2002) proposed in their stereotype content model (SCM) that all stereotypes are composed of two dimensions, *warmth* and *competence*, and different combinations of these dimensions form the basis of all available stereotypes. The identified combinations are low competence, high warmth (*pity*); high competence, low warmth (*envy*); high warmth, high competence (*admiration*); and low warmth, low competence (*contempt*) (Fiske et al., 2001). For example, one group offered as high competence, low warmth group were Asians, whereas elderly people were described at high warmth but of low competence. Fiske et al., (2002) also argued that positive stereotypes on one dimension are often matched by negative stereotypes on another, such that the stereotype remains open to prejudice from outgroup members. Fiske et al., (2002) suggest that measures of competence for any individual or group include competence, confidence, independence, competitiveness and intelligence. High ratings of these measures predict high competence, whilst low ratings on these measures predict low competence. Aspects of warmth consist of tolerance, warmth, good nature and sincerity. To determine the usefulness of these measures to describe outgroups, 73 participants were asked to rate how 23 groups of people were viewed by American society. Consistent with in-group favourability bias, members were most likely to be viewed as warm and highly competent.

One important prediction made by the SCM is that, alongside emotions and cognitions, behaviors aimed at targets will vary according to the content of the stereotype. Warm/low competence targets should be the subject of feelings of pity and active help; cold/low competence targets should elicit passive behaviors; and warm/high competence targets elicit facilitative / helping behaviors. Of most relevance to this study is that members of groups rated as cold and competent are predicted to be the target of harmful behaviors. In the context of the current study, this could mean that members of groups that are viewed as cold/competent may well receive more frequent guilty verdicts and harsher sentences than other groups.

The content of Muslim stereotypes appears to be generally negative in Western societies. Such prejudice has been linked to perceptions of fundamentalism (Laythe, Finkel & Kirkpatrick, 2001; Rowatt, Franklin & Cotton, 2005). Dwyer, Shah and Sanghera (2008) suggest recent media coverage and policy decisions often depict young British Muslims as deviant potential terrorists. These attitudes translate into evaluations and behavior. Since the terrorist attacks on September 11th 2001, there has been not only an increase in discrimination against Muslims (according to a UK survey, Sheridan, 2006), but also, since the London attacks on July 7th 2005, an increase in young Muslims subjected to observations, supervision and scrutiny by police and other security agencies (Mythen, et al., 2009). In addition, it seems that young Muslims are being defined as a 'risky, suspect population' as new counter-terrorism legislation is often introduced after terrorist attacks carried out by Muslims (Mythen, et al 2009; see also Pantazis & Pemberton, 2009). Hence, it is likely stereotypes for Muslims can include notion that they are both cold and competence, and more likely to engage in terror-related activities.

The perception of such stereotypes, and their effects on behavior, are likely to be particularly strong when Muslims are perceived to be involved in *stereotype-consistent* behaviors (e.g. terrorism) relative to when they are perceived to be involved in behaviors which are less stereotype consistent. Research in a wide range of domains demonstrates that we tend to accept (and even seek) stereotype confirming messages and information more than disconfirming ones (*confirmation bias*, see Nickerson, 1998) and can recall stereotype consistent information more easily than inconsistent information when under high levels of task demand (Dijksterhuis & Van Knippenberg, 1995). More directly, Jones and Kaplan (2003) showed that crimes which were consistent with a defendant's racial stereotype (car theft amongst African Americans and embezzlement amongst white defendants) received more guilty verdicts than stereotype inconsistent ones. In a similar vein, Khan and Ecklund (2012) found that attitudes towards Muslims American in the United States were more negative compared with individuals with unspecified backgrounds. Importantly, this effect was particularly prominent when boarding a plane – an effect which may be attributed to the stereotyped association of Muslims on planes and terrorist activities. Such effects are not always observed however: Miller et al., conducted two studies – one using a crime stereotypical of Muslims (bombing a train station) and another associated with Christianity (bombing an abortion clinic). In the Muslim stereotype crime, Muslims received the most pre-post deliberation leniency. However, contrary to expectations, the same leniency effect (for Muslims) was observed when the crime was associated with Christianity.

In summary, stereotypes towards Muslims are likely to include them being cold and competent. As this combination of traits is linked with hostile behaviors, the greater extent a defendant is viewed as being cold and competent, the more likely they are to be found guilty of a crime, and the harsher sentence they will receive. As stereotypes for Muslims include terror related crime (i.e. such crimes are stereotype consistent), such effects may be magnified when they are tried for such crimes. Given this, a number of predictions can be made. It was predicted that Muslims will be seen (by non-Muslims) as more cold and competent than other groups, and in line with the SCM, should be the subject of more hostile actions (i.e. more frequent guilty verdicts and more severe sentences). In addition, such stereotypes are more likely to be activated when behaviors are stereotype consistent (e.g. terror crimes). Thus, it was also predicted that Muslims would appear colder and more competent when terror crimes are salient, and that this increase in coldness and competence amongst suspects would be larger for Muslims than for non-Muslims. As SCM argues that both coldness and competence predict hostile behaviors, it is also predicted that ratings of suspects on such dimensions will be related to the frequency of guilty verdicts and severity of sentences and these mechanisms may mediate the effect of a suspect's religion.

METHOD

Participants

One-hundred and forty-four participants were recruited via a social networking platform to take part in an online survey. Three participants failed to finish the survey and their data were excluded. Of the 141 participants who completed the survey, 5 participants

failed a manipulation check to determine if they had correctly identified the religion of the suspects in the survey and were also excluded. Finally, two participants were excluded as they reported being Muslims which may have confounded ingroup/outgroup evaluations. Of the remaining 134 participants - 44 self-identified as Christian, 4 as Hindu, 4 as Sikh, 29 as Atheist, 48 as no religion and 5 as 'other' religion. As Muslims are outgroups for all these participants, participants' religion was not considered further in the analysis. Forty seven participants were males and 87 female. Participants were aged between 18 and 67 years ($M = 36.37$, $SD = 11.60$). Participants received no compensation for participating.

Design

A 2 x 3 between participants design was used. The independent variables were Crime type (Terror vs. Theft) and the suspect's Religion (Muslim vs. Christian vs. Atheist). The Atheist condition was employed to rule out a general anti-religion bias (although the sample/condition distribution precluded an analysis of Atheists vs. Other participants). The dependent variables comprised the verdict given by the participant, the length of the sentence suggested by the participant and the ratings of warmth and competence attributed to the suspect and group by the participant.

Measures

Verdict. Participants indicated whether they thought the participants were guilty (coded as 0) or not guilty (coded as 1). Proportions were calculated within condition such that higher proportions indicate lower proportions of guilty verdicts (or, higher rates of not guilty verdicts).

Sentence severity. Participants were asked 'If the suspect were to be found guilty, which sentence do you think would be the most appropriate?' Participants selected a sentence from a list of four options as follows; a two year suspended sentence (scored 1), up to five years in prison (scored 2), five to ten years in prison (scored 3), and a 10+ years in prison (scored as 4).

Warmth and competence. For the measure of warmth and competence as a group, participants rated the group for levels of warmth (sincerity, good nature, tolerance) and measures of competence (intelligence, competitiveness, confidence). These items were couched thusly; 'On a scale from 1-5 (where 1 = *not very xxx* and 5 = *very xxx*) how [dimension] do you think the [group] is?'. For ratings of individual suspects, [group] was replaced with the word 'suspect'. To confirm these six traits loaded differentially onto relevant factors, factor analysis was undertaken, using principal component analysis and a varimax rotation. One was undertaken for the evaluation of groups and a second for suspects. For groups, two factors emerged which combined predicted 67.3% of the scale variance. As can be seen in the rotated factor loadings in Table 1, the first factor (Eigenvalue = 2.59, 43.20% of variance) comprised the competence traits (with reasonable loadings on all expected items) whilst a second (Eigenvalue = 1.45, 24.13% of variance) representing warmth traits (with high loadings). The next highest factor Eigenvalue was .72, so further factors were not interpreted. A similar pattern was observed at the suspect level, with the model predicting 67.36% of the variance, the first (competence) factor having an

Eigenvalue of 2.53, (42.08% variance) and reasonable loadings whilst the second (warmth) had an Eigenvalue of 1.52 (25.27% variance) and good loadings. The next highest factor Eigenvalue was .65.

Table 1: *Factor loadings of various traits (at suspect and group levels of evaluation)*

Traits	Level of evaluation			
	Suspect		Group	
	Warmth	Competence	Warmth	Competence
Intelligent	.181	.78*	.31	.67*
Competitive	-.48	.59*	-.37	.63*
Confident	-.08	.78*	-.035	.77*
Sincere	.84*	-.02	.90*	<.01
Good natured	.89*	.05	.90*	.06
Tolerant	.85*	-.09	.86*	-.08

Note: Principal component analysis with varimax rotation. Within column figures marked with an asterix were included in a single scale in the final analysis.

Procedure

Participants were invited to take part in the study via a link for an online survey posted on a social networking site (www.facebook.com), using a snowball recruitment methodology. The link specified participants must be aged between 18 and 70 years old. Upon clicking the link, participants were taken to an online survey hosted via Qualtrics. After participants had consented to take part in the study they were randomly assigned to one of the six experimental conditions (i.e. terror crime/Muslim religion; terror crime/Christian religion; terror crime, Atheist; theft crime/Muslim religion; theft crime/Christian religion; theft crime, Atheist). Participants were first presented with the relevant version of a report describing a crime scenario that had supposedly taken place (manipulations for Terror/Theft conditions are presented in brackets); *'On Tuesday 12th November 2013, police were called to investigate a report of suspicious activity near an abandoned warehouse in Birmingham. A shipping container was found and was subsequently investigated by the [counter terrorism/criminal investigation] department. After a thorough investigation lasting six weeks, an arrest was later made, relating to the contents of the container. The suspect was reportedly seen by witnesses as regularly walking to the location late at night and was described as male, 5ft 10 and aged 30-40 years old. The suspect claims he was unaware of the contents in the container and had been asked to check the lock on the container by an acquaintance. The suspect pleads not guilty to the charge of [terrorism/theft]'*

The account for each crime differed only by naming either the criminal investigation department [Theft conditions] or the counter terrorism department [Terror conditions] as the investigating officers and a sentence stating the crime the suspect had been arrested for. The passage also included a general description of the perpetrator which could describe any of the suspects in the study. On the page following this information, participants were

presented with a photograph of the suspect (in the Muslim condition an Arabic male was wearing a taqiyah (skullcap), the Christian and Atheist conditions comprised Caucasian males). All targets had beards. Photos were drawn from the Centre of Longevity Face Database (Minear & Park, 2004). The image was accompanied by a statement of oath/affirmation (used in British jurisdictions) that included information of the suspect's religion. After viewing these, participants were then asked to make a judgment based on the case information they had just been presented with of guilty or not guilty. The survey then prompted all participants to decide what sentence they would deem appropriate, should the suspect be found guilty, by selecting one from the four options (see *Methods*, above). Participants were then asked if they had taken note of the suspect's religion during the survey and asked to state which religion they believed the suspect belonged to. Participants were then asked to rate individual traits of their suspect in terms of warmth and competence. After rating the suspect, participants were then asked to rate the suspect's religious group on the same traits. Finally, participants were then asked to specify their own age, religion and gender and debriefed fully.

RESULTS

A combination of logistic regression and two-way between participants ANOVAs were used to examine the effects of crime type and suspect religion on verdicts and suspect sentencing respectively.

Verdict. Logistic regression was undertaken to test the effects of Crime Type and Religion on guilty verdicts. Crime type (with terrorism as the referent category) and suspect Religion (with Muslim as the referent category) and the interaction term between the two were entered as predictors. Upper and lower 95% percent confidence intervals are reported for odds ratios. For ease of interpretation, not guilty verdicts are expressed as proportions within condition are shown in Table 2. The overall model explained 22% of the variance in verdicts (Nagelkerke $R^2 = .22$) and was statistically significant $\chi^2(5) = 20.50$, $p = .001$. It accurately predicted 80.6% of cases. An examination of the odds ratios revealed that terror defendant were 2.75 times as likely to be found guilty than non-terror crime defendants (although this was marginally significant, (CIs = 0.77-9.86, Wald (1) = 2.41, $p = .12$). Religion predicted verdict (Wald(2) = 10.59, $p = .005$). Muslims were 4.40 times as likely to be found guilty than Christian targets (CIs = 1.19 – 16.17, Wald = 4.99, $p = .01$) and 24.20 times as likely to be found guilty as Atheists (CIs = 2.74-213.94, Wald = 8.21, $p = .001$). The interaction term did not approach significance (Wald(2) = 1.37, $p = .50$).

Table 2: Mean proportion of not guilty verdicts & mean sentence severity by condition. Standard deviations in parentheses

Religion	Crime	Proportion found not guilty	Sentence severity
Muslim	Terrorism	0.48 (.51)	2.95 (1.40)
	Theft	0.71 (.46)	1.62 (0.59)
Christian	Terrorism	0.80 (.41)	2.16 (0.99)
	Theft	0.89 (.32)	1.21 (0.42)
Atheist	Terrorism	0.96 (.21)	2.00 (1.13)
	Theft	0.92 (.28)	1.36 (0.81)

To test the planned comparison that Muslim terror targets would be found guilty more frequently than all other targets, a variable was dummy coded 1 for Muslim terror defendant participants and 0 for all other targets. Logistic regression with this predictor on verdicts revealed the comparison predicted 80.6% of cases correctly in a significant model ($\chi^2(1) = 13.42, p < .001$, Nagelkerke $R^2 = .15$). Muslim terror defendants were 6.67 times as likely to be found guilty than the other conditions (CIs = 2.44-18.25, Wald(1) = 13.65, $p = .001$). Overall this analysis presents evidence that both being Muslim and being charged with a terror crime lead to higher levels of guilty verdicts, and these effects concatenate such that Muslim terror charge defendants are most likely to be found guilty. A second analysis was undertaken including variables dummy coded for the participant being a member of each religion or not in the first step (to control for participants' religion). None reached significance, $ps > .13$.

Sentence severity. A two-way ANOVA on sentence severity, with Religion and Crime Type as independent variables, revealed a significant main effect of Crime Type, $F(1,128) = 34.69, p < .001, \eta_p^2 = .21$. As would be expected, sentences given for terror crimes ($M = 2.35, SD = 1.22$) were more severe than sentences for theft ($M = 1.40, SD = 0.65$). There was also a main effect of Religion, $F(2,128) = 5.76, p = .004, \eta_p^2 = .08$ (see Table 2 for means and standard deviations). Muslim suspects were given significantly more severe sentences ($M = 2.29, SD = 1.26$) than both Christian ($M = 1.75, SD = .92, p = .016$) and Atheist suspects ($M = 1.67, SD = 1.02, p = .010$). The recommended sentences for Atheist and Christian suspects did not differ significantly ($p = 1.00$). There was no significant interaction between Crime Type and Religion on sentence severity, $F(2,128) = 1.57, p = .231, \eta_p^2 = .02$. However, simple effects analysis was undertaken to test the a-priori prediction that Muslims accused of terror crimes would be more severely sentenced than either Christian or Atheist individuals. In the terror condition, Muslims received more severe sentences than Christians ($p = .006$) and Atheists ($p = .001$). Atheists and Christians did not differ ($p = .56$). In the theft condition, no comparisons approached significance ($ps > .178$). Amongst all religions, terror crimes were sentenced more harshly than theft ($ps < .022$). Finally, the planned contrast between the Muslim / Terror condition (with a contrast weighting of -5) and other conditions (each with a contrast weighting of 1) was significant,

$F(1, 128) = 32.03, p < .001, \eta_p^2 = .20$. Muslim terror defendants were given more severe sentences than other defendants. A separate analysis including participant's own religion (coded as Christian vs. all others) revealed no effect of the covariate, $p = .64$. Similarly, analysis coding for atheist vs other showed no effect, $p = .78$.

Overall, recommended sentences were more severe for terrorism than for theft, and Muslim suspects received more severe sentences relative to Christians and Atheists suspects. This effect concatenated such that the most severe sentences were given to Muslims suspected of terrorism crimes than any other group.

Warmth and competence. Pearsons' r correlation coefficients were calculated to examine the relationship between ratings of guilt and suspect warmth and competence (see Table 3). Results indicated that suspects rated both as colder and deemed to be more competent were both more likely to be found guilty and to receive more severe sentences.

A mixed ANOVA was used to determine the effects of perceived warmth and competence of the suspect and the group. The between participants factors were Crime Type, suspect Religion and the within participants factor was the Target (suspect vs. group).

Table 3: *Correlations between the verdict, sentence, suspect warmth and suspect competence*

	Verdict	Sentence	Suspect Warmth	Suspect Competence	Group Warmth	Group Competence
Verdict	—					
Sentence	-.39***	—				
Suspect Warmth	.63***	-.44***	—			
Suspect Competence	-.26**	.17*	-.19*	—		
Group Warmth	.51***	-.36***	.74***	-.36**	—	
Group Competence	-.13	.19*	-.09	.36**	-.089	—

Note: Pearson's r shown. Statistically significant relationships are marked; * = $p < .05$, ** = $p < .01$, *** = $p < .001$.

Warmth. When measures of warmth were compared between participant's perceptions of the suspect and of the religious group, a main effect of Target was found, $F(1,128) = 5.46, p = .021, \eta_p^2 = .04$. Participants rated the suspect as colder ($M = 3.13, SD = .85$) than the group ($M = 3.27, SD = .98$). There was no main effect of Crime Type on warmth, $F(1,130) = 1.5, p = .165, \eta_p^2 = .02$, but a main effect of Religion was found, $F(2,128) = 21.36, p < .001, \eta_p^2 = .25$. Overall, Muslims were perceived as colder than both Christians ($p < .001$) and Atheists ($p < .001$) but Christians and Atheists did not differ significantly from each other ($p = .125$). There was no significant interaction between Target and Crime Type, $F(1,128) = .019, p = .890, \eta_p^2 < .01$. There was a marginally significant interaction between Target and suspect Religion, $F(2,128) = .259, p = .078, \eta_p^2 = .04$. Muslims were

perceived as colder than Christians at both the group and individual levels ($p < .001$). Christians as a group were evaluated as warmer than Atheists as a group ($p = .046$) but not at an individual target level ($p = .45$). Evaluations of warmth did not differ between individuals and groups for Muslims ($p = .94$) or Atheists ($p = .42$) but did for Christians (with the target being seen as colder than the group, $p = .002$). Crime Type and Religion conditions did not interact, $F(2, 128) = .39, p = .68$. There was no significant interaction between Target, Crime Type and Religion, $F(2, 128) = .68, p = .507, \eta_p^2 = .01$. In summary, across all religions suspects were perceived as colder than the group, but Muslim suspects were found to be colder than Christian and Atheist suspects. Muslims as a group were found to be colder than Christians and Atheists as a group. Again, additional analyses covarying out religious status showed no effect of the covariate, when coded as Christian vs. other, $p = .65$, or atheist vs. other, $p = .33$. Finally, the planned contrast between Muslim / Terror condition and other conditions was significant at both the suspect level, $F(1, 128) = 17.46, p < .001, \eta_p^2 = .12$ and group level, $F(1, 128) = 26.29, p < .001, \eta_p^2 = .17$. In summary, evaluating Muslim terror defendants led to both the group and the target being seen as colder than when other targets were evaluated.

Competence. When comparing measures of competence there was no main effect of Target (group level vs. suspect level), $F(1, 128) = .36, p = .549, \eta_p^2 < .01$. There was also no main effect of Crime Type on ratings of competence, $F(1, 128) = .80, p = .371, \eta_p^2 = .01$. There was, however, a main effect of suspect Religion on competence ratings, $F(2, 128) = 3.3.26, p = .042, \eta_p^2 = .048$. Overall, Muslims were found to be perceived as being significantly more competent than Christians ($p = .012$), but not more competent than Atheists ($p = .208$). Christians and Atheists did not differ significantly from each other ($p = .178$). There was no significant interactions between Target and Religion, $F(2, 128) = 1.06, p = .348, \eta_p^2 = .02$, or Crime Type and Religion, $F(2, 128) = .35, p = .707, \eta_p^2 = .005$. There was a marginally significant interaction between Target and Crime Type, $F(1, 130) = .3.11, p = .080, \eta_p^2 = .02$. For thefts, there was no difference in competence ratings of groups and suspects ($p = .42$). For terror crimes there was a marginally lower rating of competence for groups than individuals, ($p = .091$). Also the simple effect of Crime Type was marginally significant at the group level ($p = .055$) (with terror crimes resulting in higher competence ratings than theft crimes), but not significant at the suspect level, ($p = .80$) (see Table 4). The three-way interaction between Target, Crime Type and Religion was also not significant, $F(2, 130) = .43, p = .649, \eta_p^2 = .01$. Separate, analyses covarying out religious status showed no effect of the covariate, when comparing Christian vs. others, $p = .10$, or atheist vs. others, $p = .37$. Finally, the planned contrast between Muslim / Terror condition and other conditions was significant at the group level, $F(1, 128) = 7.56, p = .007, \eta_p^2 = .06$, but not at the suspect level, $F(1, 128) = .73, p = .396, \eta_p^2 < .01$. In summary, suspects and groups were perceived to be as competent as each other. Muslims were perceived to be more competent than Christians. Evaluating Muslim terror defendants led to higher ratings of competence for the group as a whole, but not for the target.

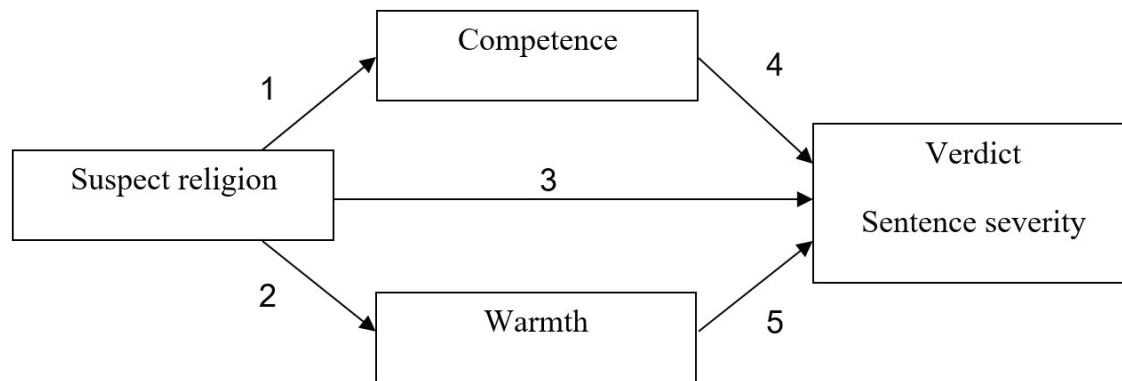
Table 4: Mean values of Warmth and Competence. Standard deviation in parentheses

Religion	Crime	Target	Warmth	Competence
Muslim	Terror	Suspect	2.48 (0.93)	3.51 (0.58)
		Group	2.38 (1.08)	3.71 (0.49)
	Theft	Suspect	2.69 (0.87)	3.46 (0.57)
		Group	2.81 (1.07)	3.51 (0.49)
Christian	Terror	Suspect	3.33 (0.83)	3.36 (0.54)
		Group	3.71 (0.86)	3.33 (0.49)
	Theft	Suspect	3.56 (0.85)	3.29 (0.54)
		Group	3.82 (0.94)	3.17 (0.48)
Atheist	Terror	Suspect	3.29 (0.58)	3.30 (0.61)
		Group	3.39 (0.58)	3.52 (0.56)
	Theft	Suspect	3.36 (0.58)	3.49 (0.78)
		Group	3.41 (0.55)	3.37 (0.53)

Mediation analysis

To test whether warmth and competence mediated the effects of suspects religion on proportion of guilty verdicts, mediation analysis was carried out using the procedure outlined using model 4 of the Hayes PROCESS macro (Hayes, 2013). Separate models were undertaken for the meditative effects of (i) suspect warmth and suspect confidence and (ii) group warmth and group confidence of (a) proportion of guilty verdicts and (b) sentence severity. The structure of these models (and path coefficient labels) can be seen in Figure 1.

One thousand bootstrap samples were taken, and the effects of mediators on the criterion variable had covariance from the other mediator and the predictor removed. Where the criterion variable was dichotomous (i.e. for verdict), an logistic regression approach was used, R^2 was calculated using the McFadden's (1974) method (where $R^2 > .02$ is regard as a good model fit), and z scores used to test path coefficients. The coefficient for each path in the models (and associated tests of significance) can be seen in Table 5.



Note: Coefficient values for paths labeled 1-5 can be found in Table 5. The indirect effect of suspect religion via competence is labeled 1,4. The indirect effect of suspect religion via warmth is labeled 2,5.

Figure 1: *Mediation model and path labels for effects of religion on verdict and sentence severity via perceived competence and warmth.*

For suspect level evaluations and sentence severity, the model was a significant predictor, $R^2 = .21$, $F(3, 130) = 11.61$, $p < .001$. For verdict, the model R^2 (McFadden) was .48, representing a very good fit. As can be seen in Table 5, in both of these models, the suspect's religion did not predict sentence severity. However, there was a significant negative indirect effect of suspects religion via perceived suspect warmth (but not competence). For group level evaluations and sentence, the model was a significant predictor, $R^2 = .17$, $F(3, 130) = 8.72$, $p < .001$. For verdict, the model was also a significant predictor, McFaddens $R^2 = .48$. For sentence severity, there was a direct effect of religion, and an indirect effect of religion via warmth (but not competence). For verdict, there was no direct effect of religion, but an indirect effect of religion via warmth (but not competence). To test whether this effect was moderated by crime type, these models repeated, adding an addition moderation of the effects of crime type on warmth and competence (Hayes 2013, Model 7). No interactions involving crime type were significant.

Table 5: Tests of path coefficients for model depicted in Figure 1, for both verdict and sentence outcomes, with group and suspect evaluations.

Level of evaluation	Criterion variable	Coefficient	Coefficient value	SE	T	p	LCI	UCI
Suspect	Sentence	1	-.04	.06	0.59	.55	-.17	.09
		2	.36	.08	4.24	<.001	.19	.53
		3	-.12	.11	1.04	.30	-.33	.10
		4	.16	.14	1.15	.25	-.12	.45
		5	-.51	.10	4.70	<.001	-.72	-.29
		1,4	-.006	.02			-.06	.01
		2,5	-.18	.07			-.33	-.07
	Verdict	1	-.04	.06	-0.59	.55	-.17	.09
		2	.36	.08	4.24	<.001	.19	.53
		3	.36	.38	0.94	.35	-.39	1.11
		4	-.53	.48	1.11*	.27	-1.46	.41
		5	2.56	.55	4.78*	<.001	1.49	3.63
		1,4	.02	.08			-.06	.32
		2,5	.92	.38			.39	1.59
Group	Sentence	1	-.08	.06	1.41	.16	-.19	.03
		2	.39	.10	3.95	<.001	.19	.58
		3	-.14	.11	1.27	.21	-.39	.08
		4	.31	.17	1.86	.066	-.02	.65
		5	-.35	.09	3.71	<.001	-.54	-.16
		1,4	-.02	.02			-.10	.00
		2,5	-.14	.05			-.28	-.05
	Verdict	1	-.08	.06	1.41	.16	-.19	.03
		2	.39	.10	3.95	<.001	.19	.58
		3	.56	.35	1.61	.11	-.12	1.25
		4	-.24	.52	-0.46*	.64	-1.26	.78
		5	1.21	.30	4.03*	<.001	.62	1.80
		1,4	.02	.07			-.08	.20
		2,5	.47	.21			.17	.99

Note: * = z score

DISCUSSION

Minority group members are found guilty and given more severe sentences more frequently than majority members (Mitchel, Haw, Pfeifer & Meissner, 2005). The current study aimed to test if this effect was observed amongst Muslim defendants who committed a terror or non-terror related crime compared to non-Muslim defendants. The current

study also sought to explore one possible mechanism which may explain such effects i.e. stereotype content.

Overall, for both proportion of guilty verdicts and sentence severity, Muslims appeared to be biased against. They were more likely to be found guilty of terror crimes than non-Muslims. Both Muslims and terror suspects received more severe sentences. These effects concatenated such that the most severe sentences were bestowed upon Muslim terror suspects than other groups. No differences were observed between Christians and Atheists in the frequency of guilty verdicts or in the severity of sentence. In both verdicts and sentence severity, Muslims defending terror charges were treated more harshly (given more frequent guilty verdicts and more severe sentences) than other conditions. This suggests that Muslims were, in the current study, the victims of a systematic discriminatory bias which, if extended to the real world, may have led to potential miscarriages of justice (particularly in terror cases).

The stereotype content model (SCM) (e.g. Fiske et al., 2002) was used to make predictions about how Muslims would be perceived when facing terror charges. In line with the existing evidence that negative stereotypes around Muslims are prevalent, Muslims as a group were perceived as both colder and more competent than other groups. The SCM suggests that such a stereotype should lead to more hostile behaviors (a pattern which was observed in the present study, see preceding paragraph). These processes were linked directly at an individual suspect level such that the colder and more competent suspects were seen to be, the more frequent guilty verdicts were received, and the greater the severity of recommended sentences. Tests of mediation revealed that the effects of a suspect's religion on these outcomes were either wholly or partly mediated by the perceptions of how warm the suspect and the suspects group were perceived. It was not mediated by perceptions of competence.

The current experiment also looked at the effects of being exposed to a Muslim or non-Muslim who was accused of theft or terror crimes on evaluations of both the defendant and the group as a whole (e.g., the stereotype of the group). Confirmation of the prevailing stereotype (e.g. a 'terrorist Muslim') led to Muslims as a group being perceived as more cold and competent. From a practical perspective this is important since the high profile nature of terror crimes may act to re-enforce such stereotypes and further exacerbate the systematic decision-making biases observed in this study. Importantly, this effect did not extend to individual defendants – evaluating a Muslim terror target did not increase the levels of competence of the defendant. In line with research on subtyping of non-typical instances of stereotyped groups (see, Richards & Hewstone, 2001), this may well be due to the fact that these defendants had been caught and already faced prosecution (and thus are counter to the stereotype and require subtyping to preserve the original stereotype). Future research could test this explanation directly by measuring levels of typicality of such targets relative to their ingroup.

The current research represents an initial exploration, so firm practical implications cannot be drawn. However, several possible considerations can be made. One important

practical implication for the criminal justice system is drawn from the observation that terror crime trials in particular may be biased against Muslim defendants. Previous research (e.g. Abshire & Bornstein, 2003; Kerr et al., 1995) suggests that bias against other outgroups (for instance, minorities) stems in large part from ingroup members' (and usually, majority ingroup members') judgments. As such, it may be important that juries in such trials represent as heterogeneous mix of individuals as possible, or that such juries are given more guidance and / or training to avoid bias than in other trials. Alternatively, instruments which measure stereotypical beliefs indirectly (e.g. the Implicit Association Tests) could be used to screen out jurors who may well hold strong stereotypes (and thus be more prone to bias). The findings that stereotype relevant crimes increase such bias could also argue for the need to reduce the use and misuse of potentially stereotypical reinforcing material, or an over emphasis of a defendant's religion. This could include not requiring a statement of faith during oaths. It would also suggest that defense counsels would perhaps be well advised to avoid as many cues as possible which highlight their client's religion.

One limitation of the current study was that, as a tradeoff between experimental control and ecological validity, identical faces could not be used in each condition. Thus, alongside differences in ethnicity and dress, other factors (e.g. levels of attractiveness) could not be controlled for completely. It is therefore possible that these differences (rather than religion) drove observed differences. We find this argument difficult to sustain since the visual source of information was one of many and, more importantly, two of the three visual targets (the two Caucasian faces) showed no differences on the majority of measures taken. Furthermore, differences occurred particularly in conditions relevant to the extant social stereotype (which one would not assume was the case if, for instance, simple differences in attractiveness or race were driving effects). Future research could use less ecologically valid stimuli to confirm the current pattern of effects.

A further limitation was that the participants had various religious affiliations (or, for many, no affiliation at all). One possibility is that participants' own religious affiliation could affect the frequency of guilty verdicts and sentence severity. This could not be ruled out in the current sample completely, although covariate analysis showing no covariate effects make it a somewhat difficult possibility to sustain. However, for all included participants, Muslims were an outgroup in terms of religion (although we concur that perceived differences / stereotypes between Muslims and each group may vary both between and within religious groups). Moreover, the current sample is broadly representative of the composition of UK juries – numerically dominated by Christians and Atheists, with a minority of followers of other religions. Thus, we would argue the results still reflect biases which are meaningful in an applied context.

A more significant limitation is the level of information given to participants and method of data collection. An online survey is not typical or realistic of a real world jury situation. Members of a jury would have access to more detailed information, including evidence, witness testimonies and suspect questioning. Although the lack of detail in the mock crime paragraph in the survey was intentional (to allow suitable ambiguity that the same scenario could be terror or theft related), this is not realistic of what a jury member

would be presented with in court. Any physical evidence presented, such as fingerprint evidence or photographs of a suspect at the crime scene would be an important factor for the jury to consider. Jury members would also be listening closely to witness testimonies and suspect questioning to be able to balance the information carefully and deliver a verdict based wholly on the information presented. Having said this, our evidence is important because it details the type of biases that individuals may be taking into the courtroom as jurors prior to exposure to any evidence per se. Such biases will likely influence how evidence is perceived and cognitively processed, as well as later decision-making.

The online survey format also does not allow for the social components that take place during jury service. Typically, jury members would deliberate the verdict such that any one member may influence the decision of another through a variety of group processes (see e.g. Devine, Clayton, Dunford, Seying & Pryce, 2001). Miller et al., 2011 showed that when judging crimes with a religious element, deliberation made mock jurors less likely to convict religious (both Christians and Muslim) defendant than non-religious defendants. In the context of the current study, whilst people may individually be susceptible to receiving stereotypes, they may be less likely to consider stereotypes in a group, especially if the stereotype is considered socially unacceptable. An individual may hold particular beliefs in private but may not be willing to declare these in public for fear of appearing racist or showing hatred towards an outgroup member. Therefore, the presence of others and group discussion may influence an individual's verdict. An online survey also lacks any real world consequences. In a real trial, a guilty verdict will have consequences for the suspect involved. At the very least, the suspect will have a criminal record, which can affect employment, housing, and relationships, and it could entail more severe consequences such as a prison sentence. When making judgments in a real world trial, jurors may be more likely to consider these consequences before delivering a verdict. In the survey, no such consequences exist so mock jury members may feel freer to be less sensitive in their judgments.

In summary this research shows, for the first time, that Muslims who are suspects in terror related crimes may be the subject of a bias leading to increased guilty verdicts and harsher sentences than members of other groups (e.g. Christians and Atheists) and also relative to Muslims accused of other sorts of crime (e.g. theft). It also suggests that the level of this bias may be driven by the extent to which defendants are viewed as cold and competent. Finally, the research suggests that exposure to Muslim terror charge defendants may well increase subsequent stereotyping of Muslims as a whole.

REFERENCES

- Abshire, J., & Bornstein, B. H. (2003). Juror sensitivity to the cross race effect. *Law and Human Behaviour*, 27, 471-480. doi: 10.1023/A:1025481905861.
- Demuth, S. and Steffensmeier, D. (2004), Ethnicity Effects on Sentence Outcomes in Large Urban Courts: Comparisons Among White, Black, and Hispanic Defendants. *Social Science Quarterly*, 85: 994–1011. doi: 10.1111/j.0038-4941.2004.00255.x.

- Devine, D. J., Clayton, L. D., Dunford, B. B., Seying, R., & Pryce, J. (2001). Jury decision making: 45 years of empirical research on deliberating groups. *Psychology, Public Policy, and Law*, 7, 622-727. doi: 10.1037//1076-8971.7.3.622.
- Dijksterhuis, A. P., & Van Knippenberg, A. D. (1995). Memory for stereotype-consistent and stereotype-inconsistent information as a function of processing pace. *European Journal of Social Psychology*, 25, 689-693. doi: 10.1002/ejsp.2420250607.
- Dwyer, C., Shah, B., & Sanghera, G. (2008). 'From cricket lover to terror suspect'—challenging representations of young British Muslim men. *Gender, Place and Culture*, 15, 117-136. doi: 10.1080/09663690701863208.
- Esqueda, C. W., Espinoza, R. K., & Culhane, S. E. (2008). The Effects of Ethnicity, SES, and Crime Status on Juror Decision Making A Cross-Cultural Examination of European American and Mexican American Mock Jurors. *Hispanic Journal of Behavioral Sciences*, 30, 181-199. doi: 10.1177/0739986308315319.
- Fiske, S.T., Cuddy, A.J.C., Glick, P. & Xu, J. (2002). A model of (often mixed) stereotype content: Competence & warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, 82, 878-902. Doi: [10.1037/0022-3514.82.6.878](https://doi.org/10.1037/0022-3514.82.6.878).
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Press.
- Johnson, S. D. (1985). Religion as a defense in a mock-jury trial. *Journal of Social Psychology*, 125, 213-220. doi: 10.1080/00224545.1985.9922874.
- Jones, C. S., & Kaplan, M. F. (2003). The effects of racially stereotypical crimes on juror decision-making and information-processing strategies. *Basic and Applied Social Psychology*, 25, 1-13. doi: 10.1207/S15324834BASP2501_1.
- Kerr, N.L., Hymes, R.W., Anderson, A.B., & Weathers, J.E. (1995). Defendant-juror similarity and mock juror judgments. *Law and Human Behaviour*, 19, 545-567. Doi 10.1007/BF01499374.
- Khan, M., & Ecklund, K. (2012). Attitudes toward Muslim Americans post-9/11. *Journal of Muslim Mental Health*, 7, 1-16. Doi: [10.3998/jmmh.10381607.0007.101](https://doi.org/10.3998/jmmh.10381607.0007.101).
- Laythe, B., Finkel, D., & Kirkpatrick, L. A. (2001). Predicting prejudice from religious fundamentalism and right-wing authoritarianism: a multiple-regression approach. *Journal for the Scientific Study of Religion*, 40, 1-10. doi: 10.1111/0021-8294.00033.
- Lynch, M. & Haney, C. (2009). Capital jury deliberation: Effects on death sentencing, comprehension and discrimination. *Law and Human Behaviour*, 33, 481-496. doi: 10.1007/s10979-008-9168-2.
- Maeder, E.M., 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.
- McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics*, (pp. 105-142). Academic Press.
- Miller, M. K., Maskaly, J., Green, M., & Peoples, C. D. (2011). The effects of deliberations and religious identity on mock jurors' verdicts. *Group Processes and Intergroup Relations*, 14, 517-532. doi: 10.1177/1368430210377458.
- Miller, M. K., Maskaly, J., Peoples, C. D., & Sigillo, A. E. (2014). The relationship between mock jurors' religious characteristics and their verdicts and sentencing decisions. *Psychology of Religion and Spirituality*, 6, 188-197. doi: [10.1037/a0036344](https://doi.org/10.1037/a0036344).
- Minear, M., & Park, D.C. (2004). A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments and Computers*, 36, 630-633. doi 10.3758/BF03206543.
- Mitchell, T.L., Haw, R.M., Pfeifer, J.E., & Meissner, C.A. (2005). Racial bias in mock juror decision making: A meta analytic review of defendant treatment. *Law and Human Behaviour*, 29, 621-638. doi 10.1007/s10979-005-8122-9.
- Mythen, G., Walklate, S., & Khan, F. (2009). 'I'm a Muslim, but I'm not a Terrorist': Victimization, Risky Identities and the Performance of Safety. *British Journal of Criminology*, 49, 736-754. doi: 10.1093/bjc/azp032.

- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2, 175-220. doi:10.1037/1089-2680.2.2.175.
- Pantazis, C., & Pemberton, S. (2009). From the 'old' to the 'new' suspect community examining the impacts of recent UK counter-terrorist legislation. *British Journal of Criminology*, 49, 646-666. doi 10.1093/bjc/azp031.
- Richards, Z., & Hewstone, M. (2001). Subtyping and subgrouping: Processes for the prevention and promotion of stereotype change. *Personality and Social Psychology Review*, 5, 52-73. doi: 10.1207/S15327957PSPR0501_4.
- Rowatt, W. C., Franklin, L. M., & Cotton, M. (2005). Patterns and personality correlates of implicit and explicit attitudes toward Christians and Muslims. *Journal for the Scientific Study of Religion*, 44, 29-43. doi 10.1111/j.1468-5906.2005.00263.x.
- Sheridan, L.P. (2006). Islamaphobia pre-and post-September 11th, 2001. *Journal of Interpersonal Violence*, 21, 317-336. doi: 10.1177/0886260505282885.
- Sommers, S. R., & Ellsworth, P. C. (2000). Race in the courtroom: Perceptions of guilt and dispositional attributions. *Personality and Social Psychology Bulletin*, 26, 1367-1379. doi: 10.1177/0146167200263005.
- Strabac, Z., & Listhaug, O. (2008). Anti-Muslim prejudice in Europe: A multilevel analysis of survey data from 30 countries. *Social Science Research*, 37, 268-286. doi: 10.1016/j.ssresearch.2007.02.00.

Date Received: 11/2017

Date Accepted: 05/2018

Suggested citation: Frings, D., Rice, K., & Albery, I.P. (2018). The Effects of Religion and Stereotype Content on Verdicts and Sentence Severity when Defending Terror Charges. [Electronic Version]. *Applied Psychology in Criminal Justice*, 14(1), 37-54.