

SIMPLY INSANE?

Attributing Terrorism to Mental Illness (Versus Ideology) Affects Mental Representations of Race

JONAS R. KUNST*

University of Oslo

Yale University

LISA S. MYHREN*

University of Oslo

IVUOMA N. ONYEADOR

Yale University

Mainstream media and public figures are often criticized for readily attributing terrorism committed by White perpetrators to mental illness, while attributing the same behavior committed by non-Whites to ideological motivation. Using a data-driven reverse-correlation approach, we show that attributing terrorism to mental illness results in a phenotypically more White mental representation of the perpetrator as compared with attributing terrorism to ideology or providing no information about its motivation. Importantly, we show that, because terrorists who are described as being motivated by mental illness are perceived as more White than those motivated by ideology, they are subsequently judged as less guilty for alleged terrorist activities. We present further evidence that this effect may be due to perceived Whiteness signaling higher socio-economic status, which reduces perceptions of culpability. In sum, our research demonstrates that extreme violence attributed to unintentional causes is perceptually associated with White perpetrators, leading to leniency in criminal judgments.

Keywords: ideology; mental illness; perpetrator; race; social perception; terrorism

In July 2011, Anders Behring Breivik left behind a 1,515 page manifesto outlining his ideological motivations for bombing the Norwegian government and killing 77 at a youth camp of the Norwegian Labour Party. Still, for months, psychiatric and legal experts and the general public heavily debated whether the actions of the White Norwegian should be attributed to mental illness or ideology. By contrast, mainstream media, law enforcement personnel, and public figures have been criticized for readily attributing acts of terror committed by non-Whites to ideology, even in the absence of suggestive evidence (Butler, 2015;

AUTHORS' NOTE: *J.R.K. and L.S.M. are shared first authors. This research was supported by a mobility grant to J.R.K. and grant 231157/F10 from the Norwegian Research Council, and by a National Science Foundation SBE Postdoctoral Research Fellowship (#1809370) to I.N.O. Correspondence concerning this article should be addressed to Jonas R. Kunst, Yale University, 2 Hillhouse Ave., New Haven, CT 06520-8205; e-mail: jonas.kunst@yale.edu.

CRIMINAL JUSTICE AND BEHAVIOR, 201X, Vol. XX, No. X, Month 2018, 1–15.

DOI: 10.1177/0093854818794742

Article reuse guidelines: sagepub.com/journals-permissions

© 2018 International Association for Correctional and Forensic Psychology

Ismail, 2016). At worst, these attributions may have negative implications for procedural justice in the context of terrorism. Using a data-driven technique, the present research aimed to investigate whether people represent an alleged terrorist as phenotypically more White when his actions are attributed to mental illness than when they are attributed to ideology. Moreover, it investigated the consequences that such a process may have for the perceived culpability of the perpetrator.

THE PORTRAYAL OF RACE IN MAINSTREAM MEDIA'S COVERAGE OF TERRORISM

In the direct aftermath of violent events, the mainstream media often adopts a “blame or explain” frame (Shahin, 2016). Whereas a “blame” frame focuses on agency, responsibility, and guilt of an alleged perpetrator, an “explain” frame focuses on forces that often lie outside of the agent’s control. Crucially, research suggests that the mainstream media’s use of these two frames often varies systematically depending on whether the perpetrator is White and American or whether he or she belongs to a racial minority group and/or is a foreigner. For instance, studies of the U.S. news media demonstrate that terrorism committed by non-Western individuals or alleged members of foreign organizations—in accordance with a blame frame—tends to be attributed to ideology (e.g., an ideological war against Christianity and the West), whereas domestic terrorists—in line with an explain frame—are humanized, individualized, and seen as mentally troubled (Powell, 2011; Turk, 2004).

This bias, however, contrasts with findings in terrorism research. Reviews of the literature suggest that mental illness might be a risk factor but is not a predictor of terrorism (Gill & Corner, 2017). Moreover, violence perpetrated by White right-wing extremists seems to be just as likely (or unlikely) to be motivated by psychopathology as violence perpetrated by Jihadists for instance (Chermak & Gruenewald, 2015; Gill, Horgan, & Deckert, 2014; Liem, van Buuren, de Roy van Zuijdewijn, Schönberger, & Bakker, 2018). Similarly, reports show that the background and profiles of suicide terrorists and individuals who are often referred to as rampage workplace and school shooters only differ superficially (Lankford, 2013; Lankford & Hakim, 2011). So, what then might be the reason people attribute violence committed by certain perpetrators to ideology but by other perpetrators to mental illness?

MOTIVATIONS TO ATTRIBUTE EXTREME VIOLENCE TO MENTAL ILLNESS

The tendency to attribute extreme violence perpetrated by in-group members to an unintentional cause such as mental illness may reflect a motivation to protect the in-group’s value (Chen, Purdie-Vaughns, Phelan, Yu, & Yang, 2015), whereas attributing the same behavior to ideology may facilitate holding out-group members morally accountable (Clark et al., 2014). For instance, when White Americans attributed the Columbine shooting to mental illness, this led to a more favorable evaluation of the perpetrators’ in-group, White American men (Chen et al., 2015). Similarly, ascribing ideological motivation to terrorists may serve the purpose of distancing the perpetrator from the in-group, whereas ascribing mental illness to a terrorist may offset this need to protect the in-group’s worth (Noor, Kteily, Siem, & Mazziotta, 2018). Finally, work in progress indicates that, among those with negative attitudes toward Muslims, non-Muslim mass shooters are seen as more driven

by mental illness than Muslim mass shooters, and that a mass shooter driven by mental illness is perceived as less likely to be Muslim than one who is driven by unknown motives (Mercier, Norris, & Shariff, 2018).

POTENTIAL CONSEQUENCES FOR MENTAL REPRESENTATIONS AND PERCEPTIONS OF CULPABILITY

The tendency to attribute terrorism committed by in-group members to mental illness but terrorism committed by out-group members to ideology may likely also be observed at a fundamental, social perceptual level. According to the dynamic interactive theory of person construal (Freeman & Ambady, 2011), higher level cognitive information about an individual can exert downstream consequences on how the race of the individual is mentally construed. Indeed, previous research demonstrated that describing an individual as engaging in stigmatized or otherwise devalued behavior, such as relying on social welfare or rejecting U.S. culture, leads people to represent the individual's race as less White and more stereotypical of racial minority groups (Brown-Iannuzzi, Dotsch, Cooley, & Payne, 2016; Kunst, Dovidio, & Dotsch, 2018). To the extent that people are motivated to protect their in-group's value, they should show a tendency to phenotypically represent a terrorist whose behavior is attributed to mental illness as racially more similar to the in-group than a terrorist whose behavior is attributed to ideology. In the case of a majority White country, this would mean that individuals belonging to the societal majority group have more White phenotypical representations of terrorists who allegedly are motivated by mental illness than of terrorists who are motivated by ideology.

Such a process may have critical downstream consequences for perceptions of culpability. The extent to which potential criminal offenders look racially stereotypical can influence how guilty they are judged to be. For instance, White people are more likely to racially profile individuals whose appearance is stereotypical of racial out-groups as potential terrorists (Kunst et al., 2018). Moreover, the less White terrorist suspects look, the harsher treatment people tend to support (Kteily, Cotterill, Sidanius, Sheehy-Skeffington, & Bergh, 2014). Crucially, analyses of actual criminal sentencing decisions show that the more defendants' racial appearance differs from the dominant racial group, the harsher sentencing they receive (Blair, Judd, & Chapleau, 2004; Eberhardt, Davies, Purdie-Vaughns, & Johnson, 2006). Thus, if terrorists who allegedly are motivated by mental illness are mentally represented as phenotypically more White than terrorists who are motivated by ideology, then this may lead to lower perceptions of culpability. Such an effect may be further explained by perceived Whiteness functioning as proxy variable for higher socio-economic status (SES). In fact, research suggests that less culpability is ascribed to offenders with high SES than to offenders with low SES (Espinoza & Willis-Esqueda, 2008). Hence, one potential reason why attributing terrorism to mental illness leads to lower culpability perceptions may be that their more White appearance signals higher SES.

THE PRESENT RESEARCH

To date, very little research exists on the effects of motivational attributions on how terrorists are racially represented and how this affects perceptions of culpability. Moreover, to the best of our knowledge, no study has so far investigated these processes at a fundamental,

perceptual level. Against this background, we tested whether attributions of terrorism to mental illness would result in a more phenotypically White and less phenotypically Middle Eastern mental representation of the perpetrator compared with when terrorism is attributed to an ideological motivation. We did so using a data-driven reverse-correlation approach (Dotsch & Todorov, 2012). This method was chosen because it allows researchers to assess people's mental representations of social categories using a bottom-up approach that imposes few a priori constraints on the potential results. Moreover, the method has proven to reliably assess biases and stereotypes across a large range of domains (see Brinkman, Todorov, & Dotsch, 2017 for a review).

Next, to investigate potential consequences for criminal justice settings, we also tested whether the mental representation of terrorists that are motivated by mental illness would be judged to be less guilty of alleged terrorist activities because they look more White. We here also tested whether perceived Whiteness would function as proxy for higher socio-economic status, which we expected to predict lower perceived guilt in accordance with previous research (Espinoza & Willis-Esqueda, 2008). We tested these relationships in two samples from Norway, a Western European, majority White country.

STUDY 1: PROBING PEOPLE'S MENTAL REPRESENTATIONS OF TERRORISTS

This first study served the purpose of probing the mental representation that people have of terrorists with different motivations. We assessed these mental representations using a reverse-correlation approach (Dotsch & Todorov, 2012). This method is designed to tap people's mental representations of social groups with relatively few a priori constraints in terms of potential results. Whereas classic survey-based verbal measurements of stereotypical beliefs, prejudices, or other types of biases naturally limit the potential answers participants can give, the reverse-correlation method allows the mental representations that participants have to emerge relatively freely from the data.

METHOD

Participants

Previous research using the reverse-correlation task suggests that 20 to 30 participants per experimental cell tend to be sufficient to observe large effects (Dotsch & Todorov, 2012; Imhoff & Dotsch, 2013). Hence, we aimed to recruit at least 60 participants for this three-condition, between-group design. In total, 75 participants were recruited via social online networks in Norway using two strategies. First, potential participants were contacted through non-political groups on Facebook for individuals interested in research and asked to participate in a study on "how people represent social categories." Second, individual users were contacted directly and asked to take part in the study. All participants were encouraged to share the link on their Facebook profile to help recruiting further participants. This procedure and description of the research was used to ensure a broad sample of participants. Because only participants who indicated a native Norwegian background were qualified given the context of this study, the final sample comprised 62 participants (72.6% females; $M_{\text{age}} = 26.18$, $SD_{\text{age}} = 10.91$). Data were collected in January 2017.¹

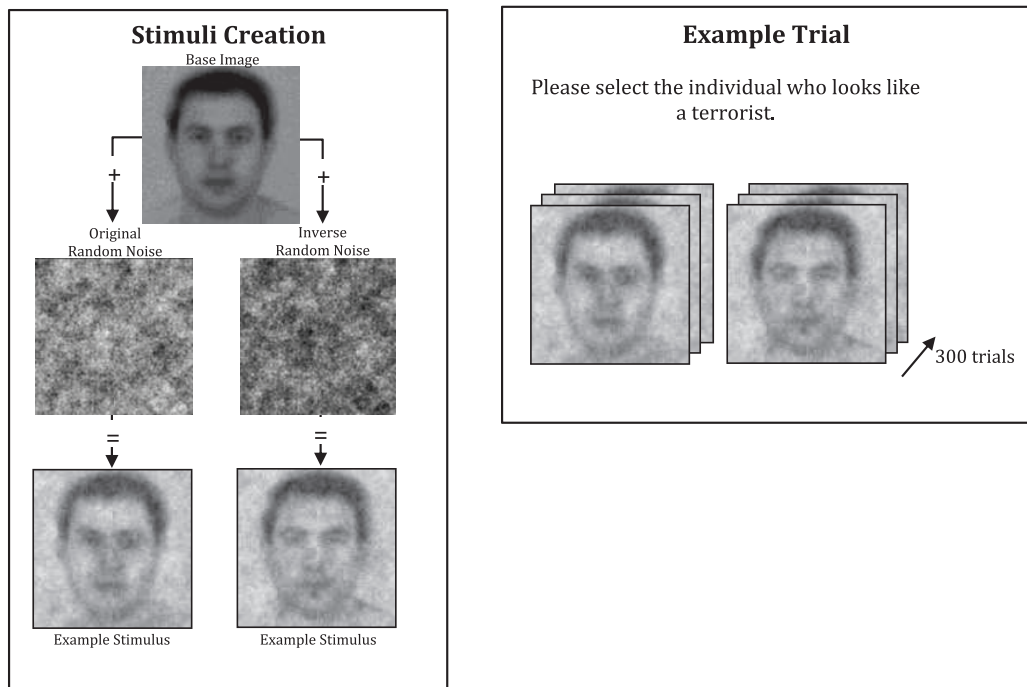


Figure 1: Procedure of Stimuli Creation and an Example Trial Are Presented for the Reverse-Correlation Task in Study 1

Note. The text in the example trial represents the instructions in the control condition.

Procedure

In reverse correlation experiments, participants are typically asked to repeatedly select from two different faces the one that they perceive best resembles a certain social group. Based on the average responses provided by all participants, researchers can then calculate the average mental representation that participants had of the social group.

Before conducting the reverse-correlation task, image stimuli have to be created. In the present research, stimuli for 300 trials were generated based on a racially ambiguous-looking base face adopted from Kunst et al. (2018). This face (see Figure 1) was created by morphing male faces from the Chicago Face Database (Ma, Correll, & Wittenbrink, 2015). Using the *rcicr* script (Dotsch, 2016), random visual noise was repeatedly generated and superimposed on the base face to create the stimuli (see Figure 1). The noise contained truncated two-cycle sinusoid patches of six orientations in all types of combinations (0° , 30° , 60° , 90° , 120° , and 150°), five spatial scales of 2, 4, 8, 16, and 32 patches per image, as well as two phases (0 and $\pi/2$) with random contrasts (see Todorov, Dotsch, Wigboldus, & Said, 2011 for a more detailed explanation of these parameters). The parameters were chosen because they produced stimuli with satisfactory visual variation in the study by Kunst et al. (2018). As displayed in Figure 2, for each trial, two stimuli images were presented, and participants were instructed to choose the image that best fit the description in the condition they had been assigned to (see details below). For each pair of stimuli, the original version of the random noise pattern was superimposed on one image, whereas the

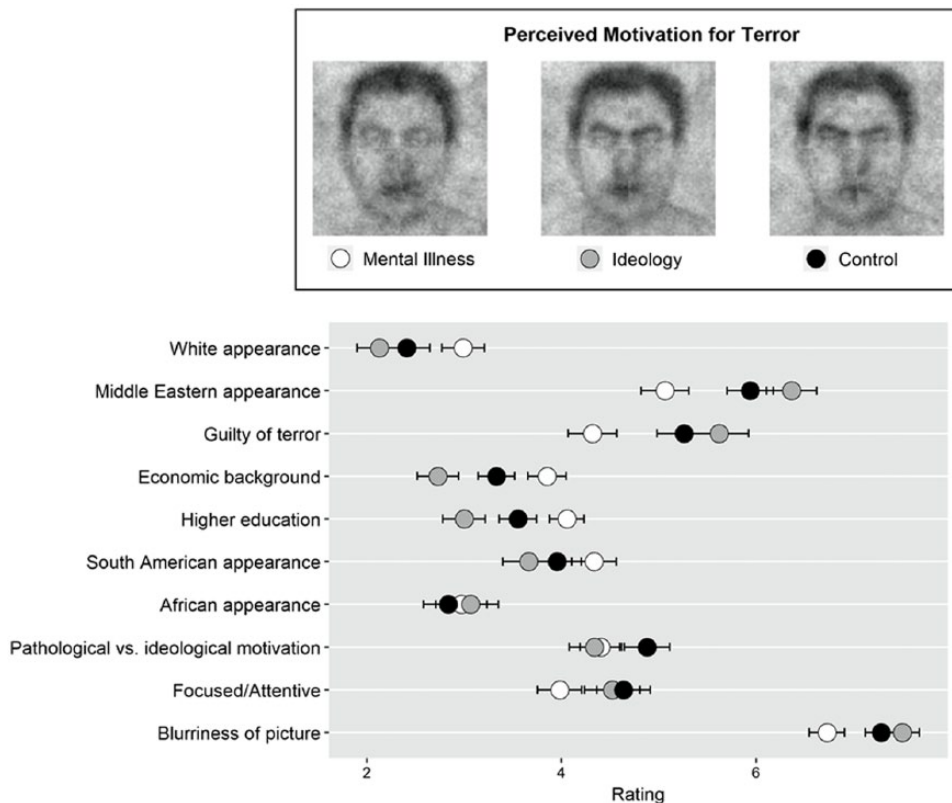


Figure 2: The Mental Representations Generated in the Three Conditions in Study 1 and Their Ratings in Study 2 Are Displayed

Note. Error bars represent ± 1 standard errors.

inverse version of this noise was superimposed on the other image, maximizing the difference between the images (Dotsch & Todorov, 2012). Whether the image with the original or inverse noise was presented first was counterbalanced. The images were horizontally presented in pairs in random order for 300 trials.

Participants were randomly assigned to a control condition, a mental illness condition, or an ideology condition. Across conditions and for each trial, participants saw the same face with a random visual noise pattern and the inverse of that random visual noise pattern superimposed, yielding images of two faces that looked different from one another (see example trial in Figure 1). To proceed to the next trial, they had to select one of the faces. Importantly, depending on the condition they had been assigned to, they were given a specific instruction for which face to select across the trials. In the control condition, participants ($n = 21$) were simply asked to pick the individual who looked the most like a terrorist. In the mental illness condition, participants ($n = 21$) were asked to pick the individual who looked the most like a terrorist driven by mental illness. In the ideology condition, participants ($n = 20$) were asked to pick the individual who looked the most like a terrorist driven by ideological reasons. As the study used a between-subjects design, participants only followed instructions from one of the three conditions. A video demonstrating the task with

instructions translated into English can be found in the online supplementary materials (see Note 1). Classification images were created by averaging the noise of the faces chosen by all participants within the different conditions and applying it to the base image using the *rcicr* script (Dotsch, 2016). This procedure has been validated within various domains of face perception (see, for example, Brinkman et al., 2017; Dotsch & Todorov, 2012; Imhoff, Woelki, Hanke, & Dotsch, 2013; Kunst, Kteily, & Thomsen, 2017; Todorov et al., 2011). The resulting classification images represent approximations of the average mental representation that participants had of the terrorist in the different conditions.

RESULTS

The aggregated classification images for each condition are presented in the upper part of Figure 2. A visual inspection of the images suggested that especially the eyebrows and nose were less pronounced and less dark in the mental illness condition, arguably resembling a more White phenotypic appearance.

DISCUSSION

This first study, which formed the first step in the reverse-correlation procedure, served the purpose of capturing the mental representations of terrorists whose motivations were attributed to different causes (i.e., ideology, mental illness, or neither in the control condition). To test our hypotheses quantitatively, in the next study, we asked an independent sample of people who were unaware of the hypotheses to rate the images on racial and other dimensions.

STUDY 2: RATINGS OF THE MENTAL REPRESENTATIONS

As is typical for reverse-correlation research, in this second study, we recruited a separate sample of participants (who were unaware of the hypotheses and how the images were generated) to rate the images generated in the first study. This procedure allows the differences in the images to be quantitatively assessed, instead of relying on subjective evaluations by the researchers. Using this procedure, we tested the following hypotheses:

Hypothesis 1: The mental representation of the terrorist motivated by mental illness will be rated as more White—and less Middle Eastern—in appearance than the mental representation of the terrorist motivated by political ideology, or the terrorist in the control condition.

Hypothesis 2: The mental representation of the terrorist motivated by mental illness will be judged as less guilty of alleged terrorist activities than the terrorists in the other two conditions.

Hypothesis 3: The latter effect will be mediated by perception of White phenotypicity and, subsequently, higher perceived socio-economic status.

METHOD

Participants

A power analysis suggested that 96 participants would provide a 90% chance to observe even small effects ($f = .15$) at a .05 significance criterion in a three-condition, within-subjects design. Thus, 151 participants were recruited from first-year lectures at a Norwegian

university for the rating of the classification images generated in the reverse-correlation task. Exclusion of 34 participants who reported a minority background and six who did not report their racial background resulted in a final sample of 111 White participants (32% females; $M_{\text{age}} = 24.14$, $SD_{\text{age}} = 8.94$). Data were collected in February 2017 (see Note 1).

Procedure

Participants rated the classification images generated in the three conditions of the reverse-correlation task in Study 1. They also rated three filler images (i.e., noise imposed images taken from the reverse-correlation stimuli of Study 1) to make the comparison of interest less salient, following the procedure of Brown-Iannuzzi et al. (2016). Participants were told that the research involved judging individuals in criminal justice settings. The participants were handed a paper questionnaire with the images presented in randomized order. The questionnaire asked participants to rate the images on several dimensions.

Perceived guilt of terrorism. First, participants read the following:

Imagine that a court invites you as an expert to evaluate evidence in a trial regarding the potential involvement in terrorism. You are told that some evidence suggests that the following individuals may be involved in suspicious activities. Using your gut feeling, to what degree do you think that the following individuals are guilty?

The reason we instructed participants to use their “gut feeling” was that we wanted to assess their spontaneous (or “automatic”) biases without evoking more elaborative information processing that may have led to socially desirable responses (Devine, Plant, Amodio, Harmon-Jones, & Vance, 2002; Plant & Devine, 2009). Having read the text, participants rated the guilt of each individual on a scale from 0 to 10 (0 = *small degree*, 10 = *large degree*).

Perceived White and Middle Eastern appearance. On a scale ranging from 0 to 10 (0 = *not at all*, 10 = *extremely*), participants rated (a) how White and (b) how Middle Eastern the individuals looked. With respect to perceived Whiteness, the term “etnisk Norsk” was used, which is the common term to describe White/Caucasian Norwegians (Phelps, Blakar, Carlquist, Nafstad, & Rand-Hendriksen, 2012).

Perceived socio-economic status. Participants rated from 0 to 10 (0 = *not at all*, 10 = *extremely*) the degree to which the individuals looked like they had higher education and the degree to which they looked financially well-off.

Distraction items. To make the hypothesis less salient to participants, they also rated on 10-point scales perceived African and South American appearance, how blurry the images looked and how focused the individuals seemed to be. Moreover, they were asked to imagine that the individuals had been convicted and to rate their perceived ideological (vs. mental illness) motivation to engage in terrorism on a scale from 0 to 10 (0 = *psychological disorder*, 10 = *ideological motivation*). This variable was measured to control for the perceived motivation of the individuals. That is, we aimed to show effects on perceived guilt that were uniquely due to changes in racial appearance by controlling for the possibility that the individual would look mentally ill. To make the focus on race less salient, the presentation order

of the items of interest and distraction items in the survey were the following: (a) perceived blurriness of the images, (b) perceived focus of the individuals, (c) perceived guilt of terrorism, (d) perceived ideological motivation, (e) perceived education, (f) perceived economic background, (g) perceived African appearance, (h) perceived Middle Eastern appearance, (i) perceived White appearance, and (j) perceived South American appearance.

RESULTS

COMPARISON OF RATINGS OF THE MENTAL REPRESENTATIONS

To compare the ratings of the mental representations, repeated-measures Analyses of Variance (ANOVAs) were conducted. Mauchly's Test of Sphericity was significant at $p \leq .010$ for all measures except for perceived African appearance ($p = .477$), perceived economic background ($p = .107$), perceived motivation ($p = .508$), and image clarity ($p = .856$). Hence, Greenhouse Geisser-corrected F values are reported for these variables. Please note that the pattern of results remained the same regardless of which correction method was used.

Ratings of the mental representations differed significantly in terms of perceived White appearance, $F(1.58, 155.90) = 22.07, p < .001, \eta_p^2 = .18$; perceived Middle Eastern appearance, $F(1.65, 163.09) = 23.17, p < .001, \eta_p^2 = .19$; perceived South American appearance, $F(1.68, 166.47) = 8.76, p = .001, \eta_p^2 = .08$; perceived economic background, $F(2, 198) = 16.34, p < .001, \eta_p^2 = .14$; perceived education, $F(1.83, 181.57) = 13.96, p < .001, \eta_p^2 = .12$; perceived image blurriness, $F(2, 198) = 17.15, p < .001, \eta_p^2 = .15$; and perceived guilt, $F(1.41, 139.93) = 26.52, p < .001, \eta_p^2 = .21$. No significant differences were observed for perceived African appearance, $F(2, 198) = 1.03, p = .359, \eta_p^2 = .01$; perceived motivation, $F(2, 198) = 1.82, p = .164, \eta_p^2 = .02$; and perceived focus, $F(1.52, 150.20) = 3.17, p = .059, \eta_p^2 = .03$.

The mean ratings for each condition are presented in Figure 2. Bonferroni-corrected post hoc comparisons showed that all three mental representations differed significantly from each other in terms of perceived White and Middle Eastern appearance. As predicted, the mental representation generated in the mental illness condition was rated as having a more White and less Middle Eastern appearance than the mental representations generated in the ideology ($ps < .001$) and control conditions ($ps \leq .001$). The latter two conditions also differed significantly, with the mental representation generated in the ideology condition rated as less White ($p = .006$) and more Middle Eastern ($p = .011$) than the mental representation generated in the control condition. Next, as predicted, the mental representation generated in the mental illness condition was rated as less guilty than the mental representations generated in the other two conditions ($ps < .001$). Moreover, the mental representation generated in the ideology condition was rated as guiltier than the mental representation generated in the control condition ($p = .006$).

In terms of the filler question, perceived South American appearance, the mental representation generated in the mental illness condition was rated as more South American than the mental representation generated in the ideology condition ($p = .001$) but did not differ significantly from the mental representation generated in the control condition ($p = .059$). Ratings of the mental representations in the control and ideology conditions did also not differ significantly from each other ($p = .100$).

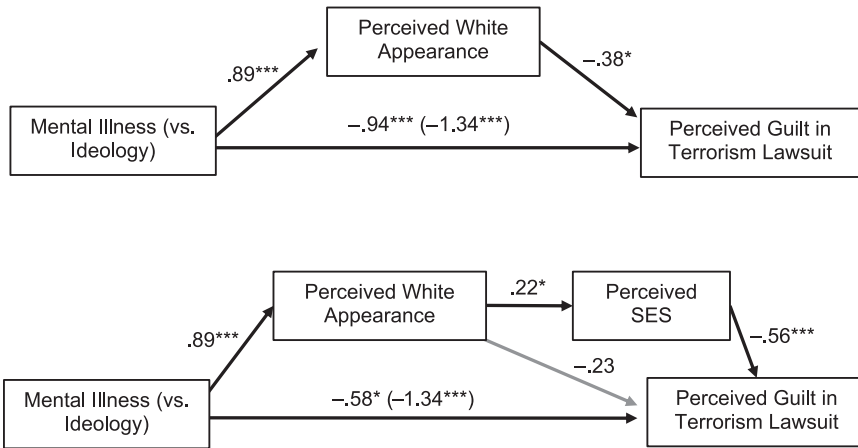


Figure 3: Simple and Extended Mediation Models Are Displayed

Note. Nonsignificant paths are presented in gray.

* $p < .05$. ** $p < .01$. *** $p < .001$.

Next, the mental representation generated in the mental illness condition was rated as being financially better off than the mental representation generated in the ideology condition ($p < .001$) but did not differ significantly from the mental representation generated in the control condition ($p = .073$). Moreover, the mental representation generated in the ideology condition was rated as being financially worse off than the mental representation generated in the control condition ($p = .001$). In terms of perceived education, all mental representations differed significantly. The mental representation generated in the mental illness condition was rated as having higher education than the mental representations generated in the ideology ($p < .001$) and control conditions ($p = .031$). The mental representation generated in the control condition was also rated as having higher education than the mental representation generated in the ideology condition ($p = .008$). Finally, the mental representation generated in the mental illness condition was rated as being less blurry than the mental representations in the other two conditions (both $ps < .001$). The mental representations generated in the control and ideology conditions did not differ significantly ($p = .338$).

TEST OF SIMPLE MEDIATION MODEL

To test whether a more White perceived appearance would explain why participants rated the mental representation of the mentally ill terrorist as less guilty, we conducted two-condition, within-participants statistical mediation analyses using the MEMORE macro (Montoya & Hayes, 2017). In all tests, we controlled for all additional rating variables, except for education and economic background, which were entered into the more complex model later on. First, we ran a model with the contrast between the mental illness and ideology conditions as independent variable. In this model (see upper model in Figure 3), perceived White appearance was the only mediator significantly predicting guilt ($p = .019$; all other mediators/filler measures: $ps > .260$). Bias-corrected bootstrapping with 10,000 random resamples showed that the contrast between the mental illness and ideology conditions indirectly led to less perceived guilt, $B = -.34$, $SE = .17$, 95% confidence interval

(CI) = $[-.75, -.05]$, due to higher perceived White appearance. All other indirect effects involving the alternate mediators were nonsignificant. An estimation of the same model comparing the mental illness condition to the control condition suggested the same trend. However, perceived White appearance here only marginally significantly predicted perceived guilt, $B = -.28$, $SE = .14$, $p = .052$ (all other mediators $ps > .103$), and the indirect effect did not reach significance, $B = -.16$, $SE = .13$, 95% CI = $[-.52, .01]$. Finally, estimating the model comparing the ideological motivation to the control condition, none of the mediators ($ps > .062$) and none of the bootstrapped indirect effects reached significance.

TEST OF EXTENDED MEDIATION MODEL

In the next step, we tested an extended version of the previous model. Specifically, the model tested whether perceived socio-economic background (i.e., the merged education and economic background variables, which were positively correlated across conditions, $rs(106) > .37$, $ps < .001$), would further mediate the effect of perceived White appearance on perceived guilt. This model was tested without covariates because the MEMORE macro currently does not support controlling for variables when testing for serial mediation. As presented in the lower model in Figure 3, perceived socio-economic status indeed fully mediated the effect of White appearance on perceived guilt. The resulting negative indirect effect of the contrast between the mental illness and ideology conditions on perceived guilt was significant, $B = -.10$, $SE = .07$, 95% CI = $[-.30, -.02]$. This extended model was not estimated with the contrast between the mental illness and control conditions, or the ideology and control conditions, given that findings from the simple mediation model were nonsignificant for these comparisons.

DISCUSSION

As predicted, the mental representation of the terrorist who was motivated by mental illness was rated as more phenotypically White and less Middle Eastern in appearance than the terrorist who was motivated by ideology or the terrorist for which no motivation was presented. Moreover, the mental representation of the mentally ill terrorist was rated as less guilty, and this was fully explained by his more White appearance and, subsequently, higher perceived socio-economic status.

Importantly, no differences were observed in perceptions of whether the potential terrorists were motivated by ideology or mental illness and the indirect effect on guilt held when controlling for this variable. This finding provides support for the discriminant validity of our effects. That is, different attributions of a terrorist's motivation led to specific and predicted changes in the perceived phenotypical appearance of those terrorists but their perceived motivation did not differ. Moreover, the mediation model demonstrated that changes in perceived Whiteness, in turn, had a unique effect on perceptions of guilt over and above the type of motivation the terrorists were perceived to have.

GENERAL DISCUSSION

The present research provided support for the hypothesis that, when terrorists are described as being driven by mental illness rather than ideology, White individuals tend to have more

White and less stereotypically Middle Eastern representations of them. In line with previous research (Blair et al., 2004; Eberhardt et al., 2006; Kunst et al., 2018), ratings by an independent sample suggested that this process may have critical downstream consequences for perceived culpability. Although in absolute terms all images were rated as more Middle Eastern than White, because the mental representation of the terrorist that was motivated by mental illness was rated as more White than the mental representation of the ideologically motivated terrorist, he was perceived as less guilty of terrorism. Further analyses suggested that the reason why increased perceived Whiteness led to less perceived guilt was that the mentally ill and consequently more White individual was perceived as having a higher socioeconomic background.

The present work contributes to an emerging area of research that focuses on perceptions of race and culpability in the context of terrorism. It has repeatedly been documented that Western mainstream media tends to readily attribute the motives of potential terrorists to ideology when the alleged perpetrator is non-White and/or foreign but attributes it to mental illness when the perpetrator is a member of the White majority group (Gill & Corner, 2017). Arguably, this bias reflects people's desire to protect the worth of their in-group and to hold out-group members morally accountable (Chen et al., 2015; Clark et al., 2014; Mercier et al., 2018; Noor et al., 2018). The present research highlights the consequences that learning about a terrorist's motivations can have at a perceptual level. Simply raising mental illness as motivation may be enough to shift the mental representations of the race of terrorists to be more White and consequently reduce perceptions of their guilt.

Our research opens up several avenues for future research on this understudied topic. At the same time, certain limitations should be noted. First, although both studies were sufficiently powered to observe effects of the expected size, the sample size in the first study and the fact that both studies used convenience samples limit the generalizability of our findings to the broader Norwegian population. For instance, the fact that parts of our recruitment strategy involved snowball sampling (i.e., we asked people to share the link to the first study on Facebook) may have led to an overrepresentation of specific subpopulations. Hence, future research will be needed to establish the replicability of our findings with larger and, preferably, representative samples.

Whereas the number of ethnic minority-group members in this research was too small to allow for a separate analysis, future studies may also test whether the effects we observed are moderated by the race of participants. For instance, given that our findings arguably represent a tendency to protect the in-group's worth, one could speculate that effects would be less pronounced among minority-group members with a Middle Eastern background (i.e., the out-group in this research). Similarly, we would also like to emphasize the importance of replicating our findings in societies that are not majority-White to test whether the terrorist who is driven by mental illness always is represented as phenotypically more similar to the racial group that is the majority in this context (irrespective of whether the context is White or not) or whether, for some reason, a terrorist driven by mental illness is generally represented as phenotypically White.

Relatedly, it would be of interest to test whether one would observe the same or similar effects when terrorism is attributed to other factors than mental illness that also tend to be perceived as being outside the perpetrator's control. For instance, it is possible that a terrorist who is described to have been brainwashed or otherwise coerced by a terrorist organization may also be mentally represented as racially more White than a perpetrator that is

described as acting out of free will. Such a scenario indeed has high societal relevance as policy makers have proposed that proof of brainwashing may spare ex-members of organizations such as ISIS of persecution (Bond, 2018).

Another avenue for future research involves the distinction between group-based and lone-wolf terrorism. Lone-wolf terrorists are substantially more likely to be mentally ill than terrorists who act as part of groups (Corner & Gill, 2015; Gruenewald, Chermak, & Freilich, 2013). Hence, it would be of interest to test whether people mentally represent lone-wolf terrorists as phenotypically more White than group-based terrorists, and whether these effects are due to people inferring that lone-wolf terrorists are driven more by mental illness and less by ideology.

In terms of the latter, another interesting direction for future research would be to investigate the effects of different types of ideologies on the mental representations of terrorists. In the present research, we did not provide information about the specific ideology of the terrorist nor did we ask participants what kind of ideology they had in mind when completing the task. Given that we used a racially ambiguous Arab–White individual and that the terrorist who was described as motivated by ideology was mentally represented as phenotypically similar to a Middle Eastern individual, it is likely that participants had Jihadist ideology in mind. In future research, it would be of interest to test the effects that different types of ideologies have on mental representations of terrorists. For instance, at least in majority White countries, we would expect right-wing ideology to lead to White mental representations but Jihadist ideology to non-White, Middle Eastern mental representations.

Although we observed systematic changes in racial appearance that proved to be highly consequential for ratings of perceived guilt, it is important to note that all terrorists in absolute terms were rated on the lower part of the Whiteness scale. We propose that a main reason for this finding is that we in all conditions asked participants to select a “terrorist,” which, due to stereotypes linking Middle Eastern individuals to terrorism, may have elicited associations with Middle Eastern individuals. Future studies should follow up on this finding, for instance, testing whether presenting participants with counter-stereotypical information (e.g., priming them with profiles of White terrorists) or asking participants to select “domestic terrorists” may lead to generally more White mental representations.

We would also like to note that the dimensions on which the images were rated are not exhaustive. It could, for instance, have been of interest to rate the images in terms of masculinity because it appears as if the mentally ill terrorist (possibly due to a less pronounced jawline and eye brows) looked less masculine than the terrorists in the other conditions. As masculine faces are more likely to be perceived as angry and threatening (Becker, Kenrick, Neuberg, Blackwell, & Smith, 2007; Gundersen & Kunst, 2018; Hess, Adams, Grammer, & Kleck, 2009), this may also mediate effects on perceived guilt, as suggested by previous research (Charlotte, Heather, & Joyce, 2012).

Finally, we would also like to note some of the potential practical implications of the present research. From a juridical perspective, the processes observed in the present paper may have negative consequences for the criminal justice system. If attributions to specific motivations change the way alleged perpetrators are racially represented, this may misguide searches and arrests at an early stage of criminal investigations before actual footage of suspects is available. For instance, attributions to ideological causes may limit profiling to Middle Eastern suspects, although the actual perpetrator may be White, and lead to initial

judgments of high culpability. By contrast, when the media and politicians readily describe mental illness as the cause of a perpetrator's mass violence, this can lead members of the public and potential future jurors to mentally construe the individual as more White and subsequently less guilty than if the mass violence was attributed to ideology. Ultimately, such biases in first impressions could influence judgments of defendants at a later stage, undermining procedural justice and the promise of a fair trial.

NOTE

1. Data, stimuli, and additional materials are available at https://osf.io/ufhne/?view_only=21e064effea544b584b60e0f5f4b4f68

REFERENCES

- Becker, D. V., Kenrick, D. T., Neuberg, S. L., Blackwell, K. C., & Smith, D. M. (2007). The confounded nature of angry men and happy women. *Journal of Personality and Social Psychology, 92*, 179-190. doi:10.1037/0022-3514.92.2.179
- Blair, I. V., Judd, C. M., & Chapleau, K. M. (2004). The influence of Afrocentric facial features in criminal sentencing. *Psychological Science, 15*, 674-679. doi:10.1111/j.0956-7976.2004.00739.x
- Bond, D. (2018). UK terror watchdog defends stance on Isis returnees. *Financial Times*. Retrieved from <https://www.ft.com/content/c02212be-0828-11e8-9650-9c0ad2d7c5b5>
- Brinkman, L., Todorov, A., & Dotsch, R. (2017). Visualising mental representations: A primer on noise-based reverse correlation in social psychology. *European Review of Social Psychology, 28*, 333-361. doi:10.1080/10463283.2017.1381469
- Brown-Iannuzzi, J. L., Dotsch, R., Cooley, E., & Payne, B. K. (2016). The relationship between mental representations of welfare recipients and attitudes toward welfare. *Psychological Science, 28*, 92-103. doi:10.1177/0956797616674999
- Butler, A. (2015, June 18). Shooters of color are called "terrorists" and "thugs." Why are white shooters called "mentally ill?" *The Washington Post*. Retrieved from https://www.washingtonpost.com/posteverything/wp/2015/06/18/call-the-charleston-church-shooting-what-it-is-terrorism/?utm_term=.43db83acbf7
- Charlotte, W., Heather, F., & Joyce, H. (2012). The effects of masculinity and suspect Gender on perceptions of guilt. *Applied Cognitive Psychology, 26*, 482-488. doi:10.1002/acp.2823
- Chen, C. Y., Purdie-Vaughns, V., Phelan, J. C., Yu, G., & Yang, L. H. (2015). Racial and mental illness stereotypes and discrimination: An identity-based analysis of the Virginia Tech and Columbine shootings. *Cultural Diversity & Ethnic Minority Psychology, 21*, 279-287. doi:10.1037/a0037881
- Chermak, S., & Gruenewald, J. A. (2015). Laying a foundation for the criminological examination of right-wing, left-wing, and Al Qaeda-inspired extremism in the United States. *Terrorism and Political Violence, 27*, 133-159. doi:10.1080/09546553.2014.975646
- Clark, C. J., Luguri, J. B., Ditto, P. H., Knobe, J., Shariff, A. F., & Baumeister, R. F. (2014). Free to punish: A motivated account of free will belief. *Journal of Personality and Social Psychology, 106*, 501-513. doi:10.1037/a0035880
- Corner, E., & Gill, P. (2015). A false dichotomy? Mental illness and lone-actor terrorism. *Law and Human Behavior, 39*, 23-34. doi:10.1037/lhb0000102
- Devine, P. G., Plant, E. A., Amodio, D. M., Harmon-Jones, E., & Vance, S. L. (2002). The regulation of explicit and implicit race bias: The role of motivations to respond without prejudice. *Journal of Personality and Social Psychology, 82*, 835-848. doi:10.1037/0022-3514.82.5.835
- Dotsch, R. (2016). rcicr: Reverse-Correlation Image-Classification Toolbox [Software]. Retrieved from <https://cran.r-project.org/web/packages/rcicr/index.html>
- Dotsch, R., & Todorov, A. (2012). Reverse correlating social face perception. *Social Psychological and Personality Science, 3*, 562-571. doi:10.1177/1948550611430272
- Eberhardt, J. L., Davies, P. G., Purdie-Vaughns, V. J., & Johnson, S. L. (2006). Looking deathworthy: Perceived stereotypicality of Black defendants predicts capital-sentencing outcomes. *Psychological Science, 17*, 383-386. doi:10.1111/j.1467-9280.2006.01716.x
- Espinoza, R. K. E., & Willis-Esqueda, C. (2008). Defendant and defense attorney characteristics and their effects on juror decision making and prejudice against Mexican Americans. *Cultural Diversity & Ethnic Minority Psychology, 14*, 364-371. doi:10.1037/a0012767
- Freeman, J. B., & Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychological Review, 118*, 247-279. doi:10.1037/a0022327
- Gill, P., & Corner, E. (2017). There and back again: The study of mental disorder and terrorist involvement. *American Psychologist, 72*, 231-241. doi:10.1037/amp0000090
- Gill, P., Horgan, J., & Deckert, P. (2014). Bombing alone: Tracing the motivations and antecedent behaviors of lone-actor terrorists. *Journal of Forensic Sciences, 59*, 425-435. doi:doi:10.1111/1556-4029.12312

- Gruenewald, J., Chermak, S., & Freilich, J. D. (2013). Distinguishing “loner” attacks from other domestic extremist violence. *Criminology & Public Policy*, *12*, 65-91. doi:10.1111/1745-9133.12008
- Gundersen, A., & Kunst, J. R. (2018). Feminist ≠ feminine? Feminist women are visually masculinized whereas feminist men are feminized. *Sex Roles*. Advance online publication. doi:10.1007/s11199-018-0931-7
- Hess, U., Adams, J. R. B., Grammer, K., & Kleck, R. E. (2009). Face gender and emotion expression: Are angry women more like men? *Journal of Vision*, *9*(12), 19-19. doi:10.1167/9.12.19
- Imhoff, R., & Dotsch, R. (2013). Do we look like me or like us? Visual projection as self- or ingroup-projection. *Social Cognition*, *31*, 806-816. doi:10.1521/soco.2013.31.6.806
- Imhoff, R., Woelki, J., Hanke, S., & Dotsch, R. (2013). Warmth and competence in your face! Visual encoding of stereotype content. *Frontiers in Psychology*, *4*, 1-8. doi:10.3389/fpsyg.2013.00386
- Ismail, N. (2016, March 15). Anders Breivik, a mass murderer who did a Nazi salute in court, isn't being called a terrorist. Why? He's white. *The Independent*. Retrieved from <http://www.independent.co.uk/voices/anders-breivik-a-mass-murderer-who-did-a-nazi-salute-in-court-isnt-being-called-a-terrorist-why-hes-a6932896.html>
- Kteily, N., Cotterill, S., Sidanius, J., Sheehy-Skeffington, J., & Bergh, R. (2014). “Not one of us”: Predictors and consequences of denying ingroup characteristics to ambiguous targets. *Personality and Social Psychology Bulletin*, *40*(10), 1231-1247. doi:10.1177/0146167214539708
- Kunst, J. R., Dovidio, J. F., & Dotsch, R. (2018). White lookalikes: Mainstream culture adoption makes immigrants “look” phenotypically White. *Personality and Social Psychology Bulletin*, *44*, 265-282.
- Kunst, J. R., Kteily, N., & Thomsen, L. (2017). “You little creep”: Evidence of blatant dehumanization of short groups. *Social Psychological and Personality Science*. Advance online publication. doi:10.1177/1948550617740613
- Lankford, A. (2013). A comparative analysis of suicide terrorists and rampage, workplace, and school shooters in the United States from 1990 to 2010. *Homicide Studies*, *17*, 255-274. doi:10.1177/1088767912462033
- Lankford, A., & Hakim, N. (2011). From Columbine to Palestine: A comparative analysis of rampage shooters in the United States and volunteer suicide bombers in the Middle East. *Aggression and Violent Behavior*, *16*, 98-107. doi:10.1016/j.avb.2010.12.006
- Liem, M., van Buuren, J., de Roy van Zuijdewijn, J., Schönberger, H., & Bakker, E. (2018). European lone actor terrorists versus “common” homicide offenders: An empirical analysis. *Homicide Studies*, *22*, 45-69. doi:10.1177/1088767917736797
- Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, *47*, 1122-1135. doi:10.3758/s13428-014-0532-5
- Mercier, B. G., Norris, A., & Shariff, A. F. (2018). *Muslim mass shooters are perceived as less mentally ill and more motivated by religion*. Unpublished manuscript. Retrieved from <https://psyarxiv.com/dwsgq/download?format=pdf>
- Montoya, A. K., & Hayes, A. F. (2017). Two-condition within-participant statistical mediation analysis: A path-analytic framework. *Psychological Methods*, *22*, 6-27. doi:10.1037/met0000086
- Noor, M., Kteily, N., Siem, B., & Mazziotta, A. (2018). “Terrorist” or “mentally ill”: Motivated biases rooted in partisanship shape attributions about violent actors. *Social Psychological and Personality Science*. Advance online publication. doi:10.1177/1948550618764808
- Phelps, J. M., Blakar, R. M., Carlquist, E., Nafstad, H. E., & Rand-Hendriksen, K. (2012). Symbolic boundaries and ideology in the Norwegian multicultural society: A longitudinal study of public discourse. *Journal of Community & Applied Social Psychology*, *22*, 187-205. doi:10.1002/casp.1126
- Plant, E. A., & Devine, P. G. (2009). The active control of prejudice: Unpacking the intentions guiding control efforts. *Journal of Personality and Social Psychology*, *96*, 640-652. doi:10.1037/a0012960
- Powell, K. A. (2011). Framing Islam: An analysis of U.S. media coverage of terrorism since 9/11. *Communication Studies*, *62*, 90-112. doi:10.1080/10510974.2011.533599
- Shahin, S. (2016). Framing “bad news.” *Journalism Practice*, *10*, 645-662. doi:10.1080/17512786.2015.1044556
- Todorov, A., Dotsch, R., Wigboldus, D. H. J., & Said, C. P. (2011). Data-driven methods for modeling social perception. *Social and Personality Psychology Compass*, *5*, 775-791. doi:10.1111/j.1751-9004.2011.00389.x
- Turk, A. T. (2004). Sociology of terrorism. *Annual Review of Sociology*, *30*, 271-286. doi:10.1146/annurev.soc.30.012703.110510

Jonas R. Kunst is a postdoctoral fellow in the Departments of Psychology at the University of Oslo and Yale University. He is also a member of the Center for Research on Extremism at the University of Oslo. His work focuses on intergroup and intercultural relations and acculturation.

Lisa S. Myhren completed her undergraduate degree in health and social psychology at the University of Oslo while the present research was conducted. Her research interests lie in the fields of social cognition, intergroup relations, and prejudice.

Ivuoma N. Onyeodor is a postdoctoral fellow in the Department of Psychology at Yale University. She is supported by NSF grant BCS-1552879 and NSF SPRF-BP Award 1809370. Her work focuses on discrimination and intergroup relations.