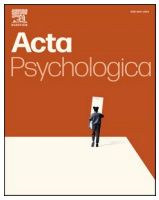
Acta Psychologica 232 (2023) 103820

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy



Shooter biases and stereotypes among police and civilians

Marleen Stelter [a,b,1], Iniobong Essien [b,c,*,1], Anette Rohmann [b], Juliane Degner [a], Stefanie Kemme [d]

[a] Universität Hamburg, Germany
[b] FernUniversität in Hagen, Germany
[c] Leuphana Universität Lüneburg, Germany
[d] University of Applied Police Sciences Hamburg, Germany

ARTICLE INFO

Keywords:
Police officer's dilemma
First-person shooter task
Shooter bias
Threat stereotypes
Anti-muslim prejudice

ABSTRACT

The present research assesses potential correlates of discriminatory police behavior, comparing police and civilian participants in a first person shooter task (FPST) as well as on various self-report measures of intergroup contact, intergroup attitudes, and ideological beliefs in three preregistered studies. Study 1 ($N = 330$), using a FPST with a short response window (630 ms), did not observe shooter biases in reaction times, error rates and signal detection parameters in neither police nor civilian participants. Study 2a ($N = 290$), using a longer response window (850 ms), observed a shooter bias in reaction times, error rates, and response criterion in both civilian and police participants. These shooter biases were largely driven by faster reactions, fewer errors, and more liberal shoot decisions for armed Arab (vs. White) targets. Study 2b ($N = 191$; 850 ms response window) closely replicated shooter biases in reaction times, error rates, and response criterion in a sample of civilian online participants. Across studies, we observed similar results in the shooter task for police and civilian samples. Furthermore, both police and civilian participants expressed anti-Muslim and anti-Arab attitudes across a variety of self-report measures. However, compared to civilians, police participants reported higher levels of anti-Muslim attitudes on some measures as well as higher levels of social dominance orientation, which might pose additional risk factors for discriminatory behavior. Lastly, while we observed reliable individual differences in self-reported intergroup attitudes, ideologies, and intergroup contact, none of these characteristics correlated with shooter biases.

1. Introduction

In 2020, incidents of police violence in the US sparked worldwide protests against police brutality (Cave et al., 2020). While protesters in many countries went to the streets in solidarity with the Black Lives Matter movement in the United States, some protests were also motivated by police misconduct and violence in protesters' own countries. In Germany, both protests and local incidents of police misconduct have stirred debates around the need to address stereotyping and systemic biases among police (Bennhold, 2020; Safronova, 2020). The present research addresses such issues by examining threat-related behavioral biases among German police and civilian samples. More specifically, we examined whether and to what extent police display behavioral biases that reflect perceptions of members of stigmatized groups as threatening.

The present research focused on threat-related behavioral biases regarding people perceived as "Arab" or "Muslim." In Germany, Arabs and Muslims face discrimination in various life domains (e.g., regarding housing, job market, and the economy, Blommaert et al., 2014; Derous et al., 2012; Koopmans et al., 2019; Mazziotta et al., 2015; Tjaden et al., 2018; Weichselbaumer, 2016). The terms "Arab" and "Muslim" are often used interchangeably in German media portrayals and public discourse (e.g., Shooman, 2012), and thus perceptions of both groups seem closely linked. Media depictions of Arabs and Muslims are mostly negative (e.g., Wigger, 2019), and focus on issues around safety threats and perceptions of cultural differences to mainstream society (Stürmer et al., 2019).

* Corresponding author at: Leuphana Universität Lüneburg, Department of Social and Organisational Psychology of Social Work, Universitätsallee 1, C1.302, 21335 Lüneburg, Germany.
E-mail addresses: marleen.stelter@fernuni-hagen.de (M. Stelter), iniobong.essien@leuphana.de (I. Essien), anette.rohmann@fernuni-hagen.de (A. Rohmann), juliane.degner@uni-hamburg.de (J. Degner), stefanie.kemme@poladium.de (S. Kemme).
[1] Marleen Stelter and Iniobong Essien have contributed equally and share first authorship.

https://doi.org/10.1016/j.actpsy.2022.103820
Received 5 October 2021; Received in revised form 21 November 2022; Accepted 19 December 2022
Available online 24 December 2022

Moreover, media framings of recent critical events[2] in Germany seem to have increased societal radicalization against Arabs, Muslims, and immigrants more generally (Stürmer et al., 2019). Such stigmatization of Arabs and Muslims is also reflected in public opinion and widespread negative attitudes towards these groups. For years, representative surveys have documented substantial levels of self-reported prejudice towards Muslims and support for anti-Muslim rhetoric (Decker et al., 2016; Savelkoul et al., 2012; Zick et al., 2016; Zick et al., 2019; Zick & Küpper, 2021). Further research suggests that people associate Arabs and Muslims mainly with concepts of danger and threat. In one study, German participants expected Muslims to be more aggressive and more sympathetic to terrorist acts than Christians (Fischer et al., 2007). Furthermore, people prompted to think of Muslims (compared to the broader category of "strangers") reported higher levels of cultural and safety threats (Spruyt & van der Noll, 2016). Taken together, current societal debates and mass media often represent Muslims and Arabs as culturally different, threatening, and dangerous. Because law enforcement institutions are immersed in this culture and police officers are socialized within the same cultural context as civilians, they are as likely as other actors in society to be affected by such negative sentiment and stereotypical representations, which may result in biased behavioral tendencies towards these social groups.

To this day, empirical field data on stereotyping and prejudice among German police is scarce (Kemme et al., 2020). Moreover, German police do not routinely document ethnicity data of the citizens they interact with (e.g., during traffic stops), which complicates assessments of biased decision-making among police. Due to this lack of data, it remains an open question whether and to what extent German police act biased towards members of stigmatized racialized groups. To address these gaps in the literature, we investigated threat-related behavioral biases in two cohorts of police students. More specifically, we employed a first-person shooter task to study spontaneous threat-related behavioral tendencies towards Arab men.

1.1. The shooter bias

In the early 2000s, high-profile incidents of police shootings of unarmed Black people in the US stimulated research examining whether people display biased behavior towards Black (vs. White) people during split-second decisions. Many of these studies have employed a first-person shooter task (FPST, Correll et al., 2002) to investigate threat-related behavioral biases. In the FPST, participants see images of armed or unarmed Black and White target individuals. Participants use a key-press to "shoot" targets carrying guns, but to "not shoot" targets carrying harmless objects. A robust finding of many of these studies is that participants are faster and make fewer errors on stereotype-congruent trials: For example, participants more quickly respond to guns in the hands of Black (vs. White) targets, but more slowly respond to harmless objects in the hands of Black (vs. White) targets. Similarly, participants more often commit errors responding to unarmed Black (vs. White) targets (and thus erroneously press "shoot"), and commit more errors responding to armed White (vs. Black) targets (and thus erroneously press "not-shoot", for a meta-analysis see Mekawi & Bresin, 2015).

Correll et al. (2002) argued that these shooter biases likely result from cultural stereotypes associating Black men in the US with threat. In split-second decisions, such cultural stereotypes are assumed to influence perception, interpretation, decision-making, and behavioral responses (Correll et al., 2002, 2016). In the FPST, these biases can be exhibited in different outcome variables: Broadly, shooter biases can be observed in reaction times and in errors, with faster responses and fewer errors in stereotype-congruent trials compared to stereotype-incongruent trials. Furthermore, drawing on signal detection theory (Green & Swets, 1966), more fine-grained analyses of errors allow disentangling the degree to which participants are able to distinguish between armed and unarmed targets (i.e., sensitivity parameter *d'*) from the degree to which participants set a lenient versus strict decision criterion to shoot (i.e, response criterion). These general patterns of effects are not always observed across all potential outcomes in the FPST and are also moderated by various contextual and methodological factors, as we will discuss in the next sections.

1.2. Shooter biases among police versus civilians

To date, most research on the shooter bias has relied on student and community samples. Only few studies have investigated police officers' racial bias in shoot decisions towards Black and White individuals, and these studies have reported mixed findings. Some studies with police officers observed shooter biases in reaction times. In a seminal study by Correll et al. (2007), police officers displayed a shooter bias in reaction times, responding more slowly to White targets holding guns and more slowly to Black targets holding harmless objects. This shooter bias among police officers was of the same magnitude as the shooter bias in a comparison sample of civilians (for other studies with police officers that observed shooter biases in reaction times, see Akinola & Mendes, 2012; Lima et al., 2018; Sadler et al., 2012).

Studies also examined shooter biases in errors (e.g., incorrectly shooting unarmed targets). For example, signal detection analyses in the above mentioned study by Correll and colleagues (2007) suggested that participants set a more lenient response criterion (i.e., a lower threshold) in their decision to shoot at Black targets compared to White targets. Although the shooter bias in the response criterion did not differ significantly between police and civilian samples, the effect was numerically less pronounced among police officers than among civilians (Correll et al., 2007). In a second study, again comparing police officers and civilians, the researchers did not replicate their previous findings. They observed that only civilians set a more lenient response criterion in their decision to shoot at Black targets compared to White targets; in contrast, police officers did not display a shooter bias in the response criterion (Correll et al., 2007; see also Ma et al., 2013). Another study by Sim and colleagues (2013) compared FPST performance between special unit officers, patrol officers, and civilians. Compared to both patrol officers and civilians, special unit officers set a more lenient response criterion in their decision to shoot at Black compared to White targets, indicating higher levels of racial bias (Sim et al., 2013).

Thus, while several studies provide evidence for shooter biases of police officers, the evidence appears inconsistent with regard to the outcome variables in which the biases occur. This inconsistency is also evident in studies that demonstrated shooter bias effects only in initial task trials but not in later trials (e.g., Plant et al., 2005; Plant & Peruche, 2005), studies which did not demonstrate racial bias among police officers in any outcome variable (Cox et al., 2014), or studies which observed reversed biases, apparently "favoring Black suspects" (James et al., 2013, p. 189, Experiment 3; James et al., 2016). Taken together, previous lab research suggests mixed findings. Some studies observed shooter biases only in reaction times, other studies observed shooter biases in errors, and yet other studies observed no or reversed shooter biases.

Two further limitations in the literature are worth mentioning. First, most studies conducted with actual police officers relied on rather small sample sizes, limiting the generalizability of each single result (see meta-analysis Mekawi & Bresin, 2015, Table 2, p. 124). The relatively low participation of police officers in studies on the shooter bias is particularly striking given increased public and scientific interest in racial bias among police. Second, almost all of the reviewed studies with police participants have been conducted in the US and to our knowledge there

[2] Critical events are historical incidents with the potential to alter public opinion regarding other, more general political issues (e.g., immigration). For example, the Cologne New Year's Eve 2016/17—after which men perceived as "Arabs" and "North Africans" were suspected of sexual assaults—substantially shifted public opinion on issues regarding immigration (Stürmer et al., 2019).

is only one study from a different societal context than the US, conducted with Brazilian military police officers (Lima et al., 2018). It is thus an open question to what extent previous research in the United States is poised to answer questions about shooter biases among police officers more generally. The question whether shooter bias findings generalize to policing outside the United States is particularly relevant given that police training and police work might differ between societies and because of evidence suggesting that shooter biases are moderated by regional differences in gun legislation (Mekawi & Bresin, 2015). Thus, societal context might matter for shooter biases. Together, these limitations call for more research investigating threat-related behavioral biases among police officers, ideally from diverse societal contexts.

1.3. Potential correlates of the shooter bias

To examine potential correlates of the shooter bias, previous research has focused on self-report measures of intergroup attitudes to examine their association with participants' performance in the FPST. For example, a study with undergraduate participants indicated that shooter biases in the FPST were moderately related to knowledge of cultural stereotypes, but were unrelated to personal beliefs (Correll et al., 2002, Study 3). Another study with police officers and civilians (Correll et al., 2007) observed that personal (threat-related) beliefs about social groups correlated with shooter biases among civilians, but no such correlations were observed among police officers. In this study, the endorsement of cultural stereotypes was uncorrelated with shooter biases among both police officers and civilians. Lastly, a meta-analysis observed only very small correlations between cultural stereotypes, personal beliefs, and shooter biases (Mekawi & Bresin, 2015). Together, these findings leave open whether or to what extent individual differences in the knowledge or endorsement of stereotypes relate to the shooter bias.

Scholars have also pointed to other psychological correlates of police behavior, such as intergroup contact, dehumanization, and distinct intergroup ideologies. Intergroup contact—particularly interactions that are perceived, interpreted, or remembered negatively—may be one contributing factor to discriminatory police behavior (Dhont et al., 2010). Similarly, in Mewaki and Bresin's meta-analysis (2015) intergroup contact was positively related to the shooter bias, with more contact with Black people being associated with larger effects. Dehumanizing attitudes—perceptions of African Americans as less human—have been linked to police officers' histories of violence against Black children and youths (e.g., Goff et al., 2014; see also Hall et al., 2016). Another potentially important correlate of police officer behavior is social dominance orientation (SDO, Sidanius & Pratto, 1999), an intergroup ideology reflecting individual preferences for social hierarchy and inequality. Previous research has linked higher levels of SDO to discriminatory behavior (Sidanius & Pratto, 1999) and to police officers' use of force (Swencionis et al., 2021). Moreover, higher levels of SDO may go hand in hand with convictions that some groups deserve harsher punishment (Sidanius et al., 2006), which may ultimately lead to more lenient shoot decisions towards members of stigmatized groups. To our knowledge, however, links between SDO and shooter biases have not previously been investigated. Taken together, measures of intergroup contact experiences, dehumanizing attitudes and SDO may also provide explanations for the occurrence of shooter biases among police officers.

1.4. The present research

The main aim of the present research was to investigate whether German police and civilian samples display similar threat-related behavioral biases towards Arab men, who are associated with threat stereotypes in many European societies (see Essien et al., 2017). Across three studies, we recruited police officers from a German police academy (Study 1 and 2a) and three civilian samples (Study 1, 2a, and 2b), who completed a FPST. Another aim of the present research was to investigate whether the magnitude of shooter biases was related to interindividual differences. Therefore, Studies 1 and 2a included questionnaires assessing stereotype endorsement, dehumanization, social dominance orientation, intergroup attitudes, and measures of intergroup contact. Preregistrations for all studies can be accessed via the Open Science Framework (OSF; Study 1: https://osf.io/n68kp Study 2a: https://osf.io/sfb7t Study 2b: https://osf.io/e3yb5). We report how we determined our sample sizes, all data exclusions, all manipulations, and all measures in the study and indicate any deviations from the preregistered method or analysis plan.[3]

2. Study 1

Study 1 had four aims. First, we investigated whether German police participants displayed shooter biases towards Arab individuals. We hypothesized that shooter biases among police participants would manifest in signal detection parameters (i.e., lower response criterion c for Arab targets than for White targets) and in reaction times (i.e., faster reactions for armed Arab targets compared to armed White targets, but slower reactions for unarmed Arab targets compared to unarmed White targets). The FPST in Study 1 limited participants' time to respond by using a restrictive response time window (630 ms; see also Mekawi & Bresin, 2015). The second goal was to test whether the magnitude of the shooter bias differed between police and civilian participants. Based on previous mixed findings regarding differences between police and civilians (Correll et al., 2007; Sim et al., 2013), we formulated a non-directional hypothesis that shooter biases would differ between police and civilian participants. Third, we investigated interindividual differences. Specifically, we hypothesized that the shooter bias would correlate positively with individual endorsement of threat-related stereotypes of Arabs (e.g., dangerous, threatening), but not with the endorsement of negative but threat-unrelated (e.g., unfamiliar), warmth-related (e.g., likable), or competence-related (e.g., competitive) stereotypes. Lastly, we hypothesized a positive correlation between the shooter bias and a measure of blatant dehumanization.

2.1. Method

2.1.1. Participants

The police sample was recruited at a police academy in Hamburg, Germany, in 2016. We aimed at recruiting the maximum number of participants available in a given cohort of police students at the police academy. As a comparison group, we also recruited a civilian sample. A sample size of $n = 100$ civilians was estimated to provide enough test power with $1–1\text{-}\beta = 0.8$, $\alpha = 0.05$ to detect a minimum effect size of $r = 0.11$ for the comparison of shooter biases between both samples (as calculated from Correll et al., 2007). The final total sample size was $N = 330$.

The sample consisted of $n = 230$ police participants ($M_{age} = 26.80$; $SD_{age} = 5.64$; 107 female; 121 male; 2 not specified) with on-the-job experience ($n = 95$; 41 %) and without on-the-job experience ($n = 135$; 59 %). Ninety-five percent of police participants reported their ethnic-cultural background to be German, 5 % reported a different ethnic-cultural background. Furthermore, the sample consisted of $n =$

[3] All analyses were conducted in R (Version 4.2.0; R Core Team, 2022) and the R-packages *citr* (Version 0.3.2; Aust, 2019), *corx* (Version 1.0.6.1; Conigrave, 2020), *cowplot* (Version 1.1.1; Wilke, 2020), *effsize* (Version 0.8.1; Torchiano, 2020), *ez* (Version 4.4.0; Lawrence, 2016), *jtools* (Version 2.2.0; Long, 2022), *lmerTest* (Version 3.1.3; Kuznetsova et al., 2017), *MBESS* (Version 4.9.1; Kelley, 2022), *papaja* (Version 0.1.1; Aust & Barth, 2022), *psych* (Mackenzie & Dudschig, 2022; Version 2.2.5; Revelle, 2022), *psychReport* (Version 3.0.2; Mackenzie & Dudschig, 2022), *Rmisc* (Version 1.5.1; Hope, 2022), *rstatix* (Version 0.7.1; Kassambara, 2022), *sjPlot* (Version 2.8.11; Lüdecke, 2022), and *tidyverse* (Version 1.3.1; Wickham et al., 2019).

100 civilian participants (M_{age} = 29.35; SD_{age} = 10.11; 65 female; 35 male; 82 % reported their ethnic-cultural background to be German; 18 % reported a different ethnic-cultural background). Based on preregistered exclusion criteria, data of three participants in the police sample were excluded from data analysis (one performed below chance level in the FSPT and two did not complete the FPST). Data of one participant in the civilian sample was excluded due to excessive non-responses (96 % of trials) in the FPST.

2.1.2. Data collection procedure

2.1.2.1. Police sample. Police participants were recruited in a police academy. Data collection took place in classrooms of the police academy and participants were tested in groups of up to 15. A White female experimenter provided initial instructions and was supported during testing by two White experimenters (one male, one female). Participants were informed that study participation was voluntary; could be canceled at any time; and individual data could be deleted upon request. After providing initial consent, participants started with the FPST. Participants then completed the following self-report measures in a fixed order: intergroup contact, stereotype endorsement, attitudes, blatant dehumanization, and demographic questions. Completing the study took approximately 30 min. After each experimental session, participants were invited to attend a lecture, in which they were informed about the aim of the study and were fully debriefed. There was no payment for study participation.

2.1.2.2. Civilian sample. The civilian sample consisted of teachers recruited at a school in Hamburg, Germany, and university students from Universität Hamburg, Germany. Data collection took place in a school building in groups of up to 15 teachers or in a computer lab on university campus. A White male experimenter provided instructions. Data collection procedures were identical to the police sample except that civilian participants were debriefed in written form. Participants were reimbursed with 5 Euros.

2.1.3. Measures

2.1.3.1. FPST. The procedure of the FPST was identical to the version used by Correll et al. (2002, Study 1) with the following modifications: First, we adapted stimulus materials used by Correll et al. (2002) but altered the skin tone of the targets and replaced heads of Black targets with Arab portraits gathered via Google Image search (see Essien et al., 2017). The resulting Arab and White targets were rated for prototypicality and perceived threat in two separate online pilot studies (see Supplement for pilot studies and analysis of prototypicality effects on responses in the FPST). Second, the FPST included 20 practice trials and 120 test trials. Targets in test trials were randomly drawn from a pool of 160 targets: 40 armed White male targets; 40 armed Arab male targets; 40 unarmed White male targets; and 40 unarmed Arab male targets. Third, the timeout window was set to 630 ms (see 2007). The FPST was run using Inquisit 4 (2014) software, based on the code from Correll et al. (2002) provided by the Millisecond Test Library.

2.1.3.2. Intergroup contact measure. As a measure of private everyday contact with Muslims, participants rated their agreement with six statements (e.g., "I know lots of Muslims." "I live or have lived in an area in which inhabitants are predominantly Muslims"; adapted from Hancock & Rhodes, 2008) on a scale from 1 (do not agree at all) to 6 (completely agree). Cronbach's α for private intergroup contact was 0.83. Using the same six-point scale, police officers with on-the-job experience also rated their amount of contact with Muslims during police work. Specifically, police officers rated their agreement with six statements (e.g., „On the job, I frequently interact with Muslims"). Cronbach's α for work-related intergroup contact was 0.77.

2.1.3.3. Stereotypes. We included a measure of stereotype endorsement based on trait ratings typically assigned to Muslim men (see Ciftci, 2012; Foroutan, 2012; Müller, 2005), which were either related to threat (*aggressive*, *considerate*, *criminal*, *dangerous*, *extremist*, *fanatic*, *harmless*, *law-abiding*, *not dangerous*, *peaceful*, *perpetrator*, *terrorist*, *threatening*, *violent*), unrelated to threat (*anti-sexist*, *civilized*, *contributing*, *undemocratic*, *unfamiliar*, *tolerant*, *traditional*), related to warmth (*good-natured*, *likeable*, *warm*), or related to competence (*competent*, *competitive*, *independent*).

Traits were presented one at a time in individually randomized sequences.[4] In a first rating block, participants estimated the percentage of German men possessing each trait on a slider from 1 to 100 (Cronbach's αs: related to threat = 0.87; unrelated to threat = 0.58; warmth = 0.73; competence = 0.60). In a second rating block, traits were presented again, and participants estimated the percentage of Muslim men possessing each trait using the same slider (Cronbach's αs: related to threat = 0.96; unrelated to threat = 0.77; warmth = 0.84; competence = 0.59).

2.1.3.4. Attitudes towards Islam. As a measure of attitudes towards Islam, participants reported their agreement with five statements (e.g., "Islam fits in German society") from Breyer and Danner (2015) on a scale ranging from 1 (do not agree at all) to 7 (completely agree). Three items measured positive attitudes and three items measured negative attitudes (one item is used by both scales). Cronbach's α for positive and negative attitudes towards Islam were 0.74 and 0.73, respectively.

2.1.3.5. Dehumanization. Participants completed a measure of blatant dehumanization (Kteily et al., 2015), in which they used sliders (ranging from 0 to 100) below a graphic depiction of human evolution to rate how human-like they perceived Muslims, Germans, and four other social groups—Americans, Greeks, Christians, and Europeans—included as distractors.

2.1.3.6. Additional variables. Participants also reported whether they had previously completed the FPST. Also, for the purpose of a separate publication (Kemme et al., 2020), we asked police students about their years of professional experience, current mission type, city districts of current and previous departments (for the past 10 years), and hours of firearms training.

2.1.3.7. Demographics. We assessed demographic information about age, gender, native language(s), nationality, and ethnicity.

2.1.4. Design

The FPST followed a 2 (Sample: police vs. civilian) by 2 (Object Type: gun vs. no-gun) by 2 (Target Ethnicity: White vs. Arab) quasi-experimental design with repeated measures on the last two factors. Dependent variables of the FPST are reaction times and response accuracies. Based on a programming error and thus deviating from the preregistered procedure, the ratio of armed versus unarmed and Arab versus White targets in the FPST varied between participants (with a minimum of 23 trials and a maximum of 37 trials in each target category combination, always adding up to a total of 120 test trials).

2.2. Results

2.2.1. Shooter bias

To analyze shooter biases for police versus civilian samples in reaction times of correct responses and in errors, we conducted preregistered

[4] Due to a programming error some traits were repeatedly presented, while other traits were not presented to the first thirteen participants in the police sample. Stereotype data from these participants were excluded from data analysis.

2 (Object Type: gun vs. no-gun) by 2 (Target Ethnicity: Arab vs. White) by 2 (Participant Group: police vs. civilians) mixed-ANOVAs with repeated-measures on the first two factors. To analyze shooter biases for police versus civilian samples in signal detection parameters, we conducted preregistered 2 (Target Ethnicity: Arab vs. White) by 2 (Participant Group: police vs. civilians) mixed ANOVAs with repeated-measures on the first factor. We first report overall shooter bias effects. Next, we report whether shooter biases differ between police and civilian participants. Additionally, we report general differences between police and civilian participants' responses in the shooter task (if there were any). Descriptive statistics are reported in Table 1 and complete ANOVA results are reported in Table 2. Additional mixed-effects model analyses are reported in the Supplement.

2.2.1.1. Reaction times. There was no significant interaction between Object Type and Target Ethnicity on reaction times in the shooter task (see Table 2), indicating no shooter bias. Furthermore, the three-way interaction between Participant Group, Object Type, and Target Ethnicity was not significant, indicating that neither participant group displayed shooter biases.

2.2.1.2. Errors. There was no shooter bias in errors, as implied by the non-significant interaction between Object Type and Target Ethnicity (see Table 2). Also, there was no interaction between Participant Group, Object type, and Target Ethnicity, indicating that neither police nor civilian participants displayed shooter biases in errors. Overall, police participants made fewer errors ($M = 0.30$, $SD = 0.16$) than civilian participants ($M = 0.35$, $SD = 0.19$; see Table 2).

2.2.1.3. Signal detection parameters. We used signal detection theory (Green & Swets, 1966) to distinguish between effects in response accuracy and response bias. We classified responses in the FPST as hits (shoot decisions for armed targets), misses (no-shoot decisions for armed targets), correct rejections (no-shoot decisions for unarmed targets), and false alarms (shoot decisions for unarmed targets). This allows calculation of the sensitivity parameter *d'* ($z_{Hits} - z_{False\ Alarms}$), indicating participants' ability to distinguish between guns and harmless objects. This also allows calculation of response criterion *c* ($-0.5 * [z_{Hits} + z_{False\ Alarms}]$), indicating participants' tendency to shoot. Hit rates of 1 or false alarm rates of 0 were adjusted following the procedure described by Macmillan and Creelman (2004).

There was no difference in response criterion c for Arab ($M = -0.16$, $SD = 0.32$) versus White targets ($M = -0.14$, $SD = 0.31$; see Table 2). Also, there was no interaction between Participant Group and Target Ethnicity, indicating that police and civilian participants did not differ in their response criteria to Arab versus White targets (see Fig. 1).

Sensitivity *d'* was generally higher for Arab targets ($M = 2.15$, $SD = 0.87$) than for White targets ($M = 1.97$, $SD = 0.81$; see Table 2). There was no interaction between Participant Group and Target Ethnicity, indicating that police and civilian participants displayed equally high levels of sensitivity for Arab versus White targets. Overall, police participants ($M = 2.14$, $SD = 0.82$) displayed a higher sensitivity than civilian participants ($M = 1.87$, $SD = 0.88$), indicating that the police sample was better at distinguishing between armed and unarmed targets (see Table 2).

Table 2
ANOVAs of reaction times, errors, and signal detection parameters in the FPST for civilians and police (Study 1).

Effect	df_n	df_d	F	p	η^2_p	90 % CI
Reaction times						
(Intercept)	1	328	137,036.14	<0.01**	1.00	[1.00, 1.00]
Participant Group	1	328	0.09	0.76	0.00	[0.00, 0.01]
Object Type	1	328	1068.96	<0.01**	0.77	[0.73, 0.79]
Target Ethnicity	1	328	0.01	0.93	0.00	[0.00, 0.00]
Participant Group x Object Type	1	328	2.61	0.11	0.01	[0.00, 0.03]
Participant Group x Target Ethnicity	1	328	1.54	0.22	0.00	[0.00, 0.02]
Object Type x Target Ethnicity	1	328	0.78	0.38	0.00	[0.00, 0.02]
Participant Group x Object Type x Target Ethnicity	1	328	1.93	0.17	0.01	[0.00, 0.03]
Errors						
(Intercept)	1	328	2301.65	<0.01**	0.88	[0.86, 0.89]
Participant Group	1	328	10.67	<0.01**	0.03	[0.01, 0.07]
Object Type	1	328	374.84	<0.01**	0.53	[0.48, 0.58]
Target Ethnicity	1	328	27.03	<0.01**	0.08	[0.04, 0.13]
Participant Group x Object Type	1	328	0.94	0.33	0.00	[0.00, 0.02]
Participant Group x Target Ethnicity	1	328	0.39	0.53	0.00	[0.00, 0.02]
Object Type x Target Ethnicity	1	328	3.32	0.07	0.01	[0.00, 0.04]
Participant Group x Object Type x Target Ethnicity	1	328	0.01	0.93	0.00	[0.00, 0.00]
Response criterion c						
(Intercept)	1	328	98.36	<0.01**	0.23	[0.17, 0.29]
Participant Group	1	328	2.74	0.10	0.01	[0.00, 0.03]
Target Ethnicity 1	1	328	0.14	0.71	0.00	[0.00, 0.01]
Participant Group x Target Ethnicity	1	328	0.32	0.57	0.00	[0.00, 0.01]
Sensitivity d'						
(Intercept)	1	328	1902.73	<0.01**	0.85	[0.83, 0.87]
Participant Group	1	328	9.18	<0.01**	0.03	[0.01, 0.06]
Target Ethnicity	1	328	20.72	<0.01**	0.06	[0.02, 0.10]
Participant Group x Target Ethnicity	1	328	0.09	0.76	0.00	[0.00, 0.01]

* $p < 0.05$; ** $p < 0.01$

2.2.2. Self-report measures

Overall, participants associated more threat-related stereotypes with Muslim men than with German men, $t(316) = -4.32$, $p < .001$, $d_z = -0.24$, 95 % CI [−0.36; −0.13]. Furthermore, participants associated more negative (threat-unrelated) stereotypes with Muslim men than

Table 1
Descriptive statistics of reaction times, errors, and signal detection parameters in the FPST for civilians and police (Study 1).

		Response time (ms)				Errors				*d'*		c	
Group	Ethnicity	*M* Gun	*SD* Gun	*M* Object	*SD* Object	*M* Gun	*SD* Gun	*M* Object	*SD* Object	*M*	*SD*	*M*	*SD*
Civilian	Arab	496	67	543	61	0.24	0.13	0.43	0.19	1.96	0.90	−0.19	0.34
Civilian	White	496	65	543	57	0.27	0.13	0.45	0.19	1.77	0.86	−0.19	0.32
Police	Arab	493	63	545	58	0.20	0.12	0.38	0.16	2.23	0.85	−0.14	0.32
Police	White	494	64	545	59	0.24	0.11	0.40	0.16	2.06	0.77	−0.12	0.31

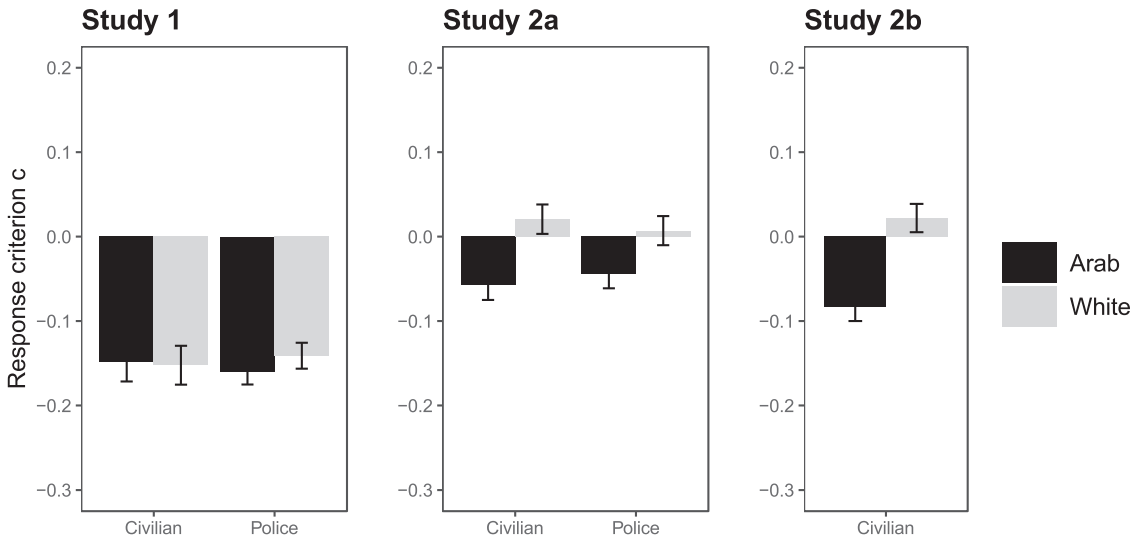


Fig. 1. Comparison of response criteria across police and civilian participants in Studies 1, 2a and 2b. Error bars show standard errors of the mean for the within-participants comparison of Arab versus White targets.

with German men, $t(316) = -19.31$, $p < .001$, $d_z = -1.34$, 95 % CI [−1.53; −1.15]. Muslim and German men were perceived as equally warm, $t(316) = -1.54$, p 0.123, $d = -0.10$, 95 % CI [−0.22; 0.03], but Muslim men were perceived as less competent than German men, $t(316) = 3.32$, p 0.001, $d_z = 0.20$, 95 % CI [0.08; 0.32]. Also, participants perceived Muslims to be less evolved than Germans, $t(598) = 5.07$, $p < .001$, $d_z = 0.38$, 95 % CI [0.28; 0.47].

Group comparisons of self-report measures of police and civilian participants are reported in Table 3. Police and civilian participants reported similar levels of private contact with Muslims, similar levels of stereotype endorsement, and similar levels of dehumanization. However, police participants reported significantly less negative stereotypes about Germans and more negative attitudes towards Islam compared to civilian participants.

2.2.3. Correlations between shooter bias and self-report measures

For preregistered correlation analyses, we calculated difference-scores of reaction times (RT) and response criterion c as indices of the shooter bias and of d' according to the following formulas: $(RT_{gun/Arab\ target} - RT_{no\ gun/White\ target}) + (RT_{gun/Arab\ target} - RT_{gun/White\ target})$, $d'_{White\ target} - d'_{Arab\ target}$, $c_{White\ target} - c_{Arab\ target}$. Similarly, we calculated four separate indices for stereotype measures (threat-related, threat-unrelated, warmth, competence), subtracting rating scores for Muslim men from rating scores for German men. We additionally calculated a dehumanization index by subtracting the dehumanization score of Muslims from the dehumanization score of Germans. Also, we added exploratory analyses of the relation between intergroup contact and the shooter bias. Correlations reported here are combined for police and civilian participants. Separate correlations for each sample are reported in the supplement (Tables S3 and S4). Results indicate that none of the three shooter bias indices were related to any of the self-report measures of stereotype endorsement, attitudes, dehumanization, and intergroup contact (see Table 4).

2.3. Discussion

In studies using the FPST, researchers typically expect to observe shooter biases in reaction times, errors, and signal detection parameters. In the FPST in the present study, participants responded overall faster and more accurately to armed targets than to unarmed targets. Contrary to our hypotheses, reaction times and error rates for both participant groups did not interact with target ethnicity. This indicates that neither police nor civilian participants displayed shooter biases in reaction times or error rates. In signal detection analyses, the response criterion did not differ between Arab and White targets, indicating that there was no shooter bias in response criterion. Taken together, police and civilian participants showed (overall) similar behavior in the FPST and the shooter bias to Arab versus White targets was less robust than expected

Table 3
Descriptive statistics for intergroup contact, stereotypes, attitudes towards Islam, and dehumanization in Study 1.

	Police		Civilian						
Measure	*M*	*SD*	*M*	*SD*	*t*	*df*	*p*	d_z	95 % CI
Contact private	3.16	0.98	3.32	1.20	1.31	328	0.19	0.16	[−0.08, 0.39]
Contact job	4.12	0.79	–	–	–	–	–	–	–
Threat stereotypes Germans	25.71	10.87	25.47	10.41	−0.19	315	0.85	−0.02	[−0.26, 0.21]
Threat stereotypes Muslims	29.69	16.24	28.72	20.35	−0.45	315	0.65	−0.05	[−0.29, 0.18]
Negative stereotypes Germans	34.43	9.70	37.49	10.41	2.56	315	0.01*	0.31	[0.07, 0.55]
Negative stereotypes Muslims	52.50	13.97	50.72	14.42	−1.04	315	0.30	−0.13	[−0.36, 0.11]
Warmth German	56.62	13.61	56.38	15.59	−0.14	315	0.89	−0.02	[−0.25, 0.22]
Warmth Muslims	57.94	17.74	58.66	19.47	0.32	315	0.75	0.04	[−0.20, 0.28]
Competence Germans	60.92	13.12	59.90	13.25	−0.64	315	0.52	−0.08	[−0.32, 0.16]
Competence Muslims	57.71	15.99	57.46	16.14	−0.13	315	0.90	−0.02	[−0.25, 0.22]
Islam (positive)	4.65	1.40	4.93	1.55	1.61	328	0.11	0.19	[−0.04, 0.43]
Islam (negative)	3.52	1.27	3.19	1.42	−2.09	328	0.04*	−0.25	[−0.49, −0.01]
Dehumanization Germans	90.90	14.61	88.93	17.11	−1.07	328	0.29	−0.13	[−0.36, 0.11]
Dehumanization Muslims	83.36	20.32	81.99	23.80	−0.53	328	0.59	−0.06	[−0.30, 0.17]

* $p < 0.05$; ** $p < 0.01$

Table 4
Zero-order correlations between shooter bias and self-report measures.

1	1	2	3	4	5	6	7	8	9	10	11	*M*	*SD*
1. d'White - d'Arab	–											−0.17	0.65
2. c White - c Arab	−0.04	–										0.01	0.33
3. Reaction time bias	−0.03	0.01	–									0.54	27.14
4. Threat stereotypes	0.04	−0.03	0.08	–								−3.75	15.45
5. Negative stereotypes	0.03	0.03	−0.03	0.66***	–							−16.55	15.26
6. Warmth stereotypes	−0.03	0.00	−0.01	−0.66***	−0.59***	–						−1.62	18.71
7. Competence stereotypes	−0.09	0.03	−0.04	−0.60***	−0.56***	0.57***	–					2.97	15.91
8. Contact (privat)	−0.03	0.09	0.04	0.22***	0.30***	−0.30***	−0.14*	–				3.21	1.05
9. Contact (job)	−0.03	0.10	0.12	−0.36***	−0.38***	0.34***	0.25*	0.14	–			4.12	0.79
10. Dehumanization	−0.02	0.00	−0.02	−0.62***	−0.48***	0.44***	0.44***	−0.24***	0.32***	–		7.36	16.53
11. Islam (negative)	0.01	0.01	−0.02	−0.55***	−0.51***	0.43***	0.40***	−0.29***	0.27**	0.47***	–	3.42	1.32
12. Islam (positive)	−0.01	−0.09	0.02	0.58***	0.52***	−0.49***	−0.42***	0.32***	−0.35***	−0.51***	−0.78***	4.73	1.45

Reaction time bias indicates the magnitude if the interaction between Object Type and Target Ethnicity; Threat stereotypes, negative stereotypes, warmth stereotypes, competence stereotypes, and dehumanization are computed as differences-scores subtracting scores for Muslim men from scores for German men.
* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

based on the literature.

Police and civilian participants did not differ in their responses in most self-reported measures: they displayed equal levels of private contact with Muslims; positive attitudes towards Islam; dehumanization of Muslims compared to Germans; and threat-related stereotypes. All participants ascribed more threat-related stereotypes to Muslim men than to German men. Similarly, all participants ascribed more competence to German than Muslim men, whereas German and Muslim men were perceived as equally warm. The only observed difference between police and civilians was regarding negative attitudes towards Islam and negative stereotypes: Compared to civilian participants, police participants reported more negative attitudes towards Islam and ascribed less negative stereotypes to German men. Taken together, police and civilian participants displayed negative evaluations and threat-related associations of Arabs and Muslims.

We observed consistent interrelations between self-report measures of stereotypes, intergroup contact, dehumanization, and attitudes towards Islam (rs = 0.11 to 0.78). This finding suggests that there were relatively robust interindividual differences in negative evaluations and threat-related associations of Arabs and Muslims. However, contrary to our hypotheses, none of these self-report measures correlated with shooter biases, neither in the police nor in the civilian sample.

There are several reasons that may account for the absence of shooter biases in Study 1. First, the short response window of 630 ms may have been too restrictive, such that differences in reaction times could no longer be detected. In the sense of a speed-accuracy tradeoff, this could have shifted effects into error rates. In the present study, however, this was not the case, with neither response latencies nor error rates revealing a shooter bias. Whereas some previous studies have observed shooter biases in errors under the same restricted response window (e.g., Correll et al., 2002, Study 2; Correll et al., 2011), others have not (Correll et al., 2007, Study 2). Thus, we cannot be certain whether the absence of a shooter bias in the present study might have been caused by the restrictive response window. Second, it is possible that participants were indeed unbiased in their responses in the FPST. Results of stereotype ratings and blatant dehumanization, however, clearly demonstrate that participants showed anti-Arab and anti-Muslim attitudes, but these appear not to have translated into response biases in the FPST. Third, procedural characteristics, such as the variance of target prototypicality and/or the unequal proportions of armed versus unarmed and Arab versus White targets in the FPST may have affected the robustness of overall effects in reaction times, errors, or signal detection parameters and thus may account for the absence of a shooter bias in Study 1. We addressed these methodological concerns in Studies 2a and b.

3. Study 2a

Study 2a used an altered FPST procedure with a less restrictive response window of 850 ms (see Correll et al., 2007; Mekawi & Bresin, 2015) and presented only highly prototypical Arab and White targets. The first aim of this study was to investigate whether police participants would display shooter biases towards Arab targets in reaction times when the procedure allows for more variability of reaction times. We hypothesized that police participants would display a shooter bias, with shorter reaction times for armed Arab (vs. White) targets, but longer reaction times for unarmed Arab (vs. White) targets. We also hypothesized that the magnitude of this shooter bias would not differ between the police and civilian sample. Because of the relatively long response time window in the FPST, we expected error rates to be lower and variance in errors to be reduced. Consequently, we did not hypothesize to observe shooter biases in errors or signal detection parameters. Another aim of Study 2a was to investigate further potential relationships between self-report measures and responses in the FPST. In addition to the measures employed in Study 1, we measured social dominance orientation (SDO), added a more subtle measure of perceived threat based on exemplar ratings, and probed participants on their positive and negative contact experiences with Muslims.

3.1. Method

3.1.1. Participants

The police sample consisted of students at a police academy in Hamburg, Germany, who had not participated in Study 1. We aimed at recruiting the maximum number of participants available in a given cohort of police students at the police academy. As a comparison group, we also recruited a civilian sample. A sample size of $n = 156$ civilians was estimated by a power analysis with $1\text{-}\beta = 0.8$, $\alpha = 0.05$, an effect size of $d_z = 0.2$ for the comparison of reaction times for armed Arab vs. armed White targets in a paired t-test. The final total sample size was $N = 290$, consisting of $n = 134$ police participants ($M_{age} = 25.94$; $SD_{age} = 5.04$; 58 female; 75 male; 1 not specified) with on-the-job experience ($n = 63$; 47 %) and without on-the-job experience ($n = 71$; 53 %). Eighty-nine percent of police participants reported their ethnic-cultural background to be German; 12 % reported a different ethnic-cultural

background. Furthermore, the sample consisted of $n = 156$ civilian participants ($M_{age} = 30.88$; $SD_{age} = 11.80$; 87 female, 69 male; 120 reported their ethnic-cultural background to be German; 35 reported a different ethnic-cultural background). One further participant in the civilian sample was excluded due to using incorrect keys in the FPST.

3.1.2. Data collection procedure

Data were collected in 2017. The setting of data collection for the police sample was identical to Study 1. Participants from the civilian sample were recruited at a local Museum and via a job board at Universität Hamburg, Germany, and tested in groups. A White female experimenter provided initial instructions. After providing consent, participants started with the FPST. After the FPST, participants completed the following measures in fixed order: A rating of Arab faces for perceived threat, a measure of the amount of private contact with Muslims (and an additional measure of on-the-job contact for police participants), a measure of positive and negative contact with Muslims, a stereotype measure, a measure of attitudes towards Islam, a scale assessing social dominance orientation, a measure of blatant dehumanization, and a demographic questionnaire. The duration of the experiment was approximately 30 min. After each experimental session, police participants and participants in the museum attended a lecture in which they were informed about the purpose of the study and were fully debriefed; participants in the computer lab were debriefed in written form. Police participants and participants in the museum were not payed for participating in the study; participants in the computer lab were reimbursed with course credit.

3.1.3. Measures

3.1.3.1. FPST. The experimental procedure of the FPST was identical to Study 1 with the following modifications: (1) The FPST included 80 test trials displaying 20 armed White male targets, 20 armed Arab male targets, 20 unarmed White male targets, and 20 unarmed Arab male targets. (2) Based on prototypicality ratings from a pilot study (see Supplement), we selected ten White faces and ten Arab faces that ranked highest in prototypicality. White and Black targets from the original FPST (Correll et al., 2002) were then each modified by replacing their faces with one of the selected White or Arab faces. As a result, we obtained ten highly prototypical Arab and ten highly prototypical White targets, each in two armed and two unarmed versions. (3) The response window was set to 850 ms (Correll et al., 2002; Correll et al., 2007).

3.1.3.2. Threat rating. We included an exemplar-based measure of threat stereotypes, in which participants rated portraits of 40 novel, male individuals (20 Arab, 20 White) on a 7-point scale (1 = not at all threatening; 7 = extremely threatening). Cronbach's α for Arab individuals = 0.95; Cronbach's α for White individuals = 0.92). Portraits were retrieved from faces databases (Langner et al., 2010; van der Schalk et al., 2011) and from a Google Image search using typical German, Turkish, and Arabic male first names as search terms. Images were presented one at a time and in random order.

3.1.3.3. Self-report measures. Participants reported the amount of contact with Muslims in private life (Cronbach's $\alpha = 0.84$) and while on the job (Cronbach's $\alpha = 0.79$), using the same items as in Study 1. In addition, we added measures for positive (Cronbach's 0.86) and negative intergroup contact (Cronbach's $\alpha = 0.87$) via ten items, asking participants to rate the frequency of positive and negative experiences with Muslims (Reimer et al., 2017). Next, participants rated threat-related (Arab men: Cronbach's $\alpha = 0.95$; German men: Cronbach's $\alpha = 0.89$) and negative stereotypes (Arab men: Cronbach's $\alpha = 0.75$; German men: Cronbach's $\alpha = 0.63$) of Arab and German men using the same items as in Study 1. Different from Study 1, we excluded measures of warmth- and competence-related stereotypes. Then, participants reported positive (Cronbach's $\alpha = 0.74$) and negative (Cronbach's $\alpha = 0.67$) attitudes towards Islam. Next, we measured social dominance orientation, using the SDO6 Scale (e.g., "Some groups of people are just more worthy than others," Pratto et al., 2006, Cronbach's $\alpha = 0.89$). Then, participants completed the same dehumanization measure as used in Study 1. At the end, participants provided demographic information about age, gender, native language, nationality, and ethnicity.

3.1.4. Design

The FPST followed a 2 (Sample: police vs. civilian) by 2 (Object Type: gun vs. no-gun) by 2 (Target Ethnicity: White vs. Arab) quasi-experimental design with repeated measures on the last two factors. Dependent variables of the FPST are reaction times and response accuracies.

3.2. Results

3.2.1. Shooter bias

To analyze shooter biases for police versus civilian samples in reaction times of correct responses and in errors, we conducted preregistered 2 (Object Type: gun vs. no-gun) by 2 (Target Ethnicity: Arab vs. White) by 2 (Participant Group: police vs. civilians) mixed-ANOVAs with repeated-measures on the first two factors. To analyze shooter biases for police versus civilian samples in signal detection parameters, we conducted preregistered 2 (Target Ethnicity: Arab vs. White) by 2 (Participant Group: police vs. civilians) mixed ANOVAs with repeated-measures on the first factor. We first report overall shooter bias effects. Next, we report whether shooter biases differ between police and civilian participants. Additionally, we report overall differences between police and civilian participants. Descriptive statistics are reported in Table 5 and results of the ANOVAs are reported in Table 6. Additional mixed-effects model analyses are reported in the Supplement.

3.2.1.1. Reaction times. We observed a shooter bias in reaction times as indicated by an interaction between Object Type and Target Ethnicity (see Table 6). Specifically, for armed targets, responses were faster for Arab targets ($M = 587$ ms, $SD = 43$) than for White targets ($M = 598$ ms, $SD = 42$), $t(291) = -7.94$, $p < .001$, $dz = -0.46$, 95 % CI [−0.59; −0.34]. Similarly, for unarmed targets, responses were faster for Arab targets ($M = 664$ ms, $SD = 41$) than for White targets ($M = 666$ ms, $SD = 43$), $t(291) = -2.31$, $p = .021$, $dz = -0.14$, 95 % CI [−0.25; −0.02], but the difference was smaller than for armed targets (see Table 5). There was no interaction between Participant Group, Object Type, and Target Ethnicity, indicating that the shooter bias in reaction times did not differ between police and civilian participants. Overall, police participants ($M = 622$ ms, $SD = 52$) responded faster than civilian participants ($M = 635$ ms, $SD = 58$).

3.2.1.2. Errors. We observed a shooter bias in errors, as indicated by an interaction between Object Type and Target Ethnicity (see Table 6). Follow-up t-tests showed that participants made fewer errors for armed Arab targets ($M = 0.06$, $SD = 0.09$) compared to armed White targets ($M = 0.10$, $SD = 0.10$), $t(291) = -7.82$, $p < .001$, $dz = -0.46$, 95 % CI [−0.58; −0.34]. However, there was no difference in errors between unarmed Arab ($M = 0.11$, $SD = 0.12$) and unarmed White targets ($M = 0.11$, $SD = 0.13$), $t(291) = 1.21$, $p = .228$, $dz = 0.07$, 95 % CI [−0.04; 0.19]. There was no interaction between Participant Group, Object Type, and Target Ethnicity, indicating that the shooter bias in errors did not differ between police and civilian participants. Overall, police participants made fewer errors overall ($M = 0.07$, $SD = 0.07$) than civilian participants ($M = 0.12$, $SD = 0.13$).

3.2.1.3. Signal detection parameters. The response criterion was overall lower for Arab targets ($M = -0.05$, $SD = 0.20$) than White targets ($M = 0.01$, $SD = 0.22$), see Table 6. This indicates that participants were

Table 5
Descriptive statistics of signal detection parameters and reaction times in the FPST for civilians and police (Study 2a).

		Response time (ms)				Errors				*d'*		c	
Group	Ethnicity	*M* Gun	*SD* Gun	*M* Object	*SD* Object	*M* Gun	*SD* Gun	*M* Object	*SD* Object	*M*	*SD*	*M*	*SD*
Civilian	Arab	590	93	666	78	0.08	0.10	0.14	0.13	3.15	0.65	−0.06	0.21
Civilian	White	604	92	670	82	0.13	0.12	0.13	0.15	2.98	0.74	0.02	0.24
Police	Arab	582	89	657	78	0.04	0.06	0.08	0.08	3.42	0.53	−0.04	0.19
Police	White	589	91	658	80	0.07	0.06	0.08	0.08	3.33	0.55	0.01	0.20

Table 6
ANOVAs of reaction times, errors, and signal detection parameters in the FPST for civilians and police (Study 2a).

Effect	df_n	df_d	*F*	*p*	η_p^2	90 % CI
Reaction times						
(Intercept)	1	290	83,780.81	<0.01**	1.00	[1.00, 1.00]
Participant Group	1	290	9.51	<0.01**	0.03	[0.01, 0.07]
Object Type	1	290	1492.15	<0.01**	0.84	[0.81, 0.86]
Target Ethnicity	1	290	55.60	<0.01**	0.16	[0.10, 0.22]
Participant Group x Object Type	1	290	0.00	0.98	0.00	[0.00, 0.00]
Participant Group x Target Ethnicity	1	290	8.05	<0.01**	0.03	[0.00, 0.06]
Object Type x Target Ethnicity	1	290	22.26	<0.01**	0.07	[0.03, 0.12]
Participant Group x Object Type x Target Ethnicity	1	290	1.43	0.23	0.00	[0.00, 0.03]
Errors						
(Intercept)	1	290	315.84	<0.01**	0.52	[0.46, 0.57]
Participant Group	1	290	23.29	<0.01**	0.07	[0.03, 0.13]
Object Type	1	290	46.28	<0.01**	0.14	[0.08, 0.20]
Target Ethnicity	1	290	18.55	<0.01**	0.06	[0.02, 0.11]
Participant Group x Object Type	1	290	1.52	0.22	0.01	[0.00, 0.03]
Participant Group x Target Ethnicity	1	290	2.05	0.15	0.01	[0.00, 0.03]
Object Type x Target Ethnicity	1	290	35.37	<0.01**	0.11	[0.06, 0.17]
Participant Group x Object Type x Target Ethnicity	1	290	1.67	0.20	0.01	[0.00, 0.03]
Response criterion c						
(Intercept)	1	328	4.23	0.04 *	0.01	[0.00, 0.04]
Participant Group	1	328	0.01	0.91	0.00	[0.00, 0.00]
Target Ethnicity	1	328	13.64	<0.01**	0.04	[0.01, 0.09]
Participant Group x Target Ethnicity	1	328	0.61	0.44	0.00	[0.00, 0.02]
Sensitivity d'						
(Intercept)	1	328	9457.38	<0.01**	0.97	[0.97, 0.97]
Participant Group	1	328	21.36	<0.01**	0.07	[0.03, 0.12]
Target Ethnicity	1	328	14.53	<0.01**	0.05	[0.02, 0.09]
Participant Group x Target Ethnicity	1	328	1.27	0.26	0.00	[0.00, 0.03]

* $p < 0.05$; ** $p < 0.01$

generally more lenient to press the "shoot" key for Arab targets compared to White targets (see Fig. 1). There was no interaction between Participant Group and Target Ethnicity (see Table 6), indicating that the shooter bias in the response criterion did not differ between police and civilian participants.

The sensitivity index *d'* was higher for Arab targets ($M = 3.27$, $SD = 0.61$) than for White targets ($M = 3.14$, $SD = 0.68$), see Table 6, indicating that participants were better able to distinguish armed from unarmed Arab targets than armed from unarmed White targets. There was no interaction between Participant Group and Target Ethnicity, indicating that police and civilian participants did not differ in their relatively higher sensitivity for Arab compared to White targets. Overall, police participants ($M = 3.37$, $SD = 0.54$) displayed a higher sensitivity *d'* in distinguishing between armed and unarmed targets than civilian participants ($M = 3.07$, $SD = 0.70$), see Table 6.

3.2.2. Self-report measures

Overall, participants associated more threat-related stereotypes with Muslim men than with German men, $t(289) = -4.91$, $p < .001$, $d_z = -0.29$, 95 % CI [−0.41; −0.17] and participants associated more negative (threat-unrelated) stereotypes with Muslim men than with German men, $t(289) = -19.64$, $p < .001$, $d_z = -1.38$, 95 % CI [−1.57; −1.18]. Furthermore, participants perceived Muslims to be less evolved than Germans, $t(537) = 4.79$, $p < .001$, $d_z = 0.37$, 95 % CI [0.29; 0.46].

Police participants reported to have significantly more private contact with Muslims than civilian participants (see Table 7). Furthermore, police participants reported to have more negative contact with Muslims than civilian participants. Also, police participants reported more negative and less positive attitudes towards Islam than civilian participants. In addition, compared to civilian participants, police participants attributed less negative stereotypes to Germans and perceived Germans as more evolved. Lastly, police participants displayed higher levels of SDO compared to civilian participants.

3.2.3. Correlations between shooter bias and self-report measures

Exploratory correlation analyses showed that shooter bias indices were uncorrelated with any of the self-report measures (see Table 8). Separate correlations for police and civilian participants are reported in Tables S6 and S7 in the Supplement.

3.3. Discussion

Consistent with our hypothesis, we observed a significant shooter bias in reaction times, with the difference between Arab and White targets being more pronounced for armed targets and less pronounced for unarmed targets. While police participants responded overall faster than civilians, the magnitude of this shooter bias in reaction times did not differ between police and civilians. Similarly, while police participants responded overall more accurately than civilians, both samples displayed a shooter bias in errors of similar magnitude, reflected in their tendency to make fewer errors for armed Arab (vs. White) targets. Moreover, both samples displayed a more lenient response criterion to Arab compared to White targets. This suggests that police and civilian participants were less careful to avoid incorrect shoot decisions for Arab (vs. White) targets, which reflects a shooter bias. These shooter biases in errors and response criterion are consistent with previous findings (e.g., Correll et al., 2007), but they were not hypothesized in our

Table 7
Descriptive statistics for intergroup contact, stereotypes, attitudes towards Islam, dehumanization, and SDO in Study 2a.

	Police		Civilian						
Measure	*M*	*SD*	*M*	*SD*	*t*	*df*	*p*	d_z	95 % CI
Contact private	3.48	1.11	3.18	1.11	−2.32	288	0.02*	−0.27	[−0.51, −0.04]
Contact job	4.27	0.91	–	–	–	–	–	–	–
Positive contact	2.67	0.88	2.84	0.95	1.55	288	0.12	0.18	[−0.05, 0.42]
Negative contact	2.43	0.91	1.9	0.91	−4.94	288	< 0.01**	−0.58	[−0.82, −0.34]
Threat stereotypes Germans	25.57	9.71	24.67	11.45	−0.71	288	0.48	−0.08	[−0.32, 0.15]
Threat stereotypes Muslims	29.87	14.82	28.86	17.99	−0.51	288	0.61	−0.06	[−0.29, 0.17]
Negative stereotypes Germans	34.07	8.36	36.82	12.84	2.12	288	0.03*	0.25	[0.02, 0.48]
Negative stereotypes Muslims	53.26	13.19	53.18	15.13	−0.05	288	0.96	−0.01	[−0.24, 0.23]
Islam (positive)	4.44	1.49	4.95	1.5	2.87	288	< 0.01**	0.34	[0.10, 0.57]
Islam (negative)	3.86	1.32	3.31	1.39	−3.43	288	< 0.01**	−0.40	[−0.64, −0.17]
Dehumanization Germans	91.67	11.8	87.24	22.05	−2.08	288	0.04*	−0.25	[−0.48, −0.01]
Dehumanization Muslims	81.5	20.17	80.19	27.05	−0.46	288	0.65	−0.05	[−0.29, 0.18]
Threat rating White	3.48	0.95	3.42	1.15	−0.48	288	0.63	−0.06	[−0.29, 0.17]
Threat rating Arab	2.74	0.82	2.72	0.9	−0.20	288	0.84	−0.02	[−0.26, 0.21]
SDO	2.68	0.84	2.29	0.98	−3.60	288	< 0.01**	−0.42	[−0.66, −0.19]

* $p < 0.05$; ** $p < 0.01$

preregistration.[5]

Different from Study 1, police participants reported more private contact with Muslims than civilian participants. Furthermore, while police and civilian participants reported similar levels of positive contact, police participants reported higher levels of negative contact with Muslims. Comparisons of attitude measures showed that police participants reported more negative and less positive attitudes towards Islam than civilian participants. Police and civilians did not differ in their stereotype expressions: Both samples ascribed more threat-related and more negative stereotypes to Muslim men than to German men. Similarly, exemplar-based threat ratings showed that participants perceived Arab faces as more threatening than White faces and this pattern did not differ between police and civilians. Also, police and civilians reported similar levels of dehumanization, with higher dehumanization of Muslims compared to Germans. Finally, compared to civilians, police officers reported higher levels of social dominance orientation. As in Study 1, we observed intercorrelations between self-report measures, but none of the measures correlated with shooter biases.

Taken together, different from Study 1, we observed consistent shooter biases in the response criterion and reaction times and these shooter biases were equally pronounced for police and civilians. As in Study 1, responses in the FPST were unrelated to any of the self-report measures.

4. Study 2b

The previous studies provide evidence that biased behavioral tendencies towards stigmatized groups might be expressed to a similar extent by both police and civilians. However, Studies 1 and 2a differ with regard to the observed FPST findings. We observed no shooter bias in Study 1, but robust shooter biases in Study 2a. To address this inconsistency, we conducted an additional close replication of Study 2a with civilian participants only. All methodological parameters of the FPST (e.g., response time window; stimuli) were identical to Study 2a. Because Study 2b focussed solely on the FPST, we did not include additional self-report measures. Study 2b had three hypotheses. First, we hypothesized that participants would display a shooter bias in reaction times, with faster reactions for armed Arab versus White targets, but slower reactions for unarmed Arab versus White targets. Second, we hypothesized that participants would make fewer errors for armed Arab versus White targets, but more errors for unarmed Arab versus White targets. Third, we hypothesized that the response criterion would be lower for Arab than for White targets.

4.1. Method

4.1.1. Participants

We aimed at recruiting 216 civilian participants via the student participant pool from a large German university and via social networks (e.g., Facebook). The sample size was estimated by a power analysis with 1-$\beta = 0.9$, $\alpha = 0.05$, for an effect size of $d_z = 0.2$ for the comparison of reaction times for armed Arab versus White targets in a paired *t*-test. We recruited a total sample of $N = 206$ participants. Fifteen participants were excluded from data analysis, because they did not complete the FPST. The final sample included $n = 191$ participants ($M_{age} = 32.90$; $SD_{age} = 10.71$; 128 female; 51 male; 1 not specified; 154 reported their ethnic-cultural background to be German; 24 reported a different ethnic-cultural background and 2 did not report their ethnic-cultural background).

4.1.2. Data collection procedure and measures

Data were collected in 2017. The experiment was conducted online using the experimental software Inquisit Web. Participants first accessed an online survey. Here, they were informed about (a) the length of the study (approx. 20 to 25 min), (b) about having to download and install the Inquisit Web Player, that (c) study participation was voluntary, (d) anonymous, and that (e) participation could be canceled at any time. Participants were then asked for their consent. Next, participants were forwarded to the Inquisit Web page, which hosted the experiment. Here, participants first downloaded and installed the Inquisit Web Player, downloaded the experimental materials, and started the experiment.

The experiment started with the FPST, identical to the version from Study 2a. After completing the FPST, participants were asked to guess the study's aim and indicated whether they had previously completed a similar task. Participants then completed a demographic questionnaire, including the following questions in fixed order: Age, gender, handedness, education, profession, German nationality, other nationality, and ethnic-cultural background. After the demographic questionnaire, participants were fully debriefed and thanked for their participation.

[5] In the preregistration, we predicted that we would not observe effects in errors or response criterion. These hypothesis formulations were based on the assumption that using a longer response time window in the FPST would shift effects from error rates into reaction times, comparable to a speed-accuracy tradeoff (e.g., Payne & Correll, 2020). However, meta-analytic evidence by Mekawi and Bresin (2015) suggests that this assumption is false. Specifically, Mekawi et al. observed that changing the response time window affected shooter biases in reaction times, but not in errors.

Table 8
Zero-order correlations between shooter bias and self-report measures.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	*M*	*SD*
1. d'White - d'Arab	–														−0.12	0.57
2. c White - c Arab	−0.02	–													0.06	0.30
3. Reaction time bias	0.08	−0.04	–												8.88	31.81
4. Threat stereotypes	−0.03	0.00	0.02	–											−4.15	14.67
5. Negative stereotypes	−0.02	0.03	0.03	0.68***	–										−17.55	15.23
6. Contact (privat)	−0.05	0.03	0.01	0.24***	0.34***	–									3.32	1.12
7. Contact (job)	−0.13	−0.11	−0.04	−0.13	−0.06	0.38***	–								4.27	0.91
8. Positive contact	−0.09	−0.04	0.06	0.37***	0.41***	0.65***	−0.01	–							2.77	0.92
9. Negative contact	0.07	0.07	0.05	−0.32***	−0.25***	0.13*	0.25	−0.13*	–						2.14	0.94
10. Dehumanization	0.00	−0.07	0.04	−0.54***	−0.45***	−0.17**	0.13	−0.22***	0.24***	–					8.52	15.39
11. Islam (negative)	0.13*	0.02	−0.06	−0.59***	−0.49***	−0.14*	0.14	−0.36***	0.44***	0.44***	–				3.55	1.39
12. Islam (positive)	−0.14*	−0.03	0.03	0.60***	0.57***	0.24***	−0.07	0.41***	−0.37***	−0.47***	−0.76***	–			4.71	1.52
13. SDO	0.03	0.02	0.08	−0.42***	−0.38***	−0.17**	0.08	−0.30***	0.39***	0.43***	0.56***	−0.56***	–		2.47	0.94
14. Threat Arabs	0.07	−0.08	0.04	−0.45***	−0.38***	−0.12*	0.02	−0.24***	0.33***	0.39***	0.44***	−0.45***	0.38***	–	3.44	1.06
15. Threat Whites	−0.03	−0.03	0.09	0.09	0.08	0.12*	0.04	0.08	0.13*	0.00	0.02	0.01	0.03	0.53***	2.73	0.86

Reaction time bias indicates the magnitude if the interaction between Object Type and Target Ethnicity; Threat stereotypes, negative stereotypes, and dehumanization are computed as differences-scores subtracting scores for Muslim men from scores for German men.
*p < 0.05; ** p < 0.01; *** p < 0.001

4.2. Results

To analyze shooter biases in reaction times of correct responses and in errors, we conducted preregistered 2 (Object Type: gun vs. no-gun) by 2 (Target Ethnicity: Arab vs. White) repeated-measures ANOVAs. To analyze shooter biases in signal detection parameters, we conducted paired *t*-tests. Descriptive statistics are reported in Table 9 and results of ANOVAs are reported in Table 10. Additional mixed-effects model analyses are reported in the Supplement.

4.2.1. Reaction times

Participants displayed a shooter bias in reaction times as indicated by an interaction between Object Type and Target Ethnicity (see Table 10). Specifically, follow-up *t*-tests showed that, for armed targets, responses were faster for Arab targets than for White targets, $t(190) = -5.85$, $p < .001$, $dz = -0.42$, 95 % CI [−0.57; −0.27]. For unarmed targets, there was no difference between reaction times for Arab targets and White targets, $t(190) = 0.83$, $p = .405$, $dz = 0.06$, 95 % CI [−0.08; 0.20].

4.2.2. Errors

Participants displayed a shooter bias in errors as indicated by an interaction between Object Type and Target Ethnicity (see Table 10). Follow-up t-tests show that, for armed targets, participants made fewer errors for Arab targets than for White targets, $t(190) = -9.07$, $p < .001$, $dz = -0.66$, 95 % CI [−0.81; −0.50]. For unarmed targets, there was no difference in errors for Arab targets and White targets, $t(190) = -1.83$, $p = .069$, $dz = -0.13$, 95 % CI [−0.27; 0.01].

4.2.3. Signal detection parameters

Results of a paired *t*-test showed that response bias *c* was smaller for Arab than for White targets, $t(190) = -4.43$, $p < .001$, $dz = -0.32$, 95 % CI [−0.47; −0.17], which suggests that participants were less careful to avoid incorrect shooting responses for Arab compared to White targets. Furthermore, sensitivity *d'* was higher for Arab compared to White targets, $t(190) = 7.56$, $p < .001$, $dz = 0.55$, 95 % CI [0.39; 0.70], which indicates that participants responded more accurately to Arab targets than to White targets. Descriptive statistics are additionally reported in Fig. 1.

4.3. Discussion

Consistent with our hypotheses and with results of Study 2a, we observed that civilian participants displayed shooter biases in reaction times, errors, and response criterion in the FPST. Shooter biases in reaction times and errors were mainly driven by participants' faster responses to armed Arab (vs. White) targets, but not by their responses to unarmed Arab (vs. White) targets. Similarly, participants displayed a shooter bias in errors, which was mainly driven by fewer errors in 'shoot'-responses to armed Arab (vs. White) targets, but not in their responses to unarmed Arab (vs. White) targets. Lastly, and consistent with our hypothesis, participants displayed a shooter bias in the response criterion, reflected in their tendency to set a lower response criterion for Arab (vs. White) targets. Taken together, findings from Study 2b fully replicate results of the civilian sample in Study 2a, thus increasing interpretability and generalizability of these findings.

5. General discussion

Three preregistered studies examined shooter biases among police and civilian samples. In Study 1, using a shorter response window, we did not observe shooter biases in reaction times, error rates, or response criterion. In Study 2a, using a longer response window, we observed a shooter bias in reaction times, error rates, and response criterion. Study 2b replicated findings from Study 2a in an online data collection. Shooter biases in reaction times (Study 2a and 2b) were largely driven by participants' faster reactions to armed Arab targets compared to

Table 9
Descriptive statistics of reaction times, errors, and signal detection parameters in the FPST for civilian participants (Study 2b).

	Reaction time (ms)				Errors				*d'*		c	
Ethnicity	*M* Gun	*SD* Gun	*M* Object	*SD* Object	*M* Gun	*SD* Gun	*M* Object	*SD* Object	*M*	*SD*	*M*	*SD*
Arab	596	97	680	81	0.12	0.12	0.20	0.17	2.82	0.79	−0.08	0.25
White	609	98	677	83	0.19	0.12	0.21	0.16	2.46	0.72	0.02	0.28

Table 10
ANOVA of reaction times, errors, and signal detection parameters in the FPST for civilian participants (Study 2b).

Effect	DFn	DFd	F	p	η_p^2	90 % CI
Reaction times						
(Intercept)	1	190	39,342.91	<0.01**	1.00	[0.99, 1.00]
Participant Group	1	190	1075.97	<0.01**	0.85	[0.82, 0.87]
Target Ethnicity	1	190	18.19	<0.01**	0.09	[0.03, 0.16]
Participant Group x Target Ethnicity	1	190	18.53	<0.01**	0.09	[0.00, 0.03]
Errors						
(Intercept)	1	190	39,342.91	<0.01**	1.00	[0.99, 1.00]
Participant Group	1	190	1075.97	<0.01**	0.85	[0.82, 0.87]
Target Ethnicity	1	190	18.19	<0.01**	0.09	[0.03, 0.16]
Participant Group x Target Ethnicity	1	190	18.53	<0.01**	0.09	[0.00, 0.03]

* $p < 0.05$; ** $p < 0.01$

armed White targets. Similarly, shooter biases in errors (Study 2a and 2b) were largely driven by fewer errors for armed Arab targets compared to armed White targets. Lastly, we observed shooter biases in response criterion (Study 2a and 2b), with a lower criterion for Arab compared to White targets. Importantly, we observed similar effects across police and civilian samples and across reaction times, errors, and response criterion.

While shooter biases did not differ between both samples, police and civilian responses in the FPST did differ systematically in two ways. Overall, police participants tended to respond faster (Study 2a) and more accurately (Study 1 and 2a) than civilian participants. This finding is consistent with previous studies indicating that trained experts perform better in the FPST (e.g., Correll et al., 2007).

In addition, we examined relationships between shooter biases and self-reported individual differences. We observed consistent interrelations among self-report measures of intergroup contact, stereotype endorsement, intergroup attitudes, dehumanization, and social dominance orientation. These correlations indicate that participants reliably differed in the extent to which they endorsed negative and threat-related stereotypes and beliefs; perceived Muslims as less human; and preferred social hierarchies and inequality. However, these individual differences were not related to the shooter bias. These findings fall in line with previous mixed findings regarding relationships between shooter biases and measures of interindividual differences (Correll et al., 2002; Correll et al., 2007; Mekawi & Bresin, 2015).

Several reasons might explain the absence of relationships between the shooter bias and self-report measures. First, it is possible that there are no reliable individual differences in shooter biases in the FPST. Shooter biases may reflect cultural knowledge of associations between threat and social groups, which are shared among most members of society. Consistent with this idea, previous research has observed similar shooter biases among Black and White participants, suggesting that shared knowledge rather than individual attitudes account for FPST performance (e.g., Correll et al., 2002). Second, responses in the FPST are made under time constraints, suggesting that effects are influenced by spontaneous responding towards members of social groups. Conversely, self-report measures may underlie more deliberate responding and may thus not only reflect stereotype activation, but also additional propositional processes (e.g., non-prejudicial goals; Gawronski et al., 2012; Gawronski & Bodenhausen, 2006). Consequently, FPST and self-report measures may capture different psychological processes. Third, relationships between self-report measures and shooter biases might be attenuated by low reliability due to high measurement error of the FPST (Payne & Correll, 2020). Taken together, several explanations might account for the fact that individual differences in intergroup contact experiences, stereotype endorsement, attitudes, dehumanization, and social dominance orientation do not correlate with the shooter bias. Future research should follow up on these different explanations by systematically examining whether or under which conditions shooter biases may reflect reliable individual differences.

Individual differences among police officers regarding their experiences of intergroup contact, stereotype endorsement, intergroup ideologies, and dehumanization are worth studying in their own right, as they might provide important psychological antecedents to intergroup behavior more generally and policing specifically (see Swencionis & Goff, 2017). Compared to civilians, police samples reported either similar or higher levels of private intergroup contact with Muslims. Furthermore, civilian and police participants reported similar levels of positive contact with Muslims, but police participants reported higher levels of negative contact with Muslims. In terms of individual differences in attitudes and ideologies, we observed some similarities between police and civilian participants. Both samples similarly ascribed threat-related stereotypes to Muslim men and perceived Arab faces as more threatening than White faces. Moreover, police and civilian participants displayed similar levels of self-reported blatant dehumanization of Muslims.

However, police and civilian participants also differed in some regards. First, all participants ascribed negative stereotypes to Muslim men, but this effect was larger for police than for civilian participants. Second, police participants reported more negative attitudes towards Islam than civilians. Third, compared to civilians, police participants displayed a stronger preference for social hierarchy and inequality (i.e., social dominance orientation). In sum, both police and civilian participants endorsed anti-Muslim and anti-Arab attitudes across a variety of self-report measures. This research provides preliminary evidence that on some measures, police participants may report even higher levels of anti-Muslim attitudes than civilians, which may pose additional risk factors for discriminatory behavior. Future studies may investigate whether these differences between police and civilians replicate and whether they are caused by self-selection biases or by socialization effects (see also Kemme et al., 2021).

Comparisons between police and civilian samples in the present research are also complicated by limitations. First, police participants were recruited at a single institution and are thus not representative of all police officers (e.g., regarding age, work experience, education, or geographic context). Second, the present research relied on a convenience sample of civilian participants, which are not representative of the general population, and thus can only provide preliminary evidence about differences between police and civilians. Consequently, it is a question for future research whether or to what extent the present findings generalize to other police samples and subgroups.

That being said, the present research provides a preliminary but thorough assessment of potential correlates of discriminatory police behavior, comparing police and civilian participants' responses on a variety of self-report measures of intergroup attitudes as well as spontaneously activated threat-related responses. We observed similar shooter biases between police and civilian samples towards Arab versus White targets. These studies suggest that shooter biases may generalize beyond societal contexts and target groups. So far, few studies have provided insights into individual differences in intergroup attitudes between police and civilians. Given increasing public interest in police behavior, future research should investigate relations between intergroup attitudes and discriminatory behavior in large-scale multi-method studies.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

We thank Katja Coser, Tobias Kayser, Jana Mangels, Nina Müller, Diana Pistoll, Marc Rommel, Natascha Rütter, and Sven Schönig for assisting with the preparation of stimulus materials and data collection.

This publication was funded by the Open Access Publication Fund of the Leuphana University Lüneburg.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.actpsy.2022.103820.

References

Akinola, M., & Mendes, W. B. (2012). Stress-induced cortisol facilitates threat-related decision making among police officers. *Behavioral Neuroscience, 126*(1), 167–174. https://doi.org/10.1037/a0026657

Aust, F. (2019). Citr: 'RStudio' add-in to insert markdown citations. https://github.com/crsh/citr.

Aust, F., & Barth, M. (2022). Papaja: Prepare reproducible APA journal articles with R markdown. https://github.com/crsh/papaja.

Bennhold, K. (2020, September 16). 29 officers suspected of sharing Hitler images are suspended in Germany. *New York Times.* https://www.nytimes.com/2020/09/16/world/europe/germany-police-far-right-chat-group.html.

Blommaert, L., Coenders, M., & van Tubergen, F. (2014). Discrimination of arabic-named applicants in the Netherlands: An internet-based field experiment examining different phases in online recruitment procedures. *Social Forces, 92*(3), 957–982. https://doi.org/10.1093/sf/sot124

Breyer, B., & Danner, D. (2015). Einstellung zum Islam (ALLBUS). In D. Danner, & A. Glöckner-Rist (Eds.), *Zusammenstellung sozialwissenschaftlicher Items und Skalen*.

Cave, D., Albeck-Ripka, L., & Magra, I. (2020). *Huge crowds around the globe march in solidarity against police brutality*. New York Times. https://www.nytimes.com/2020/06/06/world/george-floyd-global-protests.html.

Ciftci, S. (2012). Islamophobia and threat perceptions: Explaining anti-Muslim sentiment in the West. *Journal of Muslim Minority Affairs, 32*(3), 293–309. https://doi.org/10.1080/13602004.2012.727291

Conigrave, J. (2020). *Corx: Create and format correlation matrices*. https://CRAN.R-project.org/package=corx.

Correll, J., Cloutier, J., & Mellinger, C. (2016). Discriminating perception. *Psychological Inquiry, 27*(4), 281–285. https://doi.org/10.1080/1047840X.2016.1215209

Correll, J., Park, B., Judd, C. M., & Wittenbrink, B. (2002). The police officer's dilemma: Using ethnicity to disambiguate potentially threatening individuals. *Journal of Personality and Social Psychology, 83*(6), 1314–1329. https://doi.org/10.1037/0022-3514.83.6.1314

Correll, J., Park, B., Judd, C. M., Wittenbrink, B., Sadler, M. S., & Keesee, T. (2007). Across the thin blue line: Police officers and racial bias in the decision to shoot. *Journal of Personality and Social Psychology, 92*(6), 1006–1023. https://doi.org/10.1037/0022-3514.92.6.1006

Correll, J., Wittenbrink, B., Park, B., Judd, C. M., & Goyle, A. (2011). Dangerous enough: Moderating racial bias with contextual threat cues. *Journal of Experimental Social Psychology, 47*(1), 184–189. https://doi.org/10.1016/j.jesp.2010.08.017

Cox, W. T. L., Devine, P. G., Plant, E. A., & Schwartz, L. L. (2014). Toward a comprehensive understanding of officers' shooting decisions: No simple answers to this complex problem. *Basic and Applied Social Psychology, 36*(4), 356–364. https://doi.org/10.1080/01973533.2014.923312

Decker, O., Kiess, J., & Brähler, E. (Eds.). (2016). *Die enthemmte Mitte: Autoritäre und rechtsextreme Einstellung in Deutschland*. Psychosozial Verlag.

Derous, E., Ryan, A. M., & Nguyen, H.-H. D. (2012). Multiple categorization in resume screening: Examining effects on hiring discrimination against Arab applicants in field and lab settings. *Journal of Organizational Behavior, 33*(4), 544–570. https://doi.org/10.1002/job.769

Dhont, K., Cornelis, I., & van Hiel, A. (2010). Interracial public–police contact: Relationships with police officers' racial and work-related attitudes and behavior. *International Journal of Intercultural Relations, 34*(6), 551–560. https://doi.org/10.1016/j.ijintrel.2010.07.004

Essien, I., Stelter, M., Kalbe, F., Koehler, A., Mangels, J., & Meliß, S. (2017). The shooter bias: Replicating the classic effect and introducing a novel paradigm. *Journal of Experimental Social Psychology, 70*, 41–47. https://doi.org/10.1016/j.jesp.2016.12.009

Fischer, P., Greitemeyer, T., & Kastenmüller, A. (2007). What do we think about muslims? The validity of westerners' implicit theories about the associations between muslims' religiosity, religious identity, aggression potential, and attitudes toward terrorism. *Group Processes & Intergroup Relations, 10*(3), 373–382. https://doi.org/10.1177/1368430207078697

Foroutan, N. (2012). *Muslimbilder in Deutschland: Wahrnehmungen und Ausgrenzungen in der Integrationsdebatte: Expertise im Auftrag der Abteilung Wirtschafts- und Sozialpolitik der Friedrich-Ebert-Stiftung. Gesprächskreis Migration und Integration.*. Abteilung Wirtschafts- und Sozialpolitik der Friedrich-Ebert-Stiftung.

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin, 132*(5), 692–731. https://doi.org/10.1037/0033-2909.132.5.692

Gawronski, B., Brochu, P. M., Sritharan, R., & Strack, F. (2012). Cognitive consistency in prejudice-related belief systems: Integrating old-fashioned, modern, aversive, and implicit forms of prejudice. In B. Gawronski, & F. Strack (Eds.), *Cognitive consistency: A fundamental principle in social cognition* (pp. 369–389). US: Guilford Press.

Goff, P. A., Jackson, M. C., Di Leone, B. A. L., Culotta, C. M., & DiTomasso, N. A. (2014). The essence of innocence: Consequences of dehumanizing black children. *Journal of Personality and Social Psychology, 106*(4), 526–545. https://doi.org/10.1037/a0035663

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

Hall, A. V., Hall, E. V., & Perry, J. L. (2016). Black and blue: Exploring racial bias and law enforcement in the killings of unarmed black male civilians. *American Psychologist, 71*(3), 175–186. https://doi.org/10.1037/a0040109

Hancock, K. J., & Rhodes, G. (2008). Contact, configural coding and the other-race effect in face recognition. *British Journal of Psychology, 99*(1). https://doi.org/10.1348/000712607X199981.

Hope, R. M. (2022). Rmisc: Ryan miscellaneous. https://CRAN.R-project.org/package=Rmisc.

James, L., James, S. M., & Vila, B. J. (2016). The reverse racism effect: Are cops more hesitant to shoot black than white suspects? *Criminology & Public Policy, 15*(2), 457–479. https://doi.org/10.1111/1745-9133.12187

James, L., Vila, B., & Daratha, K. (2013). Results from experimental trials testing participant responses to White, Hispanic and Black suspects in high-fidelity deadly force judgment and decision-making simulations. *Journal of Experimental Criminology, 9*(2), 189–212. https://doi.org/10.1007/s11292-012-9163-y

Kassambara, A. (2022). Rstatix: Pipe-friendly framework for basic statistical tests. https://CRAN.R-project.org/package=rstatix.

Kelley, K. (2022). MBESS: The MBESS r package. https://CRAN.R-project.org/package=MBESS.

Kemme, S., Essien, I., & Stelter, M. (2020). Antimuslimische Einstellungen in der Polizei? Der zusammenhang von Kontakthäufigkeit und -qualität mit vorurteilen und stereotypen gegenüber muslimen. *Monatsschrift für Kriminologie Und Strafrechtsreform, 103*(2), 129–149. https://doi.org/10.1515/mks-2020-2048

Kemme, S., Hanslmaier, M., & Abdul-Rahman, L. (2021). Effects of selection and socialization within the police force on officers' attitudes toward punishment and authoritarianism: A comparison between police students and the general population. *Police Practice and Research, 1–21*. https://doi.org/10.1080/15614263.2021.1921582

Koopmans, R., Veit, S., & Yemane, R. (2019). Taste or statistics? A correspondence study of ethnic, racial and religious labour market discrimination in Germany. *Ethnic and Racial Studies, 42*(16), 233–252. https://doi.org/10.1080/01419870.2019.1654114

Kteily, N., Bruneau, E., Waytz, A., & Cotterill, S. (2015). The ascent of man: Theoretical and empirical evidence for blatant dehumanization. *Journal of Personality and Social Psychology, 109*(5), 901–931. https://doi.org/10.1037/pspp0000048

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the radboud faces database. *Cognition & Emotion, 24*(8), 1377–1388. https://doi.org/10.1080/02699930903485076

Lawrence, M. A. (2016). *Ez: Easy analysis and visualization of factorial experiments*. https://CRAN.R-project.org/package=ez.

Lima, M. E. O., de Araujo, C. L., & Poderoso, E. S. (2018). The decision to shoot black suspects in Brazil: The police officer's dilemma. *Race and Social Problems, 10*(2), 101–112. https://doi.org/10.1007/s12552-018-9225-5

Long, J. A. (2022). *Jtools: Analysis and presentation of social scientific data*.

Lüdecke, D. (2022). sjPlot: Data visualization for statistics in social science. https://CRAN.R-project.org/package=sjPlot.

Ma, D. S., Correll, J., Wittenbrink, B., Bar-Anan, Y., Sriram, N., & Nosek, B. A. (2013). When fatigue turns deadly: The association between fatigue and racial bias in the decision to shoot. *Basic and Applied Social Psychology, 35*(6), 515–524. https://doi.org/10.1080/01973533.2013.840630

Mackenzie, I. G., & Dudschig, C. (2022). psychReport: Reproducible reports in psychology. https://CRAN.R-project.org/package=psychReport.

Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide*. Taylor & Francis.

Mazziotta, A., Zerr, M., & Rohmann, A. (2015). The effects of multiple stigmas on discrimination in the German housing market. *Social Psychology, 46*(6), 325–334. https://doi.org/10.1027/1864-9335/a000249

Mekawi, Y., & Bresin, K. (2015). Is the evidence from racial bias shooting task studies a smoking gun? Results from a meta-analysis. *Journal of Experimental Social Psychology, 61*, 120–130. https://doi.org/10.1016/j.jesp.2015.08.002

Müller, D. (2005). Die Darstellung ethnischer Minderheiten in deutschen Medien. In R. Geißler, & H. Pöttker (Eds.), *assenmedien und die Integration ethnischer Minderheiten in Deutschland: Problemaufriss, Forschungsstand, Bibliographie* (pp. 87–126). Transcript Verlag.

Payne, B. K., & Correll, J. (). Race, weapons, and the perception of threat. In B. Gawronski (Ed.), *Advances in experimental social psychology* (Vol. 62, pp. 1–50). Academic Press is an imprint of Elsevier. https://doi.org/10.1016/bs.aesp.2020.04.001.

Plant, E. A., & Peruche, B. M. (2005). The consequences of race for police officers' responses to criminal suspects. *Psychological Science, 16*(3), 180–183. https://doi.org/10.1111/j.0956-7976.2005.00800.x

Plant, E. A., Peruche, B. M., & Butz, D. A. (2005). Eliminating automatic racial bias: Making race non-diagnostic for responses to criminal suspects. *Journal of Experimental Social Psychology, 41*(2), 141–156. https://doi.org/10.1016/j.jesp.2004.07.004

Pratto, F., Sidanius, J., & Levin, S. (2006). Social dominance theory and the dynamics of intergroup relations: Taking stock and looking forward. *European Review of Social Psychology, 17*(1), 271–320. https://doi.org/10.1080/10463280601055772

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/.

Reimer, N. K., Becker, J. C., Benz, A., Christ, O., Dhont, K., Klocke, U., Neji, S., Rychlowska, M., Schmid, K., & Hewstone, M. (2017). Intergroup contact and social change: Implications of negative and positive contact for collective action in advantaged and disadvantaged groups. *Personality and Social Psychology Bulletin, 43*(1), 121–136. https://doi.org/10.1177/0146167216676478

Revelle, W. (2022). *Psych: Procedures for psychological, psychometric, and personality research*. Northwestern University. https://CRAN.R-project.org/package=psych.

Sadler, M. S., Correll, J., Park, B., & Judd, C. M. (2012). The world is not black and white: Racial bias in the decision to shoot in a multiethnic context. *Journal of Social Issues, 68*(2), 286–313. https://doi.org/10.1111/j.1540-4560.2012.01749.x

Safronova, V. (2020, October 4). Black Germans say it's time to look inward. *New York Times*. https://www.nytimes.com/2020/10/04/style/black-germans-say-its-time-to-look-inward.html.

Savelkoul, M., Scheepers, P., van der Veld, W., & Hagendoorn, L. (2012). Comparing levels of anti-muslim attitudes across western countries. *Quality & Quantity, 46*(5), 1617–1624. https://doi.org/10.1007/s11135-011-9470-9

Shooman, Y. (2012). Das Zusammenspiel von Kultur, Religion, Ethnizität und Geschlecht im antimuslimischen Rassismus. In *Aus Politik und Zeitgeschichte* (pp. 53–57). https://CRAN.R-project.org/package=ez http://www.bpb.de/shop/zeitschriften/apuz/126515/ungleichheit-ungleichwertigkeit.

Sidanius, J., Mitchell, M., Haley, H., & Navarrete, C. D. (2006). Support for harsh criminal sanctions and criminal justice beliefs: A social dominance perspective. *Social Justice Research, 19*(4), 433–449. https://doi.org/10.1007/s11211-006-0026-4

Sidanius, J., & Pratto, F. (1999). *Social dominance: An intergroup theory of social hierarchy and oppression*. Cambridge University Press.

Sim, J. J., Correll, J., & Sadler, M. S. (2013). Understanding police and expert performance: When training attenuates (vs. Exacerbates) stereotypic bias in the decision to shoot. *Personality and Social Psychology Bulletin, 39*(3), 291–304. https://doi.org/10.1177/0146167212473157

Spruyt, B., & van der Noll, J. (2016). The "islamized stranger": On "chronic" versus "contextual" salience in the measurement of anti-Muslim prejudice. *Political Psychology*. https://doi.org/10.1111/pops.12383

Stürmer, S., Rohmann, A., Froehlich, L., & van der Noll, J. (2019). Muslim immigration, critical events, and the seeds of majority members' support for radical responses: An interactionist perspective. *Personality and Social Psychology Bulletin, 45*(1), 133–145. https://doi.org/10.1177/0146167218780989

Swencionis, J. K., & Goff, P. A. (2017). The psychological science of racial bias and policing. *Psychology, Public Policy, and Law, 23*(4), 398–409. https://doi.org/10.1037/law0000130

Swencionis, J. K., Pouget, E. R., & Goff, P. A. (2021). Supporting social hierarchy is associated with white police officers' use of force. *Proceedings of the National Academy of Sciences of the United States of America, 118*(18), Article e2007693118. https://doi.org/10.1073/pnas.2007693118

Tjaden, J. D., Schwemmer, C., & Khadjavi, M. (2018). Ride with me—Ethnic discrimination, social markets, and the sharing economy. *European Sociological Review, 34*(4), 418–432. https://doi.org/10.1093/esr/jcy024

Torchiano, M. (2020). *Effsize: Efficient effect size computation*. https://doi.org/10.5281/zenodo.1480624

van der Schalk, J., Hawk, S. T., Fischer, A. H., & Doosje, B. (2011). Moving faces, looking places: Validation of the Amsterdam dynamic facial expression set (ADFES). *Emotion, 11*(4), 907–920. https://doi.org/10.1037/a0023853

Weichselbaumer, D. (2016). Discrimination against female migrants wearing headscarves. In *IZA DP No. 10217* (pp. 1–27).

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V.Yutani, H., … (2019). Welcome to the tidyverse. *Journal of Open Source Software, 4*(43), 1686. https://doi.org/10.21105/joss.01686

Wigger, I. (2019). Anti-muslim racism and the racialisation of sexual violence: "Intersectional stereotyping" in mass media representations of male muslim migrants in Germany. *Culture and Religion, 20*(3), 248–271. https://doi.org/10.1080/14755610.2019.1658609

Wilke, C. O. (2020). Cowplot: Streamlined plot theme and plot annotations for 'ggplot2'. https://CRAN.R-project.org/package=cowplot.

Zick, A., & Küpper, B. (Eds.). (2021). *Die geforderte Mitte: Rechtsextreme und demokratiegefährdende Einstellungen in Deutschland 2020/21*. DIETZ.

Zick, A., Küpper, B., & Berghan, W. (Eds.). (2019). *Verlorene Mitte – Feindselige Zustände: Rechtsextreme Einstellungen in Deutschland 2018*. DIETZ.

Zick, A., Küpper, B., & Krause, D. (2016). *Gespaltene Mitte - feindselige Zustände: Rechtsextreme Einstellungen in Deutschland 2016*. Friedrich-Ebert-Stiftung.