Published in final edited form as:
Soc Neurosci. 2010 ; 5(2): 201–220. doi:10.1080/17470910903268931.

Decisions to Shoot in a Weapon Identification Task: The Influence of Cultural Stereotypes and Perceived Threat on False Positive Errors

Kevin K. Fleming,
Norwich University

Carole L. Bandy, and
Norwich University

Matthew O. Kimble
Middlebury College

Abstract

The decision to shoot engages executive control processes that can be biased by cultural stereotypes and perceived threat. The neural locus of the decision to shoot is likely to be found in the anterior cingulate cortex (ACC) where cognition and affect converge. Male military cadets at Norwich University (N=37) performed a weapon identification task in which they made rapid decisions to shoot when images of guns appeared briefly on a computer screen. Reaction times, error rates, and EEG activity were recorded. Cadets reacted more quickly and accurately when guns were primed by images of middle-eastern males wearing traditional clothing. However, cadets also made more false positive errors when tools were primed by these images. Error-related negativity (ERN) was measured for each response. Deeper ERN's were found in the medial-frontal cortex following false positive responses. Cadets who made fewer errors also produced deeper ERN's, indicating stronger executive control. Pupil size was used to measure autonomic arousal related to perceived threat. Images of middle-eastern males in traditional clothing produced larger pupil sizes. An image of Osama bin Laden induced the largest pupil size, as would be predicted for the exemplar of Middle East terrorism. Cadets who showed greater increases in pupil size also made more false positive errors. Regression analyses were performed to evaluate predictions based on current models of perceived threat, stereotype activation, and cognitive control. Measures of pupil size (perceived threat) and ERN (cognitive control) explained significant proportions of the variance in false positive errors to middle-eastern males in traditional clothing, while measures of reaction time, signal detection response bias, and stimulus discriminability explained most of the remaining variance.

Keywords

Weapon Bias; Stereotypes; Threat Perception; Event-Related Potentials; Pupil Size; Anterior Cingulate Cortex

---

Correspondence can be sent to: Kevin K. Fleming, Ph.D., Department of Psychology, Norwich University, 158 Harmon Drive, Northfield, VT 05663, Phone: (802) 485-2428, Fax: (802) 485-2252, kfleming@norwich.edu.

American and British military personnel are presently engaged in conflict in the Middle East, where they routinely observe a racially and culturally diverse population and must appraise and respond to perceived threats of violence and terrorism. The decision to shoot at a suspected terrorist is extremely complex and not taken lightly, but often these decisions must be made rapidly and under conditions of intense stress. What social, cognitive, affective, and even neural factors influence the decision to shoot? Bringing this phenomenon into the laboratory presents many challenges, but understanding the mechanisms underlying the decision to shoot can lead to better training protocols for military personnel and fewer casualties due to misidentification of threat or friendly fire.

Recent studies have focused on police officers and their decisions to shoot at crime suspects (Payne, 2001, 2005, 2006; Correll, Park, Judd, & Wittenbrink, 2002; Correll, Urland, & Ito, 2006). The high-profile shooting death of Amadou Diallo by four police officers in NYC called attention to potential racial bias in police officers' decisions to shoot. Using variations of a weapon identification task, researchers have found that racial primes such as black males elicit more false positive identification errors than white males under conditions of time pressure (Payne, 2001). Police officers were found to be better overall at weapon identification, and exhibited less racial bias than community members (Correll, Park, Judd, Wittenbrink, Sadler, & Keesee, 2007). More importantly, training protocols for police officers have been effective at reducing racial biases in the decision to shoot (Plant, Peruche, & Butz, 2005). In the domain of police officer training, this research has been pivotal toward our understanding of racial bias, threat perception, and cognitive control in the decision to shoot.

Police officers are trained to respond with deadly force if they perceive a threat to their own life or others. Misperceiving threat can lead to tragic deaths like that of Amadou Diallo. In the military, soldiers are also trained to respond with deadly force, but despite a perceived threat to their life, they cannot shoot until given an order by an officer. This additional constraint makes the response of the soldier more complex. Nonetheless misperceptions of threat and orders to shoot sometimes lead to casualties. According to information obtained by Army Times under the Freedom of Information Act (Tan, 2006), at least 16 American soldiers have died as a result of friendly fire during operations in Iraq and Afghanistan. Likewise, there have been many documented incidents of innocent civilians killed by soldiers who misperceived their potential threat. In combat, like police duty, the capacity to correctly appraise and respond to potential threat is a critical skill. In the military domain, it is also likely that racial and cultural biases are prevalent, particularly concerning middle-eastern males.

Our goal in the present study is to examine how stereotypes toward middle-eastern males influence the decisions of military cadets to shoot under controlled conditions. Soldiers face conflicts in many different parts of the world, and models of weapon bias effects must be generalizable to a wide range of racial and ethnic groups. To be applicable in the military domain, models of weapon bias must also include mechanisms for perceived threat and cognitive control. Soldiers must not only perceive threat accurately, but they must also be able to control their response until given a clear order to shoot.

Much of the research on weapon bias has focused on blacks and whites. Current models will be reviewed both in terms of their fit for explaining bias toward blacks, but also their fit for explaining bias toward other outgroups. To the extent that all bias is the same, then middle-eastern males will elicit the same bias as blacks. However, racial and cultural biases may not all be the same. Different sets of variables may predict different types of bias. Models are critical heuristics for explaining existing data, but they also need to predict future data and generalize to different racial and cultural groups. In the end, predictive models are essential for developing effective training protocols for police officers and military soldiers.

Dual process models have prevailed in explaining a diverse set of empirical findings from the allocation of attentional capacity (Kahneman, 1973), to semantic priming (Neely, 1977), and stereotyping (Devine, 1989). These models posit automatic and controlled processes. Automatic processes are fast, involuntary, uncontrolled, unconscious, unintentional, and are subject to implicit task demands. Controlled processes are slow, voluntary, conscious, deliberate, intentional, and subject to explicit task demands. In his original study of the weapon identification task, Payne (2001) interpreted his data in terms of a dual process model in which weapon bias is a function of both automatic and controlled processes. Payne (2005) further clarified the role of these dual processes in weapon bias by suggesting that an automatic impulse can trigger a biased response, but that cognitive control can be exerted to inhibit this biased response. Cognitive control can act as a brake on automatic, unintended stereotyping, and race bias.

Studies by Correll and colleagues (Correll, Park, Judd, and Wittenbrink, 2002; Correll, Urland, & Ito, 2006; Correll, Park, Judd, Wittenbrink, Sadler, & Keesee, 2007) further identified perceived threat as one of the driving factors behind an automatic impulse toward a biased response. Because whites represent a low threat to white participants, the automatic tendency will be to inhibit the shoot response. This model explains the finding that white participants will often miss a gun held by an armed white male. In addition, because blacks represent a high threat, the automatic tendency will be to facilitate the shoot response. This ‘shooter bias’ explains the finding that higher false positive errors occur when white participants are faced with an unarmed black male. Perceived threat and subsequent ‘shooter bias’ also explains why white participants respond faster and more accurately to armed black males.

Further studies by Amodio and colleagues (Amodio, Harmon-Jones, Devine, Curtin, Hartley, & Covert, 2004; Amodio, Devine, & Harmon-Jones, 2008) have replicated the weapon bias effect and shown that this effect is mediated by individual differences in the ability to control prejudice. Participants with less ability to control prejudice produce more false positive errors in the weapon identification task. Conversely, participants with greater ability to control their responses, produce fewer false positive errors. These findings indicate that not every impulse toward an automatic bias will be intercepted by cognitive control mechanisms, leaving room in the dual process model for weapon bias errors despite attempts for cognitive control.

The success of dual process models, however, has not prevented the development of more complex and detailed models. The quadruple process model (Quad model) developed by

Sherman and his colleagues (Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005) has been specifically applied to the weapon identification task (Sherman, Gawronsky, Gonsalkorale, Hugenberg, Allen, & Groom, 2008). The Quad model has both automatic activation and cognitive control mechanisms just like standard dual process models. In addition, task-related demands are integrated into the model with the inclusion of processes such as the ability to detect the correct response and a guessing bias if the correct response has not been detected.

The Quad model can be applied to the weapon identification task as a cascade of events or processes. Initially, a stereotype association is either activated or not activated by the face prime (denoted as AC in the model). When the gun or tool target is presented, the correct response is either detected or not detected (denoted as D in the model). If a stereotype has been activated by the face prime that is incongruent with the correct response (e.g., black/tool), then cognitive control mechanisms must either overcome bias (OB) due to the stereotype activation (e.g., shoot), or an error will be produced (e.g., a false positive error). If a stereotype is not activated and the correct response is not detected, then a guessing bias (G) toward the gun or tool response will determine the response outcome. Sherman et al. (2008) suggest that the Quad model synthesizes all of the components of the dual-process model, while providing a more detailed and nuanced description of human behavior in the weapon identification task, as well as other tasks where responses are conflicted between an automatic response and a controlled response. Errors are expected in these conflict tasks, but not all errors are due to racial or cultural biases. The Quad model offers the possibility of explaining errors that are due to task-related demands under time pressure such as stimulus discriminability, guessing bias, and failure to detect the correct response.

While the present study was not designed as a test of dual or quadruple process models, several key variables from these models were chosen as measurements of stereotype activation, cognitive control, and task-related demands. These variables were then used to predict weapon identification errors committed by military cadets when primed by middle-eastern males. Larger questions that the present research seeks to address include: How do military personnel respond in a weapon identification task? Do they have the same response tendencies as police officers? Do they have biases toward middle-eastern males that mirror the biases toward blacks held by police officers? Is clothing a mitigating factor when considering the threat posed by middle-eastern males? Is the perceived threat posed by figures such as Osama bin Laden predictive of false positive errors when primed by middle-eastern males wearing traditional clothing? Does this perceived threat generalize to middle-eastern males wearing western clothing? Do military personnel have the same level of cognitive control over their shoot responses as police officers? Is training a factor in building greater cogntive control? How do these findings generalize to other racial and ethnic groups under wartime threat? These are some of the most challenging questions posed by the present research. Answering these questions will help us understand how military personnel perceive and respond to threat, and how they make decisions to shoot at potentially armed enemy combatants under both stress and time pressure.

Participants in the present study were male, predominantly white, military cadets enrolled at Norwich University, the oldest private military academy in the United States and the birthplace of ROTC. Students come to Norwich for both academic and military training. They are not military personnel in the strict sense, but they often intend to commission into the military upon graduation. To our knowledge, they do not receive any special training related to identifying middle-eastern terror suspects, nor are they specifically trained to shoot them. However, they are very aware of the possibility that they may be called to duty in Iraq or Afghanistan and may be placed in highly volatile situations where they may have to decide to shoot at a potential enemy combatant. Like police officers in training, military cadets are sensitive to the inherent dangers of the job they hope to do.

The weapon identification task presents an ideal methodology for testing how cultural stereotypes, perceived threat, and cognitive control might influence decisions to shoot. The task also has good face validity for cadets because they can easily imagine situations in which they must decide to shoot when they see a weapon. Furthermore, they have never been in a combat situation, so their responses will be naive.

After September 11, 2001, it was feared that there would be a backlash of violence against American citizens of middle-eastern descent. This scenario, fortunately, has not come to pass. But this does not mean that a negative stereotype has not emerged toward middle-eastern males. The events of 9/11 were sudden and dramatic, creating an ideal opportunity to study stereotype development toward this generally unknown ethnic group. Our participants were between 10 and 14 years old when the attacks by Osama bin Laden were launched, and it is likely that they hold both explicit and implicit biases against middle-eastern males wearing robes and turbans. Most Americans see Osama bin Laden as the exemplar of middle-eastern terrorism, but it is also well-known that the hijackers who boarded the airliners on 9/11 wore western clothing, which is incongruent with their cultural stereotype. For this reason, we manipulated the type of clothing worn by the middle-eastern males in our stimuli. Fiske and Neuberg (1990) have suggested that stereotype incongruent information may diminish the negative affect associated with stimuli and thereby reduce bias.

In the present weapon identification task, targets included guns and tools. Cadets were told to ‘shoot’ when they saw a gun. Foils included tools of various types, including a drill and hair dryer to increase confusability. The gun and tool targets were preceded by face primes from different racial and cultural groups, including white, black, and middle-eastern males. The middle-eastern males were further divided between those wearing western clothing and those wearing traditional clothing such as robes and headgear. Reaction times and errors were recorded on each trial. Based on racial biases observed in prior studies (Greenwald, Oakes, & Hoffman, 2003), it was predicted that military cadets would make more false positive errors when primed by images of black males relative to white males. Furthermore, cultural stereotypes for middle-eastern males would predict more false positive errors when primed by images of middle-eastern males wearing traditional clothing relative to western clothing.

Results such as these would validate Fiske and Neuberg (1990) to the extent that cultural stereotypes are contingent upon cues such as clothing which moderate the effects of bias, especially if the clothing is incongruent with the stereotype. Our prediction was that cadets would share a certain cultural stereotype of middle-eastern males wearing turbans and robes, and associate that stereotype with middle-eastern terrorism. This association would in turn lead to a perception of threat and a tendency toward the shoot response. In these cases, we also predicted a greater need for cognitive control in order to overcome bias due to this activation of threat and give a correct response, as predicted by both the Quad and dual process models.

But how can perceived threat and cognitive control be measured? Correll et al. (2006) used electrophysiological measures of P200 and N200 event-related potentials (ERP’s) to measure threat perception and control processes in the decision of police officers to shoot at white and black suspects. They found that the P200 component was positively predictive of errors, indicating higher threat perception, and they found that the N200 component was negatively predictive of errors, indicating less cognitive control over the decision to shoot when faced with black suspects. Prior research by Ito, Thompson, and Cacioppo (2004), also examined a wide range of ERP components involved in face perception and the activation of race-based information in order to understand the time-course of stereotype activation. Measures that predicted race-based differences included the face-specific VPP/N170 component, the N200, and the LPP, or late positive potential. These results indicated that ERP measures could be used to examine subtle differences in brain responses to images of different racial groups.

To gain an understanding of the control processes involved in weapon identification and the decision to shoot, event-related potentials were recorded for every trial. We focused on the error-related negativity (ERN) waveform that manifests in medial frontal electrodes about 80 ms after an error has been committed. The ERN has been used previously by Amodio and his colleagues (2004, 2008) and has proven to be a sensitive measure of cognitive control. Payne (2005) suggests that the ERN is the most direct measure of cognitive control currently available. The value of the ERN is that it is believed to measure activity in the anterior cingulate cortex (ACC), an area known for its role in executive control during tasks involving arousal and conflicting information (Luu & Pedersen, 2004). Furthermore, the ACC can be subdivided into dorsal and ventral/rostral areas (Bush, 2004). The dorsal (dACC) area is associated with cognitive control, and the rostral (rACC) area is associated with emotional responses (Critchley, Tang, Glaser, Butterworth, & Dolan, 2007). This means that the ACC is where affect and cognition converge (Allman, Hakeem, Erwin, Nimchinsky, & Hof, 2001).

As further support for the use of the ERN, the weapon identification task represents the type of conflict task in which ERN’s are most often measured (Ridderinkhof, 2002; Ridderinkhof, van den Wildenberg, Wijnen, & Burle, 2004). Participants make mistakes for a variety of reasons. They are often under stress and time pressure and they must respond even before they fully process all relevant information. The ACC receives input from the emotional centers of the brain such as the amygdala. If the amygdala signals a negative emotion based on the race/ethnicity of a face prime, then the ACC may be conflicted about how to respond when the target is a tool instead of a weapon. Under time pressure, the emotional information may ‘beat’ the more time-consuming weapon identification processes and the participant may ‘shoot’ the tool erroneously. ERN’s manifest when a participant makes an error, and specifically when the participant detects the error immediately after the response has been initiated. The ERN, therefore, will be greatest in those participants who are self-monitoring their responses (Scheffers & Coles, 2000). These individuals may be less prone to weapon identification errors because they are more carefully monitoring their responses. Amodio et al. (2004) found a correlation between the ERN and false positive error rates in the weapon identification task. Participants with deeper ERN's made fewer errors. We would like to further understand how the ERN correlates with biases toward middle-eastern males. We predict that greater control will be necessary for middle-eastern males wearing traditional robes and turbans because this group is more closely associated with terrorism.

Finally, to gain a better understanding of the influence of automatic stereotype activation and perceived threat, we chose a novel measurement: pupil size. By showing images of racial and ethnic groups featuring the face and upper portions of clothing in a free-viewing paradigm while recording eye movements and pupil size, we hoped to capture the autonomic arousal generated by these images. Tracking of eye movements gives us a window onto the perception of threat as measured by ongoing changes in the actual size of the pupils as participants scan images. Pupil size is a largely untapped measure of autonomic arousal associated with increased sympathetic activity (Bradley, Miccoli, Escrig, & Lang, 2008). Because autonomic arousal is a concurrent event in the activation of threat based on a stereotype, and because the amygdala processes emotional information and feeds that information to the ACC (Adolphs, 2006), pupil size provides an excellent index of the arousal generated by faces. Pupil size measurements will also be critical for determining which features of the face and clothing produce the greatest arousal when viewing images of middle-eastern males wearing either traditional or western clothing. Using these pupil measures we may gain insight into the specific features of the face and clothing that differentiate between these racial and ethnic groups. Soldiers and cadets are trained to carefully examine details of the face and clothing, especially when these details discriminate allies from potential enemy combatants.

Our hypothesis was that cadets would show the greatest arousal to images of middle-eastern males, particularly Osama bin Laden, and that they would attend mostly to clothing features that differentiate traditional middle-eastern males from western males. In addition, we predicted that the degree of pupil dilation would correlate positively with false positive errors in the weapon identification task. If dilation of the pupils to middle-eastern men wearing robes and turbans indicates perception of threat, then more false positive errors would be predicted for cadets who show the greatest pupil dilation to these images. Finally, by including in our sample an image of Osama bin Laden, the exemplar of Middle East terrorism, we predicted that the greatest pupil dilation and concurrent arousal would be shown to this image and a clear linkage between perceived threat and autonomic arousal could be drawn.

Method

Participants

Military Cadets at Norwich University served as participants in this study. They were students enrolled in Introductory Psychology courses and their participation was voluntary. Some received extra credit from their instructors. Altogether, 37 male Cadets took part in all phases of this research: the weapon identification task, the EEG recording, and the free-viewing eye-tracking study. Four participants were removed from the behavioral data analyses due to an insufficient number of speeded responses within the allotted time. In addition, five participants failed to calibrate on the eye-tracking system and were excluded from the pupil size analyses. Calibration difficulties arise most often from corrective eye-glasses and/or excessive eye-movements or eye-blinks. Otherwise, all participants had normal or corrected-to-normal vision.

Stimuli

The face primes consisted of 20 images representing the targeted racial and ethnic groups. All images depicted males and contained the face and upper portions of the clothing. There were 5 images of whites, 5 images of blacks, 5 images of middle-eastern men wearing traditional clothing, and 5 images of middle-eastern men wearing western clothing. Traditional clothing was defined as robes, turbans, and other clothing or headgear appropriate to middle-eastern men from ethnic groups in Iraq or Afghanistan. Western clothing was defined as a business suit, collared polo shirt, or other attire found in American or Western cultures. All images were equated for attractiveness and likability by independent raters. Sample images are presented in Figure 1.

Images of guns and tools were also obtained. The 10 images of weapons depicted handguns with varying orientations. The 10 images of tools consisted of common household tools of various types: wire cutter, hedge trimmer, garden shovel, hedge clippers, hammer, cordless drill, spade, hair dryer, garden fork, and shears. Some images, such as the cordless drill and the hair dryer, were similar in shape to the handguns to increase confusability. All images of guns and tools were converted to black-and-white to control for color-based cues. Sample images of guns and tools are presented in Figure 2.

For the eye-tracking portion of the study, a subset of 14 faces were selected, 3 black, 3 white, 4 middle-eastern in traditional clothing, and 4 middle-eastern in western clothing. One image of Osama bin Laden was also included to anchor the negative stereotype associated with Middle East terrorism (see Figure 3). The image of Osama bin Laden was not shown in the weapon identification task in order to avoid bias due to a known and recognizable terror figure. None of the other middle-eastern males were recognizable, nor were they specifically chosen for being associated with terrorist groups. Those wearing western clothing might easily be viewed as American citizens, and those wearing traditional clothing might just as easily represent ethnic groups allied with American soldiers in the Middle East.

All images of faces in the eye-tracking portion of the study were presented in full-color and sized at approximately 6″ × 8″ for display on the computer screen. In contrast, the images of faces for the weapon identification task were converted to black-and-white using Photoshop to eliminate color-based cues and sized at approximately 2″× 3″ for more rapid display. A pattern mask was also created and sized at 2″× 3″ to fully cover the images of faces, guns, and tools used in the weapon identification task.

Apparatus

For the weapon identification task, a Dell laptop with a 13″ screen was used to present the face, gun, and tool stimuli. Presentation was controlled with STIM^2 software designed for use with the NeuroScan NuAmps 40-channel EEG amplifier. SCAN software was used on a second Dell laptop that served as the EEG recording computer. The two laptops were linked to the NeuroScan amplifier and stimulus triggers were sent via USB cables from the STIM^2 computer to the SCAN computer for trial initialization and identification. Behavioral responses were recorded from a two-button mouse attached to the STIM^2 computer and responses were recorded and coded for reaction time (in milliseconds) and accuracy. Scalp-recorded electrical activity was captured using a NeuroScan 32-channel cap with sintered Ag/AgCl electrodes. Conductivity was achieved using electrolyte-saturated sponges. An averaged reference was used to compute the electrical charge (in microVolts) at a sampling rate of 250 Hz. EDIT software from NeuroScan was used to epoch and to average the ERP data for subsequent ERN analysis using SPSS.

For the eye-tracking study, eye fixations and pupil size were recorded using an Eye-Link 1000 eye-tracking system from SR Research. A Dell tower computer was equipped with the Experiment Builder software from SR Research and it controlled stimulus presentation on a 19″ flat panel display. A second Dell tower computer was connected to the high-speed infrared camera that tracked the eye movements using a ratio of corneal reflection and pupil area to lock in the coordinates of the eye fixation every millisecond. This computer and related software was used to calibrate and validate tracking of the right eye for all participants. All data was recorded using DataViewer software from SR Research and analyzed using SPSS. No responses were required in the free-viewing paradigm.

Procedure

There were two phases in the present study. The first phase consisted of the weapon identification task, and the second phase consisted of the free-viewing eye-tracking task. The EEG system and the eye-tracking system were in separate, but adjacent rooms. Participants were first directed to the EEG system and the 32-electrode cap was placed on their head securely before inserting electrolyte solution into the QuikCell sponges. Once impedances for all electrodes were reduced to less than 5 kΩ, participants were seated before the computer screen and the weapon identification task was explained. They were told that they would see guns and tools presented rapidly on the computer screen. Their task was to ‘shoot’ by pressing the left mouse button if they saw a gun or, if they saw a tool, press the right mouse button to indicate a ‘no shoot’ response.

Each gun or tool target was preceded by a face prime consisting of one of the four image types (white, black, middle-eastern with western clothing, or middle-eastern with traditional clothing). Altogether, 120 trials were created: 40 trials contained guns as targets and 80 trials contained tools as targets. Each face primed two guns and four tools respectively. Trials were counterbalanced so that each image type primed each gun or tool target an equal number of times. For example, the hair dryer target appeared 8 times and was preceded by images from each of the four types of face primes two times. The sequence of trials was fixed but random with two tool trials for every gun trial on average.

Face primes were presented for 400 ms followed by a 32 ms pattern mask. A fixation cross appeared for 76 ms prior to the face prime to orient attention. Gun or tool targets were then presented for 32 ms followed by a 32 ms mask. Participants were told to press the left mouse button as quickly as possible if the target was a gun, and to press the right mouse button as quickly as possible if the target was a tool. There was an 800 ms interval between the onset of the target and the onset of the face prime for the next trial. This rapid stimulus timing meant that participants had a window of 724 ms from the onset of the gun/tool to the onset of the fixation cross for the next trial in which to respond. Speeded responses were critical in this task and participants were urged to respond immediately upon seeing the gun or tool target. The entire sequence of trials lasted less than three minutes. Upon completion of the weapon identification task, the electrode cap was removed and participants were debriefed.

The second phase of the study then began. Participants were seated in front of the Eye-Link 1000 eye-tracking system and a chin rest was used to keep their head in a fixed position approximately 24″ in front of the flat panel display. Calibration and validation of the right eye position were performed prior to initiating the free-viewing trials. Of the 37 participants, 5 failed to calibrate properly due to excessive blinks or other eye movements. Upon successful calibration and validation, participants were shown color images of the male faces selected from the previous study. Each image was shown for 10 seconds and no response was required. Participants were told simply to view the images. The entire sequence took less than three minutes to complete, and debriefing was provided immediately after the last trial.

Data Analysis

Behavioral Data—Reaction times in milliseconds were recorded for every valid trial in which the speed of the response fell within the allotted window. Responses longer than 724 ms as well as responses faster than 100 ms were coded as missing (invalid) data. Four of the 37 participants were removed from further analyses due to less than 25% valid responses. The total number of trials across the remaining 33 participants was 3960. Of these trials, 2771 were valid (70%) and 1189 were invalid (30%). Approximately 749 responses (18.8%) were invalid either because participants did not respond at all, or because they responded slower than 724 ms. Approximately 440 responses (11.2%) were invalid because participants responded too quickly. Reaction times for valid responses were averaged and analyzed for statistical differences using repeated-measures ANOVA tests.

Speeded responses were demanded in this task to induce errors. Responses were coded into four different types. Correct detection (hit) responses occurred when the participant made a correct decision to shoot when the target was a gun. Correct rejection responses occurred when the participant made a correct decision not to shoot when the target was a tool. Misses occurred when the participant made an incorrect decision not to shoot when the target was a gun. And most importantly, false positive (false alarm) responses occurred when the participant made an incorrect decision to shoot when the target was a tool.

Frequency counts for each type of response were tallied across participants for each racial and ethnic prime condition. These counts were converted into percentages by dividing the frequency counts by the number of trials of that type. Repeated-measures ANOVA tests were performed across the racial and ethnic prime conditions to test for differences in the false positive error rates and reaction times. In addition, a within-subjects repeated measures ANOVA was conducted combining the miss and false positive responses in order to test for interactions. False positive error rates, both overall and by prime condition, were used as the criterion variables in the multiple regression analyses. Reaction times for the false positive error responses were also used as predictor variables in the multiple regression analyses in order to evaluate the trade-off between speed and accuracy in this task.

Signal Detection Data—In addition to the computed percentages of errors and correct responses, measures of signal detection bias and sensitivity were computed for the hits (correct detections) and false alarms (false positive errors). Proportions of both the hits and false alarms for only the valid responses were converted to z-scores across each of the racial/ethnic face prime conditions. Bias was calculated as c, which represents the sum of the z-scores for hits and false alarms. This sum is then multiplied by a factor of −.5 (see Macmillan & Creelman, 1991). Negative bias scores indicate a more liberal criterion for the gun/shoot response, while positive bias scores indicate a more conservative criterion. Signal detection sensitivity was calculated as d′, which represents the z-scores for hits minus the z-scores for false alarms. Higher sensitivity scores indicate greater accuracy in discriminating the gun from the tool, while lower sensitivity scores indicate less accuracy. Together, these measures provide estimates of the discriminability (d′) of the gun and tool targets, as well as an estimate of the tendency to set a lower criterion (c) for making the gun/shoot response for each participant. Correlations of these measures with pupil dilation and ERN’s were also performed to evaluate the relationships between bias, sensitivity, threat perception, and cognitive control during the decision to shoot.

ERN Data—During performance of the weapon identification task, stimulus and response triggers for each trial were coded and sent to the SCAN computer and marked on the continuous EEG record. The EEG sample rate was 250 Hz with a band-pass width of .1 – 30 Hz. Epochs for each stimulus and response were 1000 ms in length and corrected to baseline using a 100 ms pre-stimulus interval. Face primes were coded for the four different image types, and responses to the targets were coded for the four different response types. To compute the ERN, the continuous EEG files were averaged for each response type and prime type using the response trigger as the starting point for each epoch. The ERN is a negative deflection in the averaged ERP waveform that occurs between 50 and 150 milliseconds after an incorrect response. Note that all 32 electrodes were used during recording, but only the ERN from the Fz electrode, midline frontal, was used for these analyses.

The peak negativity of the ERN was computed for each participant across each of the four response types and compared using paired t-tests. ERN scores were calculated as the difference between the peak negativity for the correct rejection responses and the peak for the false positive responses (FP-CR). Computation of the ERN yielded an overall difference score for each participant. In addition, ERN scores were computed across each of the racial and ethnic primes and compared using a repeated-measures ANOVA. ERN’s for each prime condition were also correlated with false positive errors following these primes and included as predictor variables in multiple regression analyses.

Note that for the white primes, the number of errors was very low, with no errors at all for several participants. This meant that no ERN’s could be computed for these cells. In addition, a minimum of three epochs were used when calculating the ERN’s across conditions. This resulted in a loss of nine participants for the white prime types (n=24). No ERN data was lost for black prime types (n=33) and only one participant was lost for each of the middle-eastern prime types (n=32). This reduction of data must be considered when interpreting the statistical analyses.

Pupil Data—The Eye-Link 1000 system recorded all eye fixation locations, eye fixation durations, and eye movements for each 10-second stimulus duration. Pupil size was measured in arbitrary units (based on the number of activated photocells in the infrared camera) and it was measured every millisecond to maintain calibration of the eye position. This ongoing measure of pupil size is a very sensitive measure of the underlying arousal evoked by the image and it is even sensitive to changes in arousal across eye fixations and interest areas.

The Eye-Link 1000 software allows a programmer to define interest areas within each image. In the images of faces used in the present study, interest areas were defined using geometric shapes circumscribing common facial and clothing features. Interest areas included: left eye, right eye, nose/mouth, chin/beard, hair/turban, and torso/clothing. To capture interest areas not defined above, such as cheeks, ears, and background, an interest area was created to encompass the entire image. Average pupil size was then calculated across all images and all interest areas within those images. Comparisons of pupil size were made using a 5×7 factorial ANOVA for the five different racial/ethnic groups and the seven different interest areas.

In addition, pupil dilation scores were computed for each participant by taking the average pupil size for white faces and subtracting it from the average pupil size for each of the other racial/ethnic groups, as well as Osama bin Laden. Dilation scores represent a measure of differential arousal to potential threat. Comparisons of dilation scores were made using a repeated-measures ANOVA, with the hypothesis that bin Laden would induce the largest pupil dilation relative to white males, followed by middle-eastern males in traditional clothing, western clothing, and black males. These dilation scores were then correlated with false positive errors and entered into a multiple regression analysis with the behavioral data to assess the relationships between arousal, threat perception, bias, sensitivity, and weapon identification errors.

Multiple Regression Analyses—Response data for each participant were compiled in order to perform multiple regression analyses. False positive error rates served as the criterion variables in all of these analyses. Separate regression analyses were performed for each racial and ethnic group. Predictor variables included measures of pupil dilation, ERN scores, reaction times, and measures of signal detection bias and sensitivity. These variables were entered into the regression models in the order listed above. Percentages of variance explained by each variable were computed along with the overall significance of the model. Different models will be obtained if stereotype activation is different for these groups. Multiple regression analyses provide a critical tool for evaluating both dual and quadruple process models of weapon identification errors.

Results and Discussion

Behavioral Data

Error Rates—False positive errors in the weapon identification task indicate an incorrect decision to shoot when the target is a tool. A total of 436 false positive errors were made across the 33 valid participants. Error rates for each racial and ethnic prime condition are presented in Table 1. Using a repeated-measures ANOVA, differences in the percentages of false positive errors across prime conditions were significant, $\underline{F}(3, 96) = 4.414$, $\underline{p} = .006$. As predicted, false positive errors were more frequent when tool targets were primed by images of black males (18.3%) and middle-eastern males in traditional clothing (19.1%). The lowest number of false positive errors occurred when tools were primed by images of white males (12.3%) while a moderate number of errors occurred for middle-eastern males in western clothing (16.4%).

Post hoc tests revealed that the differences between white and black primes were significant, $\underline{t}\,(32) = 2.60$, $\underline{p} = .014$. Similarly, the differences between white and traditional middle-eastern primes were significant, $\underline{t}\,(32) = 3.623$, $\underline{p} = .001$. Traditional middle-eastern primes also elicited more false positive errors than middle-eastern primes wearing western clothing, as predicted by Fiske & Neuberg (1990), however, this difference was not significant, $\underline{t}\,(32) = 1.375$, $\underline{p} = .179$.

Amodio et al. (2004) found an interaction between miss rates and false positive rates. High false positive rates were associated with low miss rates. A within-subjects factorial repeated measures ANOVA was used to test for this interaction and it was found to be significant, $\underline{F}(3, 96) = 4.209$, $\underline{p} = .008$. Examining the miss rates in Table 1, we can see that higher false positive error rates were associated with lower miss rates. Because the miss rates were complementary to the hit rates, examination of these rates also reveals that higher false positive error rates were found in conditions with both lower miss rates and higher hit rates. These results are also consistent with prior findings by Payne (2001) and Correll et al. (2002, 2006, 2007).

Reaction Times—There were no significant differences in the reaction times for the false positive errors across the racial and ethnic groups, $\underline{F}\,(3, 54) = .789$, $\underline{p} = .505$. Reaction times were also not significantly different for the correct rejection responses, $\underline{F}\,(3, 93) = 1.578$, $\underline{p} = .200$. Reaction times, however, for the miss responses were significantly different, $\underline{F}\,(3, 69) = 4.114$, $\underline{p} = .010$, with faster responses to white and middle-eastern males in western clothing. This result is consistent with the hypothesis that western clothing makes middle-eastern males more likely to escape detection when armed, as predicted by Fiske and Neuberg (1990). Furthermore, the reaction times for the correct detection responses were significantly higher for the black and middle-eastern males in traditional clothing, $F(3, 87) = 5.318$, $p = .002$, a result that is consistent with the hypothesis that an automatic stereotype bias toward the shoot response is beneficial when black and traditional middle-eastern males are paired with guns.

In summary, participants were biased toward a decision to shoot when primed by images of blacks and traditional middle-eastern males, and this bias had a positive benefit leading to more correct detections of the gun and fewer misses. On the negative side, however, this bias led to more false positive errors and fewer correct rejections for the tools. The speed of the response was also influenced by this priming effect, with faster decisions to shoot when primed by images of blacks and middle-eastern males in traditional clothing and slower decisions to shoot when primed by images of whites and middle-eastern males in western clothing. This trade-off between speed and accuracy is exactly what would be expected. The correlation between false positive reaction times and error rates was significant and negative, $r\,(33) = -.361$, $p = .039$. Bias in the decision to shoot leads to faster and more accurate shoot responses to guns, but it also leads to faster and erroneous shoot responses to tools.

Signal Detection Data—The converted z-scores for each racial and ethnic prime condition along with signal detection bias (c) and sensitivity (d') measures are shown in Table 2. The most liberal criterion was observed for the middle-eastern males in traditional clothing ($c = -0.02$), a result that was only marginally negative. Similarly, the criterion for black males was marginally positive ($c = 0.07$). In contrast, the bias scores were more positive for both the white and western-clad middle-eastern males ($c = 0.78$ and 0.51 respectively), indicating more conservative criteria for making a ‘shoot’ response. These measures of bias were significantly different across all of the racial and ethnic conditions, $F\,(3, 96) = 6.334$, $p < .001$.

Sensitivity, as measured in signal detection, is independent of bias, and it reflects the discriminability of the stimuli. Guns and tools were most discriminable following white primes ($d' = 2.19$) and least discriminable following black primes ($d' = 0.98$), a result that may explain some of the false positive errors following black primes. Discriminablility was only slightly higher for traditional middle-eastern primes ($d' = 1.45$) relative to western-clad middle-eastern primes ($d' = 1.24$). These differences in sensitivity across the racial and ethnic primes were not significant, $F\,(3, 96) = 1.859$, $p = .140$.

ERN Data

When soldiers make a decision to shoot, it is reasonable to ask what is happening in their brains, especially when speed is required and there is conflicting information about whether or not to shoot. Error-related negativity (ERN) is found predominantly in the medial frontal areas of the brain associated with the anterior cingulate cortex (ACC). The ACC is often found to regulate responses in forced choice decision tasks involving conflicting information. ERN’s typically manifest as negative waveforms peaking at approximately 80 ms following an erroneous response (Holroyd, Nieuwenhuis, Mars, & Coles, 2004). The ERN's recorded in the present study manifested peak negativity in the Fz electrode at 67 ms post-response.

Two analyses were performed on the ERN data from the Fz electrode. First, the peak negativity for the correct rejection responses was compared to the peak negativity following the false positive responses using a paired t-test. This difference was significant, $\underline{t}$ (36) = 3.975, $\underline{p}$< .001. Deeper ERN's manifested in the region of the Fz electrode following false positive responses (−3.272 μV) relative to correct rejection responses (−0.777 μV). Second, the peak negativity for the miss responses (−2.914 μV) was compared to the peak negativity for the correct detection responses (−2.598 μV). As expected, this difference was not significant, $\underline{t}$ (36) = 0.321, $\underline{p}$ = .750. ERN's do not occur for misses because these are errors of omission, rather than commission, and that seems to make a difference. The data support a general view that errors of commission are monitored and corrected in regions of the brain associated with the ACC.

To test the hypothesis that ERN's are deeper for blacks and traditional middle-eastern males relative to white males, we computed the ERN's for each of the groups. Overall, no significant differences were found, $\underline{F}$(3, 66) = 1.412, $\underline{p}$ = .247. Middle-eastern males evoked the deepest ERN's when depicted in traditional clothes (−4.68 μV) and western clothes (−3.64 μV). White males evoked moderate ERN's (−2.62 μV), while black males evoked the least negative ERN's (−1.11 μV). Post hoc comparisons revealed that ERN differences between the black and traditional middle-eastern males were significant, $\underline{t}$(30) = 2.104, $\underline{p}$ = .044. But, both groups were expected to evoke the deepest ERN's.

Lack of significance in the overall analysis was partly due to low error rates for the white primes and the resulting loss of power (1 − β= .358). Computing ERN's based on errors leaves many empty cells for the white primes. This is problematic because the overall analysis depends on full cells across each of the prime conditions. Even more troubling is the fact that the absence of errors would be the best evidence for the efficacy of cognitive control, so when participants make no errors, not only can we not measure their ERN's, we do not have any measure of the best evidence for cognitive control. On the other hand, there is no reason to exert cognitive control if there is no automatic activation of bias. Perhaps the white, black, and middle-eastern males in traditional clothing did not evoke any bias. This would explain the lack of significance as well.

Despite the lack of overall differences in ERN's across the prime conditions, a positive and significant correlation was found between the overall number of false positive errors and the overall depth of negativity in the ERN recorded in the Fz electrode, $\underline{r}$ (33) = .410, $\underline{p}$ = .018. In other words, participants who made fewer false positive errors also showed deeper ERN's when they made these errors. This result is consistent with the findings of Amodio et al. (2004). To further understand the relationship between ERN's and false positive errors, individual ERN scores were computed for each participant across each prime condition. ERN scores for each condition were then correlated with the number of false positive errors for that condition. ERN's computed following traditional middle-eastern primes were significantly and positively correlated with false positive errors to this ethnic group, $\underline{r}$ (30)

= .545, p = .002. Deeper ERN’s were associated with fewer errors to middle-eastern males in traditional clothing. Even though the correlations for the other racial and ethnic groups were not significant, these results are encouraging. This is very strong evidence that cadets were attempting to overcome bias toward middle-eastern males wearing traditional clothing.

Pupil Data

The mean pupil size overall was 565 units with a standard deviation of 249. A factorial ANOVA was performed using pupil size as a measure of arousal across the 5 image types and 7 interest areas. No interaction effect was found, F (20, 2291) = 0.172, p = 1.00. Lack of an interaction suggests that interest areas such as eyes, mouth, hair, beard, and clothing contribute to changes in pupil size independently of the racial and ethnic categories. This bears some discussion because an interaction of this type was anticipated. It is reasonable to define racial and ethnic categories in terms of differences in facial features and forms of clothing. Certain features may stand out as diagnostic of race and ethnicity (i.e., beard and turban). Lack of interaction suggests that racial and ethnic categories are not perceived based on defining features, rather ethnicity and race are cumulative perceptions based on all features taken together (see Figure 4).

Examining the differences in pupil size across the racial and ethnic groups, a significant main effect of image type was found, F (4, 2291) = 3.22, p = .012. Pupil size was greatest for the image of Osama bin Laden (598), as would be expected given that he represents the face of terrorism in the Middle East. The next largest measures of pupil size were obtained for middle-eastern males wearing traditional clothing (583). Again, this is consistent with a stereotype for terrorism. Importantly, pupil size was smaller for the middle-eastern males in western clothing (571) relative to those in traditional clothing. If pupil size is a direct index of arousal, then western clothing seems to ameliorate the threat perceived in middle-eastern males, but not completely. As expected, the images of whites induced the smallest measures of pupil size (539), a result that is consistent with a lack of perceived threat from own-race faces. The pupil size for blacks (546) was only slightly larger than for whites. This result was not expected because it is generally assumed that blacks are perceived as threatening by whites. Norwich University is primarily white, but the Corps of Cadets does have a representative number of black cadets who train side-by-side with the white cadets. One explanation for this discrepancy is that cadets in our sample perceive blacks as comrades rather than enemies, and this could also explain why blacks did not evoke deeper ERN’s. Nonetheless, pupil size does seem to be an excellent index of arousal and threat perception because the greatest pupil sizes were found for Osama bin Laden, a known terror threat, and middle-eastern males in traditional clothing.

For the separate interest areas within the images, a significant main effect was also found, F (6, 2291) = 2.64, p = .022. This effect revealed larger pupil size measurements for clothing relative to other interest areas (see Figure 4). In fact, clothing produced the largest measures of pupil size across all racial and ethnic groups, and the trend toward larger pupil size across these groups was reflected in the clothing. The largest pupil size overall was observed for the traditional clothing worn by middle-eastern males, which is consistent with the potential threat posed by the stereotype for this group and the diagnostic salience of clothing. Before concluding that clothing was the defining feature that induced the perception of threat in cadets, however, it is very important to note that pupil size was also largest for the clothing in the images of both whites and blacks. These raw values were even larger than for the clothing worn by Osama bin Laden. This result makes it is hard to claim that the mere presence or absence of a robe or turban predicted the perception of threat. The threat perceived in Osama bin Laden was not due simply to a man wearing a turban and robe. Threat perception is more holistic. Osama bin Laden is threatening because all of his features combine to produce an average pupil size that is larger than any other image. Clothing emerged as the most salient threat, especially for the traditional robes and headgear for middle-eastern males, but this feature was not the only threat perceived.

If threat perception can influence the decision to shoot, then how is pupil size related to false positive errors in the weapon identification task? To answer this question, pupil size was averaged across participants. The difference between the pupil size for Osama bin Laden and the pupil size for whites was calculated to obtain a measure of pupil dilation. The mean increase in pupil size for Osama bin Laden was +59 units. Likewise, calculations of the dilation scores for the remaining racial and ethnic groups yielded increases in pupil size of +44 units for middle-eastern males in traditional clothing, increases of +32 units for middle-eastern males wearing western clothing, and increases of +7 units for black males. In a repeated measures ANOVA, these dilation scores differed significantly across the racial and ethnic groups just as the raw pupil sizes did in the previous factorial analyses, $\underline{F}(3, 93) = 8.208$. $p = .001$.

Dilation scores across participants were then correlated with the false positive error rates in the weapon identification task. More false positive errors overall were associated with larger pupil dilation toward the image of Osama bin Laden, $\underline{r}\ (28) = .364$, $\underline{p} = .028$. This correlation was positive and significant, offering a tantalizing relationship between pupil dilation and the overall number of false positive errors found in the behavioral data. Cadets who exhibited greater arousal to images of Osama bin Laden were more likely to make false positive errors overall. Specific predictions of pupil dilation for each of the racial and ethnic primes will be described further in the multiple regression analyses.

Multiple Regression Analyses

Multiple regression is a powerful tool for testing the fit of a model to existing data. In the present study, five important predictors of false positive errors in a weapon identification task were measured based on the autmatic and controlled processes identified in traditional dual and quadruple process models. These included pupil dilation scores (autonomic arousal/perceived threat), ERN scores (brain measures of cognitive control), reaction times for error responses (speed-accuracy trade-off), along with measures of signal detection bias (decision criterion) and signal detection sensitivity (stimulus discriminability). These five measures were used as predictor variables in a multiple regression analysis. Our predictions are based on a multi-process model in which more false positive errors are predicted by greater perceived threat (higher autonomic arousal) and less cognitive control (less activity in the ACC), when participants with more liberal criteria for the shoot response are asked to make speeded decisions for less discriminable stimuli.

Multiple regression analyses were conducted on the overall false positive error rates as well as the error rates to each specific racial and ethnic prime. This allowed for the determination of a best fit model for false positive errors in general and for each racial and ethnic group in particular. In the overall analysis, false positive error rates were predicted by pupil dilation scores relative to Osama bin Laden, ERN scores, and reaction times. When pupil dilation scores were entered into the equation first, 13.3% of the variance was explained, $\underline{R} = .364$, $\underline{R^2} = .133$. Adding ERN scores next explained an additional 5.8% of the variance, $\underline{R} = .437$, $\underline{R^2} = .191$. To complete this model, reaction times for the false positive errors were added to the equation and an additional 12.1% of the variance in the model was explained, $\underline{R} = .558$, $\underline{R^2} = .312$, revealing a characteristic speed-accuracy trade-off. The overall model was significant, $\underline{F}(3, 24) = 3.623$, $\underline{p} = .027$, and the three variables explained 31.2% of the variance in the false positive error rates. All three of these variables were significant correlates of false positive errors as reported in Table 3.

Predicting false positive errors to middle-eastern males in traditional clothing, pupil size was entered into the equation first and it explained 14.5% of the variance, $\underline{R} = .381$, $\underline{R^2} = .145$. ERN scores explained an additional 19.7% of the variance, $\underline{R} = .585$, $\underline{R^2} = .342$. Reaction times also added 5.3% of explained variance, $\underline{R} = .629$, $\underline{R^2} = .395$. The regression model including these three variables was significant, $\underline{F}(3, 20) = 4.576$, $\underline{p} = .013$. Note that each of these three predictor variable were significant correlates of the false positive error rates as shown in Table 3. When signal detection bias and sensitivity were added to the equation, a total of 81.0% of the varaince was explained, $\underline{R} = .900$, $\underline{R^2} = .810$. Bias explained an additional 6.9% of the variance and sensitivity explained the remaining 34.6 % of the variance. These results are not surprising given that bias and sensitivity are calculated using false positive errors relative to hit rates. More importantly, weapon bias toward middle-eastern males in traditional clothing was influenced by automatic activation of threat (larger pupil dilation) and greater cognitive control (deeper ERN's) as predicted by both dual and quadruple process models. The speed-accuracy trade-off was also evident as faster responses in this speeded task were associated with higher error rates.

Predicting false positive errors to middle-eastern males in western clothing reveals a slightly different model. Pupil dilation scores predicted an initial 13.4% of the variance, $\underline{R} = .366$, $\underline{R^2} = .134$. ERN's, however, only predicted an addition 0.3% of the variance, $\underline{R} = .370$, $\underline{R^2} = .137$. Reaction times predicted 5.4% of additional variance, $\underline{R} = ..438$, $\underline{R^2} = .191$. Bias explained an additional 3.1% of variance, while sensitivity explained 51.9% of the variance, yielding a model that predicted 74.2% of the variance, $\underline{R} = .861$, $\underline{R^2} = .742$. In this model, pupil dilation predicted errors to middle-eastern males wearing western clothing, but ERN did not.

Predicting false positive errors to the white primes reveals a similar model. Pupil dilation scores predicted an initial 29.0% of the variance, $\underline{R} = .539$, $\underline{R^2} = .290$. This variable alone produced a significant model, $\underline{F}(1, 15) = 6.135$, $\underline{p} = .026$. ERN scores predicted an additional 0.5% of the variance, $\underline{R} = .544$, $\underline{R^2} = .295$. Reaction times predicted 1.0% of variance, while bias and sensitivity explained 34.9% and 15.9% of the variance respectively, yielding a model that predicted 81.3% of the variance, $\underline{R} = .902$, $\underline{R^2} = .813$. This model for white males is similar to the model for the middle-eastern males wearing western clothing, as can be seen in the pattern of correlations in Table 3. It is tempting to conclude that both groups are treated the same. No automatic stereotype bias should be activated by white faces, and therefore no need for overcoming bias. This may explain why ERN is not a predictor of errors for white primes and middle-eastern primes wearning western clothing.

Black primes reveal an altogether different model. The only variables that predicted false positive errors following black primes were signal detection bias and sensitivity. When pupil dilation scores were entered into the model, 2.5% of the initial variance was explained, $\underline{R} = .157$, $\underline{R}^2 = .025$. ERN scores added only 1.0% of explained variance, $\underline{R} = .188$, $\underline{R}^2 = .035$. Even reaction times only added 0.2% of explained variance, $\underline{R} = .191$, $R^2 = .037$. Bias explained an additional 3.3% of the variance, while sensitivity explained the majority of the variance at 71.1%. Despite the lack of prediction by the primary variables, signal detection sensitivity (d′) helped the model explain 78.1% of the variance in false positive errors following black primes, $\underline{R} = .884$, $\underline{R}^2 = .781$.

Of all the data reported in this study, the lack of prediction for false positive errors following the black primes is the most perplexing finding. It is possible that cadets at Norwich University have no bias toward blacks, and this is supported by the lack of pupil dilation to black images, and by the lack of ERN’s following errors to tools primed by black images. Yet, higher error rates were observed for black primes in the weapon identification task. It is possible that the presence of highly arousing images of middle-eastern males changed the demands of the task for the cadets. They may have adopted an enemy-detection strategy and in this context black males are more likely to be allies than enemies. This may explain why arousal was not high for blacks, but given that decisions to shoot were still made, deeper ERN’s would have been predicted because cognitive control would help reduce friendly fire casualties. It is hard to draw conclusions based on the lack of an observed effect, but because decisions to shoot have real-world implications, it is hard not to wonder if black soldiers are more likely to be casualties of friendly fire.

In summary, multiple process models such as the dual and quadruple process models of stereotype activation are supported by the pattern of correlations and the multiple regression models. Most salient is the finding that false positive errors to middle-eastern males in traditional clothing are predicted by all elements of the model: stereotype activation, cognitive control, and task-related demands. In addition, the lack of evidence for cognitive control for the white and middle-eastern males in western clothing supports Fiske and Neuberg (1990). No cognitive control is needed if no negative stereotype is activated by these groups. Clothing appears to make a difference when activating a stereotype, and western clothing produces responses to middle-eastern males that are indistinguishable from white males. Traditional clothing worn my middle-eastern males, however, appears to induce weapon misidentification errors. Military cadets who make the most errors of this type, tend to show greater pupil dilation to images such as Osama bin Laden, produce weaker ERN waveforms in their frontal cortex, and tend to respond faster. They are also more susceptible to task demands, showing less ability to discriminate the gun/tool targets and more liberal bias toward the shoot response.

General Discussion

Armed conflicts place soldiers at extreme risk. They are often deployed in regions of the world where cultures collide and they must discriminate friend from foe. They must also carry weapons with the presumptive authority to shoot and kill. For psychologists who study human social behavior, the decision to shoot raises a swarm of questions, but surprisingly few studies have been done with soldiers to understand the social, cognitive, affective, and physiological basis for these life-or-death decisions. Several pioneering researchers have recently addressed these questions as they pertain to police officers and their decisions to shoot crime suspects (Payne, 2001, 2005, 2006; Plant, Peruche, & Butz, 2005; Correll, Urland, & Ito, 2006; Correll, Park, Judd, Wittenbrink, Sadler, & Keesee, 2007). The present research enters the military domain and addresses questions pertaining to soldiers and how they make decisions to shoot potentially armed enemy combatants. This is a timely investigation given British and American involvement in a protracted Middle East conflict.

Stereotypes are useful heuristics for predicting the behavior of strangers. Just as police officers are faster and more accurate when asked to ‘shoot’ images of black crime suspects holding guns in a weapon identification task, we found that military cadets were faster and more accurate when gun targets were primed by images of middle-eastern males wearing traditional clothing. Clearly there is a benefit to holding a stereotype if indeed that stereotype predicts the threat posed by a racial or ethnic group. Police officers and soldiers are trained to identify weapons and to respond to those persons holding weapons with deadly force. If indeed black crime suspects and/or middle-eastern males wearing turbans are more likely to be found holding weapons, then the perception of threat would be warranted and a bias toward shooting would benefit the police officer or soldier. The forensic evidence does show that black males are shot and killed by police officers at a rate that is disproportionate to the rate at which white males are shot and killed. But this rate is not disproportionate to the rate at which police officers are shot and killed by black males relative to white males (Kleck, 2007). From this perspective, the stereotype of black males and the threat that they pose may be beneficial to the survival of police officers. In a similar manner, a negative stereotype of middle-eastern males wearing traditional robes and the threat that they pose to soldiers deployed in the Middle East may also be beneficial.

On the negative side, however, stereotypes can lead to false positive errors. As predicted, we found that military cadets made more false positive errors when tools were primed by images of middle-eastern males wearing traditional clothing. From a signal-detection perspective, it would be ideal if cadets made no decisions to shoot when confronted with tools, and always made decisions to shoot when confronted with guns. But any bias in the decision to shoot will benefit correct detections at the expense of false positive errors, or correct rejections at the expense of misses. Our results do in fact indicate a bias toward the decision to shoot when confronted with images of middle-eastern males wearing traditional clothing. This accords with the view that a negative stereotype has emerged toward middle-eastern males to the extent that they are depicted wearing traditional robes and turbans congruent with the cultural stereotype.

Cultural stereotypes present many difficulties for researchers because they depend on a wide range of racial and ethnic differences between people. As predicted by Fiske and Neuberg (1990), the bias of military cadets toward middle-eastern males was evident when the clothing was congruent with the traditional robes and headgear of the Middle East. However, when the middle-eastern males depicted in our images wore western clothing, such as culturally incongruent suits and ties, this observed bias was reduced. Despite the fact that the actual terror suspects in the 9/11 attack were middle-eastern males wearing western clothing, cadets showed minimal bias in the decision to shoot when primed by these images. Instead, the negative stereotype associated with Osama bin Laden as the exemplar of Middle East terrorism did seem to influence their decisions to shoot when primed by middle-eastern males wearing traditional clothing.

The influence of cultural stereotypes on weapon bias is mediated by perceived threat. We chose to use the novel measure of pupil size in order to examine autonomic arousal and perceived threat. Pupil size was found to vary significantly across the racial and cultural groups represented in our stimuli, and these differences reflected the trends found in the weapon bias data. Middle-eastern males in traditional clothing induced larger measures of pupil size, with the image of Osama bin Laden showing the largest pupil dilation as might be expected given his role in the terrorist attacks of 9/11. Interestingly, the middle-eastern males wearing western clothing induced moderate levels of threat as measured by pupil size. Arousal toward these images may have been ameliorated by the incongruent cue of the clothing.

The Quad model of weapon bias, as recently articulated by Sherman and his colleagues (Sherman, Gawronsky, Gonsalkorale, Hugenberg, Allen, & Groom, 2008), provides a good fit for the racial and cultural biases observed in our study. Pupil dilation provides an excellent physiological index of arousal that can predict false positive error rates to different racial and ethnic groups. The predictiveness of pupil dilation in the multiple regression analysis is even more impressive given that these measures were obtained from a separate free-viewing paradigm. ERN’s, on the other hand, were the strongest predictors of the false positive errors to the middle-eastern males, a result that is consistent with the Quad model if a strong negative stereotype toward this group must be overcome to avoid bias. Further studies with better measures of ERN are needed to fully evaluate the Quad model and its predictions for whites, blacks, and other potential outgroups.

Finally, the brain is ultimately the organ that controls the decision to shoot, and we have posited, along with other researchers who have used ERN’s as indices of cognitive control (Amodio, Harmon-Jones, Devine, Curtin, & Hartley, 2004; Amodio, Devine, & Harmon-Jones, 2008), that the ACC plays a central role in this executive decision. Our ERN data were consistent with prior research on brain activity related to errors. Deeper peak negativity was found in the midline frontal electrode Fz when participants made false positive errors. These ERN’s were correlated positively with the number of errors committed by our participants. Cadets who made more false positive errors tended to have more positive (i.e., less negative) ERN’s. They did not seem to learn from their mistakes and they continued to commit errors. Participants who made fewer errors, however, showed more negative (i.e., less positive) ERN’s. They seemed to demonstrate high levels of self-monitoring and greater control over errors. Our conclusion is that weapon bias is a function of underlying brain activity. The anterior cingulate cortex (ACC) monitors responses and shows greater activity (lower ERN's) following error responses when participants show greater restraint from making false positive errors.

The long-term goal of this research is to develop training protocols that can reduce bias and thereby limit casualties of war due to misidentification of threat and friendly fire. This goal presumes that soldiers have control over their responses to the extent that they can over-ride perceived threat and cultural stereotypes in order to make correct decisions to shoot. Historically, the view of the brain as a social organ has shaped the modern field of social neuroscience. The triune brain theory proposed originally by MacLean in 1949 (MacLean, 1989) is essentially a model of brain function that places cognitive neocortical control over emotional limbic arousal. Arguably, humans need a large neo-cortex to suppress impulses generated by the limbic system. Decisions to shoot are perhaps the ultimate social responses to fight-or-flight situations. Humans are really the only threat to humans, and that threat is real when humans are holding weapons.

More recent views on the social brain hypothesis have targeted specialized structures within the neo-cortex that interface between emotion and cognition. Allman (1998) extensively examined the anterior cingulate cortex and the specialized spindle cells in this region that are phylogenetically recent and perhaps unique to great apes and humans. There are not a large number of these spindle cells in the ACC and pre-frontal cortex, but they project to many sites throughout the brain and have been implicated in numerous psychopathologies that affect cognitive control over impulses such as attention-deficit disorder, schizophrenia, clinical depression, and obsessive-compulsive disorder. For veterans of combat, post-traumatic stress disorder (PTSD) has also been linked to the ACC, to the extent that there is a breakdown in the ability to control emotions. It is vital that we begin to understand the role of the ACC in controlling emotions and executing accurate decisions to shoot. There are profound implications that long-term deployment of soldiers in volatile situations can literally wear down the normal functioning of the ACC and lead to less accurate decisions to shoot. Soldiers returning home after long stretches of service are also vulnerable to fluctuations in ACC functioning that may lead to suicide or even homicide. To assume that every soldier has an equal strength and resiliency is not warranted. The brain is a complex social organ and the decision to shoot places great strain on its cognitive, affective, and social functions.

Acknowledgments

This publication was made possible by the Vermont Genetics Network through Grant Number P20 RR16462 from the INBRE Program of the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of NCRR or NIH. Many thanks are due to Yuliya Omarov, who assisted in all phases of the ERP data analysis, and to Will Dorion, who assisted in writing the eye-tracking programs.

References

Adolphs R. Perception and emotion: How we recognize facial expressions. Current Directions in Psychological Science. 2006; 15(5):222–226.

Allman, JM. Evolving Brains. New York: W. H. Freeman and Company; 1998.

Allman JM, Hakeem A, Erwin JM, Nimchinsky E, Hof P. The anterior cingulate cortex: The evolution of an interface between emotion and cognition. Annals of the New York Academy of Sciences. 2001:107–117. [PubMed: 11411161]

Amodio DM, Harmon-Jones E, Devine PG, Curtin JJ, Hartley SL, Covert AE. Neural signals for the detection of unintentional race bias. PsychologicalScience. 2004; 15:225–232.

Amodio DM, Devine PG, Harmon-Jones E. Individual differences in the regulation of intergroup bias: The role of conflict monitoring and neural signals for control. Journal of Personality and Social Psychology. 2008; 94(1):60–74. [PubMed: 18179318]

Bradley MM, Miccoli L, Escrig MA, Lang PJ. The pupil as a measure of emotional arousal and autonomic activation. Psychophysiology. 2008 (In Press).

Bush, G. Multimodal studies of cingulate cortex. In: Posner, MI., editor. Cognitive neuroscience of attention. New York: The Guilford Press; 2004.

Conrey FR, Sherman JW, Gawronski B, Hugenberg K, Groom C. Separating multiple processes in implicit social cognition: The quad-model of implicit task performance. Journal of Personality and Social Psychology. 2005; 89:469–487. [PubMed: 16287412]

Correll J, Park B, Judd CM, Wittenbrink B. The police officer's dilemma: Using ethnicity to disambiguate potentially threatening individuals. Journal of Personality and Social Psychology. 2002; 83:1314–1329. [PubMed: 12500813]

Correll J, Urland GR, Ito TA. Event-related potentials and the decision to shoot: The role of threat perception and cognitive control. Journal of Experimental Social Psychology. 2006; 42:120–128.

Correll J, Park B, Judd CM, Wittenbrink B, Sadler MS, Keesee T. Across the thin blue line: Police officers and racial bias in the decision to shoot. Journal of Personality and Social Psychology. 2007; 92(6):1006–1023. [PubMed: 17547485]

Critchley HD, Tang J, Glaser D, Butterworth B, Dolan RJ. Anterior cingulate activity during error and autonomic response. Emotion. 2007; 7(4):724–729. [PubMed: 18039039]

Devine PG. Prejudice and stereotypes: Their automatic and controlled components. Journal of Personality and Social Psychology. 1989; 56:5–18.

Fiske, ST.; Neuberg, SL. A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In: Zanna, MP., editor. Advances in experimental social psychology. Vol. 23. New York: Academic Press; 1990. p. 1-74.

Greenwald AG, Oakes MA, Hoffman HG. Targets of discrimination: Effects of race on responses to weapon holders. Journal of Experimental and Social Psychology. 2003; 39:399–405.

Holroyd, CB.; Nieuwenhuis, S.; Mars, RB.; Coles, MGH. Anterior cingulate cortex, selection for action, and error processing. In: Posner, MI., editor. Cognitive neuroscience of attention. New York: The Guilford Press; 2004.

Ito TA, Thompson E, Cacioppo JT. Tracking the timecourse of social perception: The effects of racial cues on event-related brain potentials. Personality and Social Psychology Bulletin. 2004; 30(10): 1267–1280. [PubMed: 15466600]

Kahneman, D. Attention and Effort. Englewood Cliffs, NJ: Prentice -Hall; 1973.

Kleck G. Are police officers more likely to kill African-American suspects? Psychological Reports. 2007; 100(1):31–34. [PubMed: 17451000]

Luu, P.; Pederson, SM. The anterior cingulate cortex: Regulating actions in context. In: Posner, MI., editor. Cognitive neuroscience of attention. New York: The Guilford Press; 2004.

MacLean, P. The Triune Brain in Evolution. New York: Plenum Press; 1989.

Macmillan, NA.; Creelman, CD. Detection theory: A user's guide. Cambridge: Cambridge University Press; 1991.

Neely JH. Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited capacity attention. Journal of Experimental Psychology: General. 1977; 106:226–254.

Payne BK. Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. Journal of Personality and Social Psychology. 2001; 81:181–192. [PubMed: 11519925]

Payne BK. Conceptualizing control in social cognition: How executive functioning modulates the expression of automatic stereotyping. Journal of Personality and Social Psychology. 2005; 89:488–503. [PubMed: 16287413]

Payne BK. Weapon bias: Split-second decisions and unintended stereotyping. Current Directions in Psychological Science. 2006; 15(6):287–291.

Plant EA, Peruche BM, Butz DA. Eliminating automatic racial bias: Making race non-diagnostic for responses to criminal suspects. Journal of Experimental and Social Psychology. 2005; 41:141–156.

Ridderinkhof KR. Micro- and macro-adjustments of task set: Activation and suppression in conflict tasks. Psychological Research. 2002; 66:312–323. [PubMed: 12466928]

Ridderinkhof, KR.; van den Wildenberg, WPM.; Wijnen, J.; Burle, B. Cognitive neuroscience of attention. Posner, MI., editor. New York: The Guilford Press; 2004.

Scheffers MK, Coles MGH. Performance monitoring in a confusing world: Error-related brain activity, judgments of response accuracy, and types of errors. Journal of Experimental Psychology: Human Perception and Performance. 2000; 26:141–151. [PubMed: 10696610]

Sherman JW, Gawronsky B, Gonsalkorale K, Hugenberg K, Allen TJ, Groom CJ. The self-regulation of automatic associations and behavioral impulses. Psychological Review. 2008; 115(2):314–335. [PubMed: 18426292]

Tan, M. USA Today. 2006. Families of friendly fire casualties on mission for truth.

Figure 1.

Sample images of white, black, and middle-eastern faces used as primes.

Figure 2.
Sample images of gun and tool targets.

Figure 3.
Sample image of Osama bin Laden used in the eye-tracking study.

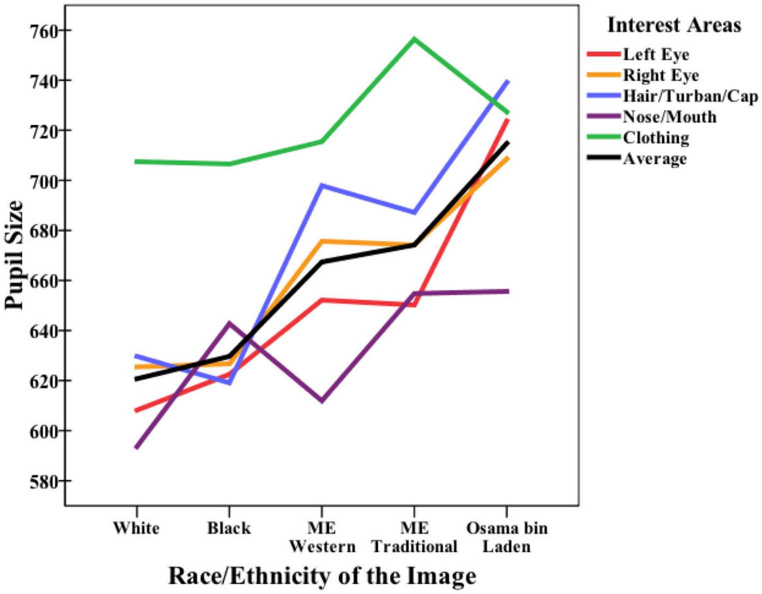

Figure 4.

Mean pupil size as a function of image type and interest areas.

Table 1

Valid Response Measurements in the Weapon Identification Task as a Function of Face Primes

	Face Primes			
Response Types	**White**	**Black**	**Middle-Eastern (Western)**	**Middle-Eastern (Traditional)**
Tool Trials				
False Positive Errors				
Percentage	12.3 %	18.3 %	16.4 %	19.1 %
Mean RT	259 ms	267 ms	258 ms	257 ms
Correct Rejections				
Percentage	57.0 %	46.7 %	54.1 %	50.0 %
Mean RT	330 ms	353 ms	348 ms	353 ms
Gun Trials				
Misses				
Percentage	28.8 %	26.7 %	29.7 %	23.3 %
Mean RT	277 ms	282 ms	290 ms	247 ms
Correct Detections				
Percentage	42.7 %	47.6 %	39.4 %	53.9 %
Mean RT	353 ms	330 ms	384 ms	332 ms

Table 2

Signal Detection Bias and Sensitivity Measures as a Function of Face Primes

	Face Primes			
Scores	**White**	**Black**	**Middle-Eastern (Western)**	**Middle-Eastern (Traditional)**
False Alarms (z)	−1.87	−0.66	−1.13	−0.70
Hits (z)	0.32	0.43	0.11	0.75
Bias (c)	0.78	0.07	0.51	−0.02
Sensitivity (d')	2.19	0.98	1.24	1.45

$c = -.5(z_H + z_{FA})$

Using the SPSS computational formula, $c = -.5*[\text{PROBIT(Hits)} + \text{PROBIT(False Alarms)}]$

$d' = z_H - z_{FA}$

Table 3

Correlations of Multiple Regression Predictor Variables with False Positive Errors by Condition

	False Positive Errors by Face Prime Condition				
Predictors	**White**	**Black**	**Middle-Eastern (Western)**	**Middle-Eastern (Traditional)**	**Overall**
Pupil Dilation	.539 **	.157	.366 *	.381 *	.364 *
ERN	−.065	.107	.134	.495 **	.410 *
Reaction Time	−.220	.023	−.215	−.456 *	−.361 *
Bias (c)	−.779 **	−.241	−.103	−.257	---
Sensitivity (d′)	−.606 **	−.417 *	−.463 *	−.483 **	---

* $p < .05$

** $p < .01$