Robots Can Be More Than Black And White: Examining Racial Bias Towards Robots

Arifah Addison, Kumar Yogeeswaran, Christoph Bartneck
University of Canterbury, Private Bag 4800, 8140 Christchurch, New Zealand

Abstract

Previous studies showed that using the 'shooter bias' paradigm, people demonstrate a similar racial bias toward dark colored robots over light colored robots (i.e., Black vs. White) as they do toward humans of similar skin tones (Bartneck et al. 2018). However, such an effect could be argued to be the result of social priming. Additionally, it raises the question of how people might respond to robots that are in the middle of the color spectrum (i.e., brown) and whether such effects are moderated by the perceived anthropomorphism of the robots. We conducted two experiments to first examine whether shooter bias tendencies shown towards robots is driven by social priming, and then whether diversification of robot color and level of anthropomorphism influenced shooter bias. Our results showed that shooter bias was not influenced by social priming, and interestingly, introducing a new color of robot removed shooter bias tendencies entirely. However, varying the anthropomorphism of the robots did not moderate the level of shooter bias, and contrary to our expectations, the robots were not perceived by the participants as having different levels of anthropomorphism.

Introduction

Although some research has been done around how gender and personality of artificial intelligence is perceived and reacted to, e.g. (Tay, Jung, and Park 2014), little research has covered the area of perceived race. This is likely because most robots developed so far are not very human-like. However, those that are could be seen to have a race, as shown in Figure 1, and the first studies have begun to emerge on the topic. For example, by analysing free-form comments about highly human-like robots, it has been shown that people more frequently dehumanise robots racialized as Asian and Black, than they do robots racialized as White (Strait et al. 2018).

The shooter bias paradigm was first introduced in a 2002 paper titled 'The Police Officer's Dilemma: Using Ethnicity to Disambiguate Potentially Threatening Individuals'. In this paper, a clear bias was found when participants were asked to 'shoot' images of Black or White human males that were holding a gun, and 'not shoot' Black or White human males holding benign objects (e.g., a soda can, wallet, cell phone). Both Black and White participants from the USA showed a similar tendency to be quicker to shoot armed Black males and quicker to not shoot unarmed White males (than the other). Such an effect is collectively referred to as a shooter bias, via an interaction between race and object type. A recent meta-analysis of 42 studies supports this general tendency of a shooter bias in the psychological literature (Mekawi and Bresin 2015). In a recent follow up, the shooter bias paradigm was used to show that the racial bias shown towards Black humans was transferred to humanoid robots of different colors (Bartneck et al. 2018). This study used half human images from the original 2002 experiments, and half 'racialized' robots that were re-colored with Black and White human skin tones, and the results were found to be comparable to the original human-only study. Specifically, people showed a similar shooter bias toward robots racialized as Black relative to White in a similar fashion as they showed toward Black vs. White humans, no matter their own race.

Figure 1: Erica: Erato Ishiguro Symbiotic Human-Robot Interaction Project

The current research expands on this preliminary work by examining several unanswered questions. (1) First, would such a shooter bias effect emerge if humans were removed from the task entirely? Ogunyale, Bryant, and Howard (2018) argued that the robot shooter bias experiment was a "classic case of social priming" because participants completed ratings of Black Americans and White Americans before completing the shooter bias task. Additionally, participants made judgements to shoot both Black and White humans and robots in the same task, so stereotypes about Black people may have simply been applied to robots of the same skin tone. Therefore, one of the primary goals of the present work was to address this criticism by having participants complete the shooter bias task prior to any questions about race or ethnicity and by removing the humans from the experiment entirely so that participants complete the task with the robot images only.

(2) Second, does shooter bias follow a continuum from dark to light skin tones, or is it largely based on stereotypes and prejudices that people have toward a certain group? Previous research suggests that there is a strong shooter bias towards Black males even when Latino and Asian males are present as well as Whites, but there was no significant difference found between Latino, Asian and White males (Sadler et al. 2012). Therefore, in the present work, we examined whether shooter bias would be particularly evident toward robots racialized as Black relative to those that are Brown or White.

(3) Third, do shooter bias effects observed in earlier work depend on the degree of perceived anthropomorphism of the robot? Anthropomorphism has an important role in HRI. As shown by Zlotowski et al. (2015), when interacting with humans, the higher the level of anthropomorphism of robots the more effective they are in achieving their goals. Therefore it could be said that people demonstrate even stronger shooter bias tendencies toward robots that are more human-like relative to those that are more machine-like. However, as prior work had only used one robot type in the experiments (i.e., the Nao robot), it is unclear whether shooter bias effect would vary as a function of the perceived anthropomorphism of the robot (i.e., from less to more human-like). The present work addresses this question by examining shooter bias using three different robots (i.e., Inmoov, Nao, Robosapian) that we perceived as varying on a continuum of anthropomorphism from human-like to machine-like.

As a secondary goal, we wanted to find out if the number of trials had an effect on shooter bias tendencies. Previous work has shown that the more cognitively fatigued people are, the more their implicit biases play a part in their decisions. For example, cognitive busyness both inhibits activation of stereotyping and increases the likelihood that any activated stereotypes will be applied (Gilbert and Hixon 1991), and Macrae, Milne, and Bodenhausen (1994) suggests that stereotyping evolved to help preserve mental processing resources. Correll et al. (2013) found that cognitive fatigue has an influence on shooter bias, but they investigated pre-game cognitive loading rather than the number of trails. Therefore, we wanted to find out whether number of trials as a measure of fatigue had an influence on shooter bias in order to reduce the risk of unknown variables across experiments.

Experiment A: Social Priming and Fatigue

In this experiment we sought to find out whether social priming may have caused the shooter bias effect in the initial robot shooter bias study. We did this by removing any questions about race or ethnicity before the shooter bias task and only asked participants any race-related questions after they had played the game. We further ensured that there was no social priming by removing all human images from the experiment so that people would only complete the shooter bias task with the racialized robots (i.e., robots that were black or white in color). We further sought to find out whether task length had an influence on the shooter bias effect by analysing the trials in each half of the game-play independently.

Method

We repeated the initial Robot Racism shooter bias experiment without the human images - i.e., with only the robots racialized as Black and White, each image displayed in random order within each of 2 blocks of 64 trials. The experiment factors were 2 (racialization: black vs white) $\times$ 2 (object in hand: gun vs. benign object) x2 (fatigue: block1 vs block2) within-subjects design. Prior to playing the game, participants provided their basic demographics data, and after the game they were asked to ascribe a race to images of the Black and White robots.

Participants

A total of 113 participants from the USA were recruited from Amazon Mechanical Turk (AMT)[1]. As this was a different platform from the previous robot shooter bias studies, there was a low chance of the same people being recruited, but we also included a post-task question as to whether they had been part of any similar studies as a check. There is evidence that results obtained by recruiting participants through AMT are similar to those obtained by running a study in a lab (Bartneck et al. 2015). Participants received $1.50 USD for completing the experiment and additional bonuses were offered to the top players (based on speed and accuracy). We also advised them that we may restrict payment for those who did not put genuine effort into playing the game correctly (e.g. high no-response rate or hitting the same key repeatedly). After analysing the data accordingly, we excluded all those who achieved less than 80% success rate to filter out all suspicious response patterns. This left a sample of 106 participants (42 female, 64 male). Participant age range was 18 to 58 years ($m = 38; sd = 8.26$). A majority of these participants reported being of White/Caucasian descent ($n = 84$), with others identifying as Black/African American ($n = 9$), Asian ($n = 6$), Latino/Hispanic ($n = 5$), Native American ($n = 2$), and mixed race ($n = 2$).

[1] https://www.mturk.com/

Stimuli

We used the Nao robot images from the initial robot racism study, which had been re-colored with human skin tones from the African and Caucasian women in the professional photograph of multiracial women shown in (Bartneck et al. 2018, Figure 2). Each robot was either holding a gun, a remote control, a candy bar or a soda can (Figure 2). Using 16 backgrounds × 2 skin colors (Black vs White) × 2 objects (gun vs benign object) gave us a total of 64 images.

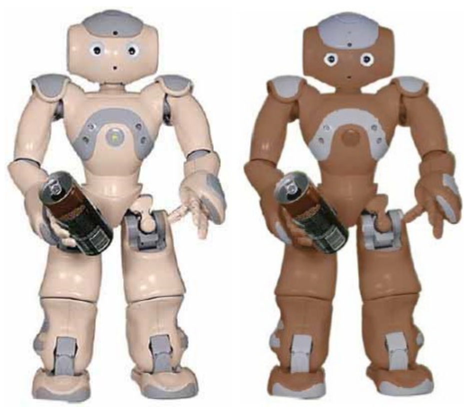

Figure 2: 'White' and 'Black' Nao robots.

Procedure

The task was carried out via Inquisit Web [2] which allowed us to record reaction times with millisecond precision. Participants were first provided with an information sheet to read and asked to give consent in Mechanical Turk for their participation in the study before being directed to Inquisit Web.

Participants first answered a series of demographics questions, then went through 20 practice trials with random conditions to ensure a lack of routine. Each trial started with a 500ms fixation, followed by 1-4 empty backgrounds for a random duration (500 to 1000ms), then finally the target image was shown for 850ms. Participants received feedback on how they performed on the trial for 2000ms.

After the practice round, they were allowed to rest and continue to the main test when they were ready - two rounds of 64 trials, each image repeated once per block in random order. After playing the main test the participants were congratulated and given their final score. Participants were then shown images of the Nao robot with each of the two skin tones used in the game in random order and asked to ascribe them a race from a list, including the option of "does not apply" at the very top of the list.

[2] http://www.millisecond.com/

Measures

Demographics Participants completed a demographic questionnaire including questions about their age, race, gender, nationality, and religion.

Shooter Bias We measured participants' reaction times in deciding whether to shoot/not shoot, and accuracy (correct identification of aggressors versus non-aggressors) while completing the shooter bias task. The reaction time is measured as the time between the end of the stimuli being shown on the screen and the time when a key was pressed. Following the procedures outlined in (Correll et al. 2002), we calculated the average reaction time and accuracy for different conditions for the variables of: robot racialization (Black vs. White), and object in hand (gun vs. benign object).

Robot Race Participants were shown the Black and White Nao robots and asked to ascribe a race to each from a list of options, which also included "Does not apply" at the top of the list.

Results and Discussion

As per the initial robot racism study which followed the procedure in (Correll et al. 2002), we analysed the average log-transformed reaction times for correct trials, and the average accuracy rates. In addition we analysed the two blocks of 64 trials separately as a measure of fatigue. The means and standard deviations for all conditions are shown in Table 1.

Robot Race Participants were asked to ascribe a race to each of the colors of Nao robot. Although the first option in the list was "Does not apply", 86% of participants did ascribe a race to the robots. The robot racialized as 'White' was mostly identified as White/Caucasian (65%) with the next highest identification being Asian (9.7%). The robot racialized as 'Black' was mainly ascribed to be Black/African American (47%) or Latino/Hispanic (21%), and 8% identifying it as Indian. This suggests that our manipulation of the skin color did serve to alter the perceived race of the robot in the eyes of most participants.

Shooter Bias A 2×2 analysis of variances (ANOVA) revealed the expected significant 2-way interaction between racialization and object in hand for reaction time ($F(1,105) = 7.428, p = 0.008, \eta^2 = 0.066$), but not for accuracy ($F(1,105) = 0.114, p = 0.736, \eta^2 = 0.001$). Similar to previous work (Bartneck et al. 2018), participants took significantly longer to not shooter unarmed robots that were racialized as Black compared to those racialized as White ($p < .001$), but there was no difference in time taken to shoot an armed robot that was racialized as Black or White ($p = .89$).

We then proceeded to examine whether fatigue emerging from task length would impact on shooter bias tendencies. A $2 \times 2 \times 2$ ANOVA revealed non-significant 3-way interaction effect between racialization, object, and fatigue on both reaction time ($F(1,105) =<$

Table 1: Means and Standard Deviations for Reaction Times and Accuracy within Blocks

		Reaction Times		Accuracy	
		Black	White	Black	White
Gun	Block1	581 (52)	582 (51)	0.95 (0.07)	0.95 (0.06)
	Block2	581 (53)	580 (55)	0.97 (0.05)	0.96 (0.06)
Benign Object	Block1	638 (54)	633 (52)	0.91 (0.09)	0.91 (0.09)
	Block2	627 (54)	619 (54)	0.95 (0.07)	0.94 (0.09)

$.001, p = 0.992, \eta^2 < 0.001$) or accuracy ($F(1, 105) = 0.285, p = 0.595, \eta^2 = 0.003$). However, there was a significant main effect of fatigue on reaction time, ($F(1, 105) = 10.129, p < 0.001, \eta^2 = 0.088$), such that participants were generally faster in the second block than the first block, but this did not interact with robot color. See Table 1) for full results.

To summarise, our first study sought to investigate whether shooter bias towards robots may be a result of social priming by re-running the previous robot shooter bias study with robot images only, and with no survey questions regarding participants attitudes regarding race prior to the shooter bias task. Our results showed that the shooter bias was still present towards the darker colored robot even with no social priming, therefore eliminating it as an influencing factor.

We also sought to find out whether playing a greater number of trials influenced the shooter bias as it had been previously shown that although cognitive fatigue did have an effect, task length had not been factored in to any analysis despite a variation of the number of trials in the different experiments. By analysing the data over two blocks independently, we found that fatigue from trail length did not have an effect on shooter bias tendencies.

Experiment B: Diversification and Anthropomorphism

In this second experiment, we sought to find out whether adding a wider range of human skin tones influenced shooter bias. Previous research with human targets revealed that participants showed similar levels of shooter bias only toward Black males even when including Latino and Asian targets to the original shooter bias paradigm (Sadler et al. 2012). Therefore, including brown colored robots may not significantly impact shooter bias tendencies. In addition to testing whether brown colored robots influenced shooter bias, we also wanted to test if perceived anthropomorphism of robots interacted with robot color and object in hand. Specifically, we wanted to examine whether shooter bias fell across a spectrum from darker to lighter and if this was especially pronounced when a robot was more human-like relative to less human-like.

Method

We repeated the robot shooter bias experiment, but also added robots with a brown skin tone taken from the same image of multiracial women used previously, to achieve a spectrum of racialization from Black to White. Furthermore, we added two more types of robots (Inmoov and Robosapien) to achieve a range of anthropomorphism from less to more human-like where we suspected that Inmoov would be perceived as most human-like, followed by Nao, then Robosapien. The experiment was therefore a 3 (skin tone: Black vs. Brown vs. White) $\times$ 2 (object in hand: gun vs. benign object) $\times$ 3 (robot agent: Inmoov vs. Nao. vs Robosapian) mixed design, with robot agent as the only between-subjects factor so that each participant would encounter only one type of robot in the task.

Participants

A total of 340 participants from the USA were recruited from Amazon Mechanical Turk, and no person was allowed to take part in both experiments. This was done using AMT's ability to exclude workers with a particular 'Qualification Type' that we assigned to each worker who took part in either experiment. As per Experiment A, we did not use the data for all those who achieved less than 80% success rate. This left a sample of 312 participants (162 female, 150 male). Participant age range was 18 to 78 years ($m = 29; sd = 10.8$). A majority of these participants reported being of White/Caucasian descent ($n = 226$), with others identifying as Black/African American ($n = 35$), Asian ($n = 27$), Latino/Hispanic ($n = 20$), Native American ($n = 2$), and mixed race ($n = 2$). Participants received $1.25 USD for completing the experiment which was slightly less than Experiment A due to the shorter playing time (1 round of 60 trials).

Stimuli

We photographed all combinations of the three types of robot (Inmoov, Nao, and Robosapian) and re-colored them with the skin tones of the African, Indian and Caucasian women from the same multi-racial photograph used previously. Using 5 backgrounds $\times$ 3 skin colors $\times$ 2 objects (gun vs non-gun) $\times$ 3 robot type (levels of anthropomorphism) gave us a total of 90 images (30 images for each of the three types of robot).

Procedure

Similar to Experiment A, the test was carried out via Inquisit Web. The main study comprised of one block of 60 trials - 5 backgrounds x 6 conditions, each repeated twice in random order. Each participant saw only one type of robot - Inmoov, Nao or Robosapian. After playing the game as well as ascribing a race to each of the three colors of robot, participants were shown each of the three types of robot and asked to rate them using the five anthropomorphism questions from the Godspeed Questionnaire (Bartneck et al. 2009), a standardised measuring tool.

Measures

Demographics As per Experiment A, participants completed a demographic questionnaire including questions about their age, race, gender and nationality.

Shooter Bias As per Experiment A, we measured the reaction time and accuracy across the different trial conditions for the variables of: agent (robot type: Inmoov vs. Nao vs. Robosapian), racialization (Black vs. Brown vs. White), and object (gun vs. benign object).

Robot Race Participants were shown Black, White and Brown Nao robots and were asked to ascribe each one a race from a list of options, which also included 'Does not apply'.

Robot Anthropomorphism Participants were asked to rate their impression of each of the three robot types using five anthropomorphism questions. Each question used a Likert scale from 1-5 where 0=less; 5=more anthropomorphic (i.e., a high rating is more human-like) to get an overall anthropomorphism rating.

Results and Discussion

As per Experiment A, we analysed the average log-transformed reaction times for correct trials, and the average accuracy rates for the variables of: robot type (Inmoov vs. Nao vs. Robosapian), color (Black vs. Brown vs. White), and object in hand (gun vs. benign object). The means and standard deviations for all conditions are shown in Table 2.

Robot Race After playing the game, participants were asked to ascribe a race to each of the three colors of Nao robot. Although the first option in the list was 'Does not apply', 75% of participants did ascribe a race to the robots. The robot racialized as 'White' was highly identified as White/Caucasian (66%), while the 'Black' robot was mainly ascribed to be Black/African American (28%) or Latino/Hispanic (22%). The light brown robot was mainly ascribed to be Latino/Hispanic (29%) with 11% ascribing it to be African American, or one of other races in the list that would normally be associated with a brown skin tone (e.g. Indian, Asian, Native American). Therefore, it appears that introducing a brown colored robot led to greater variability in the racialization of both the Black and Brown colored robots where although participants clearly identified both as non-White, there was more variability in their racialization as African American, Latino/Hispanic, or another ethnic/racial group.

Robot Anthropomorphism We performed a manipulation check to investigate if the robots we used in the experiment were perceived differently according to their level of anthropomorphism. An ANOVA revealed no significant difference between the anthropomorphism ratings of the three robots, ($F(2, 622) = 2.183, p = 0.114, \eta^2 = 0.007$). We then performed a second ANOVA in which we compared only the ratings that each participant gave for the robot that they interacted with. Here again, there was no significant effect of the robot type on the perceived anthropomorphism ($F(2, 309) = 2.163, p = 0.117, \eta^2 = 0.014$). Since our manipulation check revealed that the participants did not perceive the robots to be different in terms of their anthropomorphism, we excluded the perceived anthropomorphism from the further analysis. However, as even though these different types of robot do not appear to vary on a continuum of anthropomorphism they may vary on an unmeasured construct, see footnote [3].

Shooter Bias A 2 × 3 ANOVA revealed a non-significant effect of robot color (i.e., racialization) and object on both reaction time ($F(2, 618) = 2.662, p = 0.073, \eta^2 = 0.008$), and accuracy ($F(2, 618) = 0.481, p = 0.619, \eta^2 = 0.002$). Contrary to the previous studies on robot shooter bias, there was no significant difference in shoot and don't shoot responses for robots that appeared Black or White. As the only major differing factor in this study is the inclusion of brown colored robots having ruled out social priming and task length as explanations for shooter bias, we suspect that the inclusion of another colored robot reduced shooter bias tendencies. We particularly suspect this to be the case because we would have at least expected a shooter bias with the Nao robot similar to Experiment A and previous research (Bartneck et al., 2018), but here we find no evidence of shooter bias even when considering the Nao robot separately (color by object interaction for Nao robot was ($F(2, 97) = 0.358, p = 0.700, \eta^2 = 0.007$)).

[3] A $2 \times 3 \times 3$ ANOVA revealed a significant 3-way interaction between racialization, object, and robot type on both reaction time ($F(4, 618) = 3.881, p = 0.004, \eta^2 = 0.025$) and accuracy ($F(4, 618) = 3.609, p = 0.006, \eta^2 = 0.023$) in the shooter bias paradigm. Decomposing this 3-way interaction, t-tests with Bonferroni corrected alpha revealed that there was only two significant simple effects, both toward Robosapien robots. Specifically, participants took significantly longer to refrain from shooting the Black Robosapien robot than the White ($p < .001$) or Brown Robosapien robots ($p < .001$). However, all other simple effects were non-significant across all three robots. Additionally, we tested whether there were differences in overall responses in the game toward the three types of robots and found that robot type significantly impacted upon reaction time ($F(2, 309) = 6.3962, p = 0.002, \eta^2 = 0.040$) and accuracy ($F(2, 309) = 5.248, p = 0.006, \eta^2 = 0.033$) on the shooter bias task. Post-hoc t-tests with Bonferroni corrected alpha showed that participants took significantly longer to respond to the Nao robot than the Robosapien robot ($p = 0.001$), but they were significantly more accurate when responding to the Robosapien than the Inmoov robot ($p = 0.004$).

Table 2: Means and Standard Deviations for Reaction Times and Accuracy across all Conditions

		Reaction Times			Accuracy		
		Black	Brown	White	Black	Brown	White
Gun	Inmoov	558 (58)	552 (55)	548 (58)	0.93 (0.08)	0.93 (0.08)	0.94 (0.08)
	Nao	561 (51)	559 (51)	562 (50)	0.95 (0.07)	0.97 (0.05)	0.96 (0.07)
	Robosapian	538 (55)	541 (55)	536 (54)	0.96 (0.07)	0.97 (0.06)	0.96 (0.06)
Benign Object	Inmoov	600 (54)	595 (54)	600 (53)	0.93 (0.08)	0.93 (0.08)	0.92 (0.10)
	Nao	612 (51)	613 (56)	615 (49)	0.91 (0.12)	0.92 (0.11)	0.92 (0.09)
	Robosapian	595 (51)	583 (46)	586 (51)	0.93 (0.10)	0.94 (0.10)	0.96 (0.07)

Conclusions, Limitations, and Future Work

The present research sought to address a few important goals. We first wanted to address the issue of possible social priming in the initial "Robots and Racism" research and whether player fatigue from task length had any influence on the shooter bias. Our results showed that the shooter bias effect was still present for robots racialized as Black and White when there is no social priming, and that trial length does not have an influence on this race-based shooter bias.

We then wanted to find out if the shooter bias varied across a racial spectrum from Black to White as well as a range of anthropomorphism from less to more human-like. Contrary to our expectations and prior work with multiple racial targets in a shooter bias context (Sadler et al. 2012), we found that the shooter bias towards Black robots disappeared when a brown robot was present no matter which robot type was encountered. Additionally, participants were neither faster to shoot armed robots that were darker in color, nor to not shoot unarmed robots that were lighter in color. As the main variable that had not been excluded as an influence was the addition of the brown robot, this potentially means that diversification of robots might lead to a reduction in racial bias towards them. However, future work is needed to further examine if the inclusion of a diverse range of colors on robots can indeed erase any color-based bias that emerges from robot racialization. Prior social scientific research in human-human contexts suggest that increased exposure to diversity may have both positive and negative implications for implicit and explicit biases (e.g. (Tadmor et al. 2012; Rae et al. 2015; Lai et al. 2014; Hewstone 2015)). In the context of HRI, it would be particularly valuable for future work to directly test whether exposure to robots of a diverse range of colors and perceived racialization can indeed lead people to show no differences in their implicit and explicit responses. Future work could also run the experiment with robots colored with non-human skin tones, e.g. true black and white, and primary colors such as blue, green, and red to see if there is any particular bias shown toward such robots. Interestingly, participants across several studies seem to racialize robots even when given the option to indicate that the robot does not have a race suggesting a tendency to racialize robots, but having robots in non-human colors (e.g., green or blue) ought to remove such tendencies.

And finally, we also wanted to see whether the shooter bias was influenced by how human-like the robot was. Contrary to our expectations, participants did not see the three robots as differing in their perceived anthropomorphism. Zlotowski et al. (2015) showed that humanoid robots were subject to the inversion effect, a way of measuring whether objects are implicitly recognised as human, and the idea that participants may be more hesitant to shoot a highly human-like robot compared to a machine-like robot could be explored further. Our study revealed that although there was no shooter bias shown to any of the robot types, there was a consistent difference in reaction times. They were faster to react to Robosapiens and slower to react to Nao no matter which colour the robot was. This suggests that there may be a different dimension on which these robots implicitly varied, but future work is needed to more systematically examine the potential moderating role of robot anthropomorphism on shooter bias.

One of the challenges of the present work was that the brown robot was ascribed a plurality of racial and ethnic categories making it difficult to assess if participants' responses were due to stereotypes and prejudice toward one group or another. Indeed, these could lead to contradictory implicit biases making it important for future work to disentangle shooter biases toward different groups separately (i.e., among those that perceived the brown robot to be Latino vs. Indian vs. Native American). Future studies could also examine automatic perception of race in robots rather than rely on explicit means of race ascription. For example, the methods outlined in (Freeman et al. 2010) in which mouse-tracking was used to measure how much subjects skewed towards one image or the other when asked to ascribed a race to human faces across a racial spectrum, could be one technique that allows researchers to better understand racialization in the context of robotics. Also, as the present research uses online convenience samples, future work should examine these effects in a more representative national sample. Prior research by Correll et al 2002 show that both African American and White American participants demonstrate comparable levels of shooter bias, so a more ethnically diverse sample may not show differing effects. However, a more representative sample would allow researchers to better understand whether certain characteristics moderate such effects.

References

[Bartneck et al. 2009] Bartneck, C.; Kulić, D.; Croft, E.; and Zoghbi, S. 2009. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International journal of social robotics* 1(1):71–81.

[Bartneck et al. 2015] Bartneck, C.; Duenser, A.; Moltchanova, E.; and Zawieska, K. 2015. Comparing the similarity of responses received from experiments conducted in amazon's mechanical turk to experiments conducted with traditional methods. *PLOS One* 10(4):e0121595.

[Bartneck et al. 2018] Bartneck, C.; Yogeeswaran, K.; Sparrow, R.; Wang, S.; and Eyssel, F. 2018. Robots and racism. In *ACM/IEEE International Conference on Human-Robot Interaction*, 196–204.

[Correll et al. 2002] Correll, J.; Park, B.; Judd, C. M.; and 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.

[Correll et al. 2013] Correll, J.; Ma, D. S.; Wittenbrink, B.; Bar-Anan, Y.; and Nosek, B. A. 2013. When fatigue turns deadly: The association between fatigue and racial bias in the decision to shoot. *Basic and Applied Psychology* 35:515–524.

[Freeman et al. 2010] Freeman, J. B.; Pauker, K.; Apfelbaum, E. P.; and Ambady, N. 2010. Continuous dynamics in the real-time perception of race. *Journal of Experimental Social Psychology* 46:179–185.

[Gilbert and Hixon 1991] Gilbert, D. T., and Hixon, G. J. 1991. The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology* 60(4):509–517.

[Hewstone 2015] Hewstone, M. 2015. Consequences of diversity for social cohesion and prejudice: The missing dimension of intergroup contact. *Journal of Social Issues* 71(2):417–438.

[Lai et al. 2014] Lai, C.; Marini, M.; Lehr, S.; Cerruti, C.; Shin, J.-E.; Joy-Gaba, J. A.; and Nosek, B. 2014. Reducing implicit racial preferences: I. a comparative investigation of 17 interventions. *Journal of Experimental Social Psychology: General* 143(4):1765–1785.

[Macrae, Milne, and Bodenhausen 1994] Macrae, N. C.; Milne, A. B.; and Bodenhausen, G. V. 1994. Stereotypes as energy-saving devices: A peek inside the cognitive toolbox. *Journal of Personality and Social Psychology* 66(1):37–47.

[Mekawi and Bresin 2015] Mekawi, Y., and 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.

[Ogunyale, Bryant, and Howard 2018] Ogunyale, T.; Bryant, D.; and Howard, A. 2018. Does removing stereotype priming remove bias? a pilot human-robot interaction study. *ArXiv e-prints*.

[Rae et al. 2015] Rae, J. R.; Newheiser, A.-K.; ; and Olson, K. R. 2015. Exposure to racial out-groups and implicit race bias in the united states. *Social Psychological and Personality Science* 6(5):535–543.

[Sadler et al. 2012] Sadler, M. S.; Correll, J.; Park, B.; and 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.

[Strait et al. 2018] Strait, M.; Sanchez Ramos, A.; Contreras, V.; and Garcia, N. 2018. Robots racialized in the likeness of marginalized social identities are subject to greater dehumanization than those racialized as white. In *IEEE International Conference on Robot and Human Interactive Communication*.

[Tadmor et al. 2012] Tadmor, C.; Hong, Y.-y.; Chao, M.M., W. F.; and Wang, W. 2012. Multicultural experiences reduce intergroup bias through epistemic unfreezing. *Journal of Personality and Social Psychology* 103(5):750–772.

[Tay, Jung, and Park 2014] Tay, B.; Jung, Y.; and Park, T. 2014. When stereotypes meet robots - the double-edged sword of robot gender and personality in human-robot interaction. *Computers in Human Behaviour* 38:75–84.

[Zlotowski et al. 2015] Zlotowski, J.; Bartneck, C.; Yogeeswaran, K.; and Proudfoot, D. 2015. Anthropomorphism: Opportunities and challenges in human-robot interaction. *International journal of social robotics* 7(7):347–360.