Prime-Target Salience Matching as an Alternative Driver of Implicit Bias on the Weapon Identification Task

Lindsay Tuttle

Defense Date: April 9th, 2018

University of Colorado, Boulder

Department of Psychology and Neuroscience

Committee Members:

Thesis Advisor: Dr. Joshua Correll, Department of Psychology and Neuroscience

Honors Council Representative: Dr. Eliana Colunga, Department of Psychology and Neuroscience

Outside Committee Member: Dr. Glenda Walden, Department of Sociology

Paper Overview

Implicit tasks purport to measure attitudes that explicit measures might not be sensitive to. In particular, people may hold associations that they are unaware of or unwilling to disclose. There are several such tasks, and social psychology has come to accept the assumption that they are reliable assessments of mental associations (Nosek, 2011). In contrast to that assumption, Rothermund and Wentura (2004) have suggested that semantic matches between stimuli may not be the only driver of the response patterns seen on implicit tasks. They identified one feature in particular that may partially explain bias: salience of stimuli. I seek to extend this work on semantic association by exploring salience matching in prime-target pairs. I propose that salience matched pairs are what actually allows for faster reaction time and less errors on stereotypic trials, for example a gun following a Black prime, in the Weapon Identification Task. This paper will review the ways I tested this using two different studies.

Weapon Identification Task (WIT)

In the Weapon Identification Task, participants are told they will be doing a computer task where they will be categorizing objects. They are shown images of guns and non-gun objects, and their job is to categorize them as such by pressing a key. Before each object presentation, the participants are primed with a face. It appears on the screen only briefly, and they must look at it in order to see the object that will succeed it. Typically, participants respond faster and more accurately to gun stimuli after being shown a Black prime relative to a White prime. In contrast, after a White face is shown, on average participants respond faster and more accurately to non-gun stimuli (e.g., Payne, 2001, Payne, Shimizu and Jacoby, 2005, Payne 2005, Huntsinger, Sinclair and Clore 2009).

The WIT was designed as a measure of mental associations between race and the concepts of danger or violence. The results of this experiment supported the hypothesis that the race of face stimuli influenced the response tendencies of participants (Payne, 2001). The task provided insight into the mental associations underlying several important phenomena, including police bias in shootings of Black suspects. Experiments have shown repeatedly that people tend to misidentify everyday objects as weapons when the objects are associated with Black faces, (Payne, 2005). Additionally, being prompted as to be "aware of and motivated to avoid using stereotypes" only served to increase the reaction time difference based on prime (Payne, 2002). This suggests that the more accessible racial bias is in one's associations in an attitudinal network, the stronger the effect.

The WIT poses a challenge to participants because of its speed of presentation and required response. Because of the race prime presented, the participant is unable to respond based on the actual target they saw, (Payne, 2005). This pattern is present when stereotype inconsistency is created (White face, gun or Black face, non-gun). Payne has also found that when under time restriction, Black primes caused more race specific errors than White primes. The distractor stimuli (rotated tools) presented were more likely to be misclassified as guns when primed by a Black face (Payne, 2001).

Rothermund's research indicates a semantic match between race and danger (2004). One possible explanation for the results of the original WIT experiment may be salience pairings rather than racial bias. Salience can be defined as the mental prominence or accessibility of an idea or concept. I propose that similarity in salience could be one mechanism driving biased responding on the WIT. A salience pairing would exist when the salience of a prime and a stimulus match (e.g., both are highly salient, or both are low in salience). It is the degree to

which a stimulus attracts mental attention and demands a mental response. Salience is influenced by both visual prominence and relevance. Stimuli with differences in color or configuration, have greater salience, (Guillermo, 2017).

The match between a high or low pairing could result in a faster reaction time. The question I explored was whether WIT bias could be mimicked via manipulations of salience. In other words, can bias be induced via salience matching between primes and targets? The research conducted in this experiment explored this concept of salience through two different approaches, the figure-ground model (Study 1) and illusory correlation (Study 2).

The Figure-Ground Model

The Figure-Ground model proposes that people process a scene in terms of figures and grounds, (Wagemans, 2012). Figures are objects of the scene that represent closer, more relevant, and therefore more salient objects. Grounds, in contrast, are those objects that are less salient, often farther away, or less relevant. An example of this is observing a red dot against a white background (Guillermo, 2017). The red dot in this presentation appears to pop out because of its figure properties. It may be the case that race-threat can prompt mental alertness similarly to how the contrast of a figure against a ground prompts attention. It is possible that some of the existing research on the WIT overestimates the effects of race priming by creating conditions where Black faces are highly salient compared to White faces simply due to their low-level, visual features, (Guillermo, 2017). This idea can be extended to conceptual figures and grounds. These would be objects that people's cognition assigns high relevance to in general (figures), or low relevance to (grounds). For example, most people have been conditioned to view police cars as highly salient. In contrast, streets are filled with ordinary, low cost, civilian vehicles and as such these cars have low salience.

I propose that Black faces may possess stronger conceptual "figure" properties than White faces. One reason for this could be the prevalence of negative stereotypes of Black people in the news and mainstream media. As a result, many people have been conditioned by the media to view Black people as highly salient. Additionally, most Americans are guilty of staying within their racial bubble, not knowing many or any people of another race. As a result, people of their same race are not very salient to them but people of other races are "rarer," or more salient stimuli.

The average person in relatively liberal college-student samples likely experiences guns as highly salient objects because of the abnormality of seeing a gun in their day to day life. As a result, guns act as a stimulus that puts the viewer into a higher level of mental awareness of their surroundings. The salience congruence between the rare stimuli, guns, and Black faces, creates a faster reaction time when one is used as a prime for the other. Similarly, White faces and less salient objects, such as tools or household objects, when paired stimuli, result in faster reaction time.

Study 1: Creating Salience with Contrast Manipulations

In order to remove the element of race, all of the stimuli I used were altered images of White male faces. These images were grayscaled and figureness was created by editing them to have either high or low contrast (see Figure 1) and were presented before images of guns and non-guns (similar to the original WIT). I predicted that I would see a faster reaction time when high contrast faces were paired with gun stimuli. In contrast, I also hypothesized I would observe a faster reaction time for low contrast faces paired with non-guns.

There is evidence that this racial bias can be imitated by altering images in order to have a stronger conceptual "figure" property (Rothermund and Wentura, 2004). By increasing the

contrast, an image can be created that will have stronger figure properties (relative to low-contrast faces). These will be presented alongside low contrast faces in order to create a more ground-like stimulus. It was hypothesized that when high-contrast images and gun images were paired, and the low contrast and non-gun images were paired their similar salience would result in fast reaction times. When images with asymmetrical saliencies are paired the participants will have a harder time connecting them, and this will result in a relatively slower reaction time. I wanted to explore whether the glaring appearance of high contrast faces could mimic the results of high-salience Black primes. Similarly, I wanted to explore whether faded, low-contrast images could mimic the low salience of White primes.

Figure 1. Study 1: Examples of White Male Face Primes. On the left is a grayscale "low contrast" prime, and on the right is a grayscale high contrast prime. Each face oval was edited to be used as both a high and low salience prime.

Method

Participants

The participants in this study were 67 male and 64 female students (n=131) at the University of Colorado, Boulder, enrolled in an introduction to psychology class. The students in

the course signed up for the study online and received class credit for participation. Upon arrival, they completed a 15 minute computer task, followed by a short survey collecting demographic information. 72% percent of the participants identified as White, 12% as Asian/Pacific Islander, 7% as Hispanic, 3% as Black, 4% as other and 2% as Native American. The participants were between the ages of 18 and 26 (M=19.79, SD=3.44).

Materials and Procedure

The WIT task I created was adapted from the original done by Payne (2001). The study was programmed in Javascript, HTML and PHP. Each trial began with a mask presentation for 200ms, followed by 200ms of prime presentation (high or low contrast face), and then target (Gun or non-gun) presentation for 200ms. After this, the mask reappeared for 300ms. If participants did not respond by the end of this mask, they were shown a "too slow" warning. Thus, the participant was given a 500ms response window (200ms target + 300ms mask), but responses were recorded until the end of the trial. Upon pressing the button corresponding to the correct category, the participant would see a green check mark. If the participant submitted an incorrect answer, a red “x” would appear. The inter-trial-interval, was 1000ms. Primes were 55 unique White male faces adapted into a high and low contrast version (100 stimuli total). Participants completed 220 trials, seeing each unique face 4 different times (2 high contrast and 2 low contrast). The WIT and the survey that followed were run on 2014 Mac desktops running OS X. The task was run online using Safari 7.

Design

This experiment was a 2 (prime: high vs. low contrast) x 2 (object: gun vs. non-gun) within-participants design. I analyzed the error rate of participant responses.

Results

A significant object effect was found, $F(1, 92) = -2.76$, $p = .007$, such that participants were faster and more accurate to gun objects than to non-gun objects. This is consistent with the higher salience of guns and is consistent with prior literature (Payne, 2005). There was an non-significant effect for prime, $F(1, 92) = -0.27$, $p = .78$. There was no evidence of an interaction between object and prime, $F(1, 92) = 1.13$, $p = .26$. See Figure 2 for means.

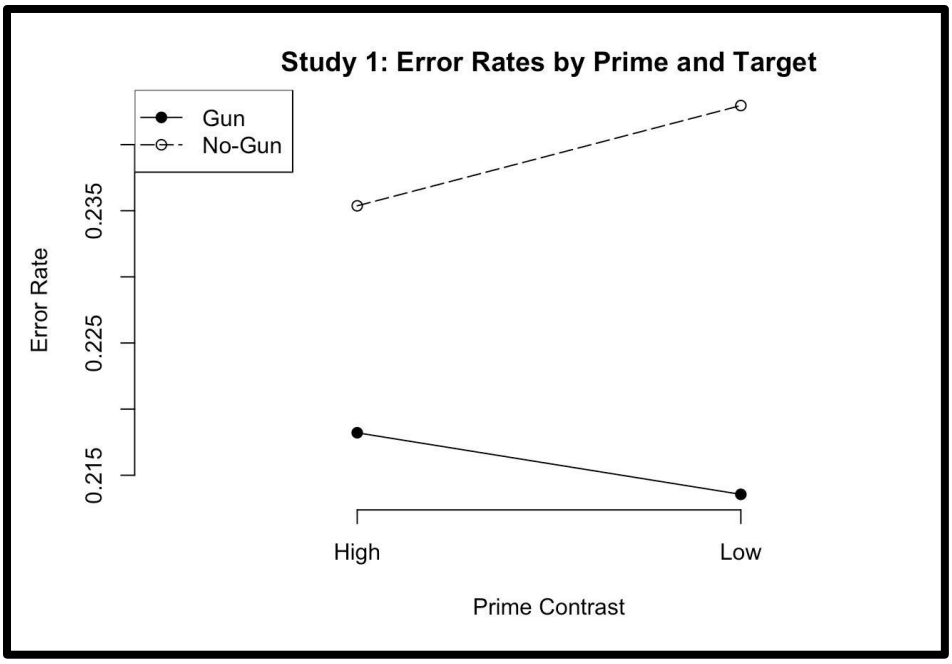

Figure 2. Study 1: graph of error rates by prime (high or low contrast faces) and target type (Gun or No-Gun objects). There was a significant object effect, such that on average, participants had fewer errors to Gun objects.

Discussion

Although there was support for an object effect, this is quite typical. In the original WIT, a object effect was found (Payne, 2001). This is consistent with more attention being paid to gun objects. An object effect can be seen in the average participant's high accuracy at identifying gun objects. We did not observe the predicted interaction between prime and target. One explanation for this could be that the manipulation simply did not work. In other words, the contrast

difference may have not caused a difference in salience. In order for a conceptual salience match between faces and guns to occur, differing conceptual salience must be perceived.

STUDY 2: Creating Salience via the Illusory Correlation

After not finding evidence of a contrast effect in Study 1, I sought a new approach to manipulating salience. Our research was inspired by the research by Hamilton and Gifford (1976) on the illusory correlation. The illusory correlation is the relationship that individuals perceive between two groups of events that have little to no correlation (Smith and Alpert, 2007). Their research suggested that the different perceptions of majority and minority groups "could result solely from the cognitive mechanisms involved in processing information about stimulus events that differ in their frequencies of co-occurrence," (Hamilton and Gifford, 1976, p. 392). This illusion manipulates salience by changing the frequency of stimuli. The research has shown that by doing this, the perception of a relationship between rare categories can develop (Hamilton and Gifford, 1976, 405). I propose that the illusory correlation could similarly be applied to the WIT.

For each participant, one color of primes was made to be rare. It was hypothesized that since guns are relatively rare objects in real life, those two rarities together should prompt salience matching. The match between the two categorizes perceived as rare generates the illusory correlation.

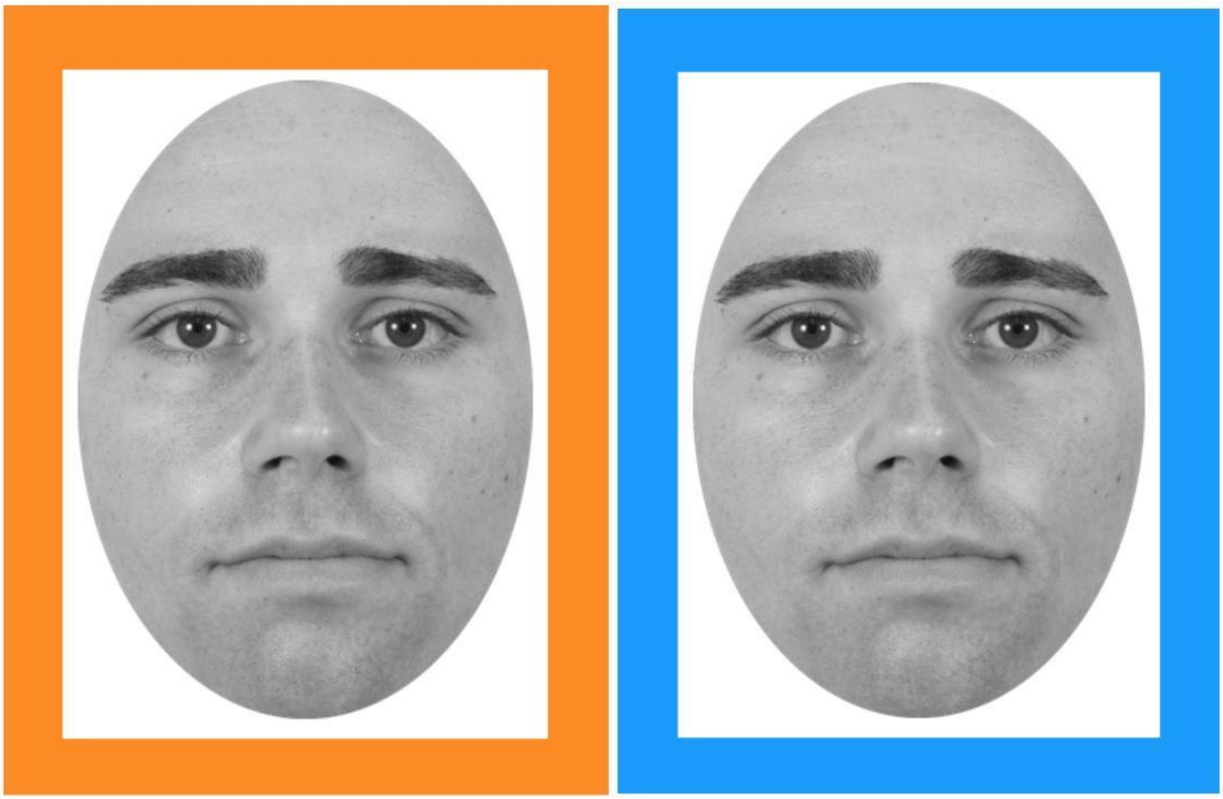

Figure 3. Study 2: examples of White male face primes with colored frames. Orange and blue were chosen for the frame colors because of their limited associations or "neutralness." Like the study 1 primes, each male face was used as both an orange and a blue stimulus. Participants were randomly assigned to see either more orange primes or more blue primes.

Method

Participants

The participants in part 2 of this study were 72 male and 118 female students (*n*=195) at the University of Colorado Boulder. Like study 1, these participants received class credit for participation. They also completed a 15 minute WIT, followed by a short survey collecting demographic information. In Study 2, 69% of the participants identified as White, 11% as Asian/Pacific Islander, 10% as Hispanic, 4% as Black, 4% as other (Indian or Middle Eastern) and 1% as Native American. All of the participants were between the ages of 18 and 23

(*M*=18.21, *SD*=1.03). 11 people were omitted from our analysis due to reporting recent alcohol or drug use.

Materials and Procedure

The task was administered in a lab setting. The task program was created via web infrastructure, including javascript, HTML, and PHP. It was accessed from the lab over the internet using Mozilla Firefox. There were a total of 35 unique White male faces. Each of these originals was adapted into both an orange-frame and blue-frame version (see Figure 3). Participants completed a total of 320 trails. Because effects based on color were uninteresting for the current questions, the frequent color-group was counterbalanced between participants. Each participant saw 70% of their primes in one color group, and the remaining 30% as the other color group. In other words, some participants saw a higher proportion of blue-framed faces and some saw a higher proportion of orange-framed faces. An even number of guns and non-gun objects were shown after each color of face. That is, 50% of orange-framed faces were followed by guns, and 50% of blue-framed faces were followed by guns regardless of which prime color was more frequently encountered overall. Each participant was told that at the end of the task they would be asked to report the percentage of guns/non-guns they saw and percentage of orange/blue framed faces they saw. This was intended to direct their attention to the color of the frames and the frequency of the stimuli in an attempt to encourage the illusory correlation. After each prime, a gun or non-gun image appeared and participants were to classify the objects.

Design

This experiment was a 2 (prime: high vs. low contrast) x 2 (object: gun vs. non-gun) x 2 (frequent prime: blue vs. orange) with the first two factors manipulated within participants and the latter manipulated between participants. I analyzed participant error rates.

Results

There was a marginally significant color effect, such that orange primes resulted in less errors, $F(1, 112) = -1.86$, $p = .065$. No evidence of an object effect was found, $F(1,112) = 0.016$, $p = .98$. The interaction effect between prime, object and frequency of each prime group was non-significant, $F(1, 111) = -0.59$, $p = .56$.

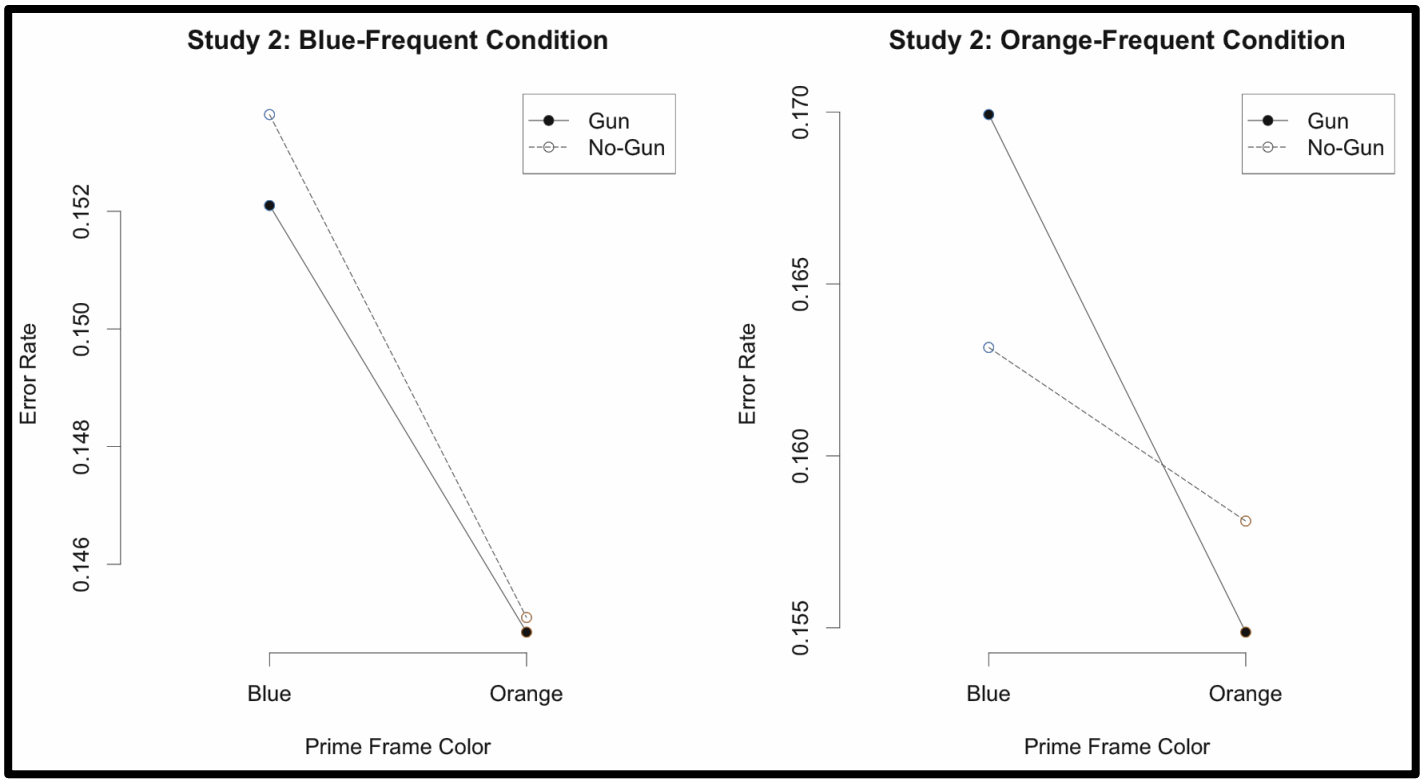

Figure 4. Study 2: error rates of prime x object x color. Orange gun and non-gun primes resulted in significantly less errors than blue gun and non-gun primes.

Discussion

Study 2 explored whether the results of the original WIT could be recreated by creating salience matches using illusory correlation manipulations. This was done by presenting 2 equivalent color groups at different frequencies. It was expected that the rarer prime group would have high salience and would result in less errors when paired with gun primes. Similarly it was hypothesized that highly frequent color primes would have low salience and would result in a lower error rate when matched with non-gun objects. The analysis revealed no significant

interaction effect between prime, object and frequency of the prime group, nor an object effect. However, there was a significant color effect such that on average, orange prime frames resulted in less errors for both gun and non-gun items. The salience manipulation using the illusory correlation was not able to mimic the results of the original WIT experiment.

Conclusion

Typically, WIT responses have been interpreted to be due to associations between the face prime and the object that follows. Both study 1 and study 2 were designed to test whether an alternative explanation, salience matching, could be the underlying mechanism. No support for this account was observed. There are different ways to interpret the results that were found. The first is that this experiment had sampling error. Another issue was that the undergraduate psychology students measured are not representative of the general population, meaning I might not have found an effect that is, in fact, reliable in the population. Another possibility is that our experiment had low construct validity. In our exploration of new areas of the WIT, it may have been the case that the manipulations of the primes I designed did not affect salience in the way I believed it would.

This research took two unique approaches to trying to measure the effect of salience differences. The results did not match the findings associated with the typical WIT although there was prior evidence to indicate this might occur. Research on the Implicit Association Task, which purports to measure a person's automatic associations between mental concepts, has found salience effects (Rothermund, 2004). That being said, the WIT is a different task, which employs different manipulation. These structural differences between the two tasks could explain our absence of significant findings. It may be worthwhile to further explore non-race salience pairings with the WIT.

It would be premature to base any conclusions on our null results. That being said, one interpretation of the findings from study 1 and 2 would be further support of the Weapon Identification Task measuring implicit racial bias. In order to rule out the idea of salience matching on the WIT conclusively, it will be important for future research to dive a bit deeper. One way to do this would be to remove face primes altogether. Instead, arbitrary primes such as fruit could be used so that the stimuli are either rare (star fruit, jackfruit, kumquat) or ordinary (apple, orange, banana). A WIT could be designed that presented these objects before gun and non-gun primes and measured the resulting error rates.

References

Guillermo, S., & Correll, J. (2017). Beyond Stereotypes: The Complexity of Race-Based Attention, 1–62.

Hamilton, D., & Gifford, R. (1976). Illusory Correlation in Interpersonal Perception: A Cognitive Basis of Stereotypic Judgments. *Journal of Experimental Social Psychology*, *12*(4), 392–407.

Huntsinger, J. R., Sinclair, S., & Clore, G. L. (2009). Affective regulation of implicitly measured stereotypes and attitudes: Automatic and controlled processes. Journal of Experimental Social Psychology, 45(3), 560–566. doi: 10.1016/j.jesp.2009.01.007

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, *47*, 1122–1135. doi: 10.3758/s13428-014-0532-5

Milner AN, George BJ, Allison DB (2016) Black and Hispanic Men Perceived to Be Large Are at Increased Risk for Police Frisk, Search, and Force. PLoS ONE 11(1): e0147158. https://doi.org/10.1371/journal.pone.0147158

Nosek, B. A., Hawkins, C. B., & Frazier, R. S. (2011). Implicit social cognition: From measures to mechanisms. Trends in Cognitive Sciences, 15(4), 152–159. doi: 10.1016/j.tics.2011.01.005

Payne, K. (2001). Prejudice and Perception: The Role of Automatic and Controlled Processes in Misperceiving a Weapon. *Journal of Personality and Social Psychology*, *81*(2), 181–192.

Payne, K., Lambert, A., & Jacoby, L. (2002). Best laid plans: Effects of goals on accessibility bias and cognitive control in race-based misperceptions of weapons. *Journal of Experimental Social Psychology, 38*(4), 384–396.

Payne, B. K., Shimizu, Y., & Jacoby, L. (2005). Mental control and visual illusions: Toward explaining race-biased weapon misidentifications. *Journal of Experimental Social Psychology, 41*(1), 36–47.

Payne, B. K. (2005). Conceptualizing control in social cognition: How executive functioning modulates the expression of automatic stereotyping. Journal of Personality and Social Psychology, 89(4), 488–503. doi: 10.1037/0022-3514.89.4.488

Payne, K. (2005). Conceptualizing Control in Social Cognition: How Executive Functioning Modulates the Expression of Automatic Stereotyping. *Journal of Personality and Social Psychology, 89*(4), 488–503.

Ratliff, K. A., & Nosek, B. A. (2010). Creating distinct implicit and explicit attitudes with an illusory correlation paradigm. *Journal of Experimental Social Psychology, 46*, 721-728.

Rothermund, K., & Wentura, D. (2004). Underlying Processes in the Implicit Association Test: Dissociating Salience From Associations. *Journal of Experimental Psychology: General, 133*(2), 139–165.

Smith, M., & Alpert, G. (2007). EXPLAINING POLICE BIAS: A Theory of Social Conditioning and Illusory Correlation. *CRIMINAL JUSTICE AND BEHAVIOR, 34*(10), 1262–1283. https://doi.org/10.1177/0093854807304484