Multidimensional Approach to Implicit Bias and the Underlying Cognitive Mechanism

by

Hansol Rheem

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved October 2019 by the
Graduate Supervisory Committee:

D. Vaughn Becker, Chair
Scotty D. Craig
Robert S. Gutzwiller

ARIZONA STATE UNIVERSITY

December 2019

ABSTRACT

Social categories such as race and gender are associated by people with certain characteristics (e.g. males are angry), which unconsciously affects how people evaluate and react to a person of specific social categories. This phenomenon, referred to as implicit bias, has been the interest of many social psychologists. However, the implicit bias research has been focusing on only one social category at a time, despite humans being entities of multiple social categories. The research also neglects the behavioral contexts in which implicit biases are triggered and rely on a broad definition for the locus of the bias regulation mechanism. These limitations raise questions on whether the current bias reduction strategies are effective. The current dissertation sought to address these limitations by introducing an ecologically valid and multidimensional method. In Chapters 1 and 2, the mouse-tracking task was integrated into the implicit association task to examine how implicit biases were moderated in different behavioral contexts. The results demonstrated that the manifestation of implicit biases depended on the behavioral context as well as the distinctive identity created by the combinations of different social categories. Chapter 3 laid groundwork for testing working memory as the processing capacity for the bias regulation mechanism. The result suggested that the hand-motion tracking indices of working memory load could be used to infer the capacity of an individual to suppress the influence of implicit bias. In Chapter 4, the mouse-tracking paradigm was integrated into the Stroop task with implicit associations serving as the Stroop targets. The implicit associations produced various effects including the conflict adaptation effect, like the Stroop targets, which suggested that implicit associations and

Stroop stimuli are handled by overlapping cognitive mechanisms. Throughout these efforts, the current dissertation, first, demonstrated that a more ecologically valid and multidimensional approach is required to understand biased behaviors in detail. Furthermore, the current dissertation suggested the cognitive control mechanism as a finer definition for the locus of the bias regulation mechanism, which could be leveraged to offer solutions that are more adaptive and effective in the environment where collaboration and harmony are more important than ever.

ACKNOWLEDGMENTS

The completion of this dissertation could not have been possible without the support I received from many people I encountered during my wonderful years at Arizona State University.

I acknowledge and thank my doctoral advisor and dissertation chair, Dr. D. Vaughn Becker for his support. I was able to achieve more than I expected, thanks to his unconditional support. Dr. Scotty D. Craig was not just an encouraging advisor but was also someone I could talk to whenever I was stressed out. Dr. Robert S. Gutzwiller always gave me inspiration for research, which helped me grow as a better researcher. Dr. Chris Blais, although he was not my doctoral advisor, sat with me for hours every week to advise me on research, which helped me push my limits.

I would also like to extend my sincere gratitude to the faculty members at the Human Systems Engineering. Dr. Nancy J. Cooke, as the Graduate Program Chair, made sure that I had the department's full support, for which I am grateful. Dr. Bing Wu motivated me to pursue a Ph.D. degree in the department and gave me crucial feedback that strengthened my research ability. It was an honor of mine to work with the brightest minds in the department.

Finally, I would like to thank my family members and friends who have been there with me and supported me spiritually. I've had many ups and downs so far and was able to make it to the "ups" and overcome the "downs" thanks to everyone.

TABLE OF CONTENTS

Page

LIST OF TABLES vi

LIST OF FIGURES vii

CHAPTER

1 INTRODUCTION 1

Measures of Implicit Bias 2

Intersectionality of Gender, Race, and Emotion 5

The Current Dissertation 7

2 CHAPTER1 9

Overview: Experiments 1 & 2 9

Method 10

General Discussion 25

3 CHAPTER2 33

Overview: Experiments 3 & 4 33

Method 34

General Discussion 47

4 CHAPTER3 51

Overview: Experiments 5 & 6 51

Method 55

General Discussion 79

5 CHAPTER4 85

CHAPTER Page

Overview: Experiments 7 & 8 85

Method 90

General Discussion 109

6 FURTHER DISCUSSIONS AND CONCLUSIONS 116

Significance of the Study and Application 116

Limitations 118

Possible Extensions of the Study 124

Summary and Conclusions 130

REFERENCES 133

APPENDIX

A CONGRUENCY EFFECTS IN EXPERIMENT 1 145

B CONGRUENCY EFFECTS IN EXPERIMENT 2 147

C CONGRUENCY EFFECTS IN EXPERIMENT 3 149

D CONGRUENCY EFFECTS IN EXPERIMENT 4 151

E MEAN RESPONSE TIMES AND AUCS FOR EXPERIMENT 5 153

F MEAN RESPONSE TIMES AND AUCS FOR EXPERIMENT 6 155

G LIST OF HAND MOTION FEATURES FED TO SVM CLASSIFIERS 157

H CONGRUENCY EFFECTS IN EXPERIMENT 7 159

I CONGRUENCY EFFECTS IN EXPERIMENT 8 161

J IRB PERMISSION FOR HUMAN SUBJECT TESTING 163

LIST OF TABLES

Table Page

1. Comparisons of Congruency Effects in the Emotion and Gender Tasks 114

2. Summary of the Dissertation 117

LIST OF FIGURES

Figure Page

1. Example of the Mouse-Tracking Task Used in Experiments 1 and 2 13

2. Descriptive Data for Experiment 1 19

3. Descriptive Data for Experiment 2 21

4. Interaction Plots (RT) for Experiment 1 22

5. Interaction Plots (RT) for Experiment 1 23

6. Interaction Plots (AUC) for Experiment 1..23

7. Interaction Plots (RT) for Experiment 2 24

8. Interaction Plots (AUC) for Experiment 2..25

9. Descriptive Data for Experiment 3 37

10. Descriptive Data for Experiment 4 42

11. Interaction Plots (RT) for Experiment 3 43

12. Interaction Plots (AUC) for Experiment 3 44

13. Interaction Plots (RT) for Experiment 4 45

14. Interaction Plots (RT) for Experiment 4 46

15. Interaction Plots (AUC) for Experiment 4 46

16. Experimental Setting in Experiments 5 and 6 58

17. The subjective Rating Scale Created Based on Paas (1991) 58

18. Examples of the Hand Motion Features Fed to SVM Classifiers 62

19. Descriptive Data for Experiment 5 64

20. Mean Trajectories as a Function of Orientation, WM load, and Size 66

Figure Page

21. Mean Velocities as a Function of Orientation, WM load, and Size 68

22. Mean Trajectory Angles as a Function of Orientation, WM load, and Size 68

23. Mean Velocities and Angles at Each Time Step by the Condition 69

24. Stacked Bar Chart of Subjective Mental Effort Ratings in Experiment 6 71

25. Descriptive Data for Experiment 6 72

26. Mean Trajectories as a Function of Orientation, WM load, and Size 73

27. Mean Velocities as a Function of Orientation, WM load, and Size 75

28. Mean Trajectory Angles as a Function of Orientation, WM load, and Size 75

29. Mean Velocities and Angles at Each Time Step by the Condition 76

30. Two-Way Interaction of Target Size and WM Load in Experiment 5 81

31. Illustration of a Conflict Adaptation Effect Under Different WM Load 89

32. Example of a Trial Sequence in Experiments 7 and 8 92

33. Stacked Bar Chart of Subjective Mental Effort Ratings in Experiment 7 95

34. Mean RTs as a Function of Conditions in Experiment 7 97

35. Mean AUCs as a Function of Conditions in Experiment 7...............................100

36. Mean Conflict Adaptation Effects at Each Time Step 101

37. Stacked Bar Chart of Subjective Mental Effort Ratings in Experiment 8 103

38. Mean RTs as a Function of Conditions in Experiment 8 104

39. Mean AUCs as a Function of Conditions in Experiment 8...............................106

40. Mean Conflict Adaptation Effects at Each Time Step 107

41. Anatomy of the Conflict Adaptation Effects 124

INTRODUCTION

As members of the diverse world, we meet new people every day, who might be a potential friend, foe, or a random person just passing by. Among these potential relationships, how do people distinguish a potential friend from a foe, so that they can decide to approach or avoid people they meet? Previous research has consistently shown that different races or genders are readily identified and can give rise to misunderstandings (Allport, 1954; Billig, 1985; Ehrlich, 1973; Tajfel, 1981; Devine, 1989; Smith & Branscombe, 1984). In return, these associations, whether stereotypical or not, affect how people form impressions of others (Branscombe & Smith, 1990; Olson & Fazio, 2004) and guide behaviors in reaction to others (Amodio & Devine, 2006). Even in the absence of explicit biases, these associations can give rise to implicit biases, which can be triggered automatically and without awareness (Devine, 1989; Fazio, 1990; Greenwald & Banaji, 1995; Jacoby, 1991; Payne, 2005), thereby reducing the time required to evaluate a person.

One-hundred years have passed since the first American woman was elected to serve in the U.S Congress (Geggel, 2017). Nevertheless, implicit bias is still affecting our society. Women earn only 80 percent of what men earn (Preble, 2017), hold less than 25 percent of total seats both in the Senate and the House of Representatives (Cohn, 2016). The gender inequality, in return, poses threats to the entire society. Research suggests that a 50 percent increase in the gender wage gap leads to a 35 percent decrease in income per capita (Cavalcanti & Tavares, 2015). Moreover, Hsieh et al. (2012) argued that the aggregate productivity gains in U.S. between 1960 and 2008, was negatively correlated with the level of discrimination the African Americans and women faced in the labor market. Despite the issues induced by implicit bias, there is limited understanding of implicit bias and strategy to reduce it. This motivated the current dissertation project to review previous implicit bias studies and their methods for assessing implicit bias. In the process, the current dissertation suggests an alternative method for measuring implicit bias: the mouse-tracking method. Furthermore, the mouse-tracking method is leveraged to unveil new findings which may contribute to the development of new strategies for reducing bias that are less-effortful and more enjoyable compared to the current strategies.

Measures of Implicit Bias

Several methods have been proposed as measures of implicit bias. The implicit association test (IAT), for example, was used to measure the impact of implicit racial bias on human behavior (Greenwald et al., 1998). In the study by Greenwald and his colleagues, Korean American, and Japanese American participants were asked to categorize target words which were either a name typically used within a particular ethnic group (i.e. Korean, Japanese) or a word eliciting a certain emotion (i.e. pleasant, unpleasant). These two types of words were presented in alternate trials and participants were instructed to categorize these words using their both hands. The researchers found a delayed response when the same hand was used to respond to both an outgroup name and a pleasant word, compared to when different hands were used for the name and the word.

Similar to the IAT, the evaluative priming task is another method that measures implicit bias. In the study by Fazio et al. (1995), participants viewed the face of an ingroup or outgroup member, which was followed by a negative or positive adjective word. Participants were instructed to categorize the adjective either as pleasant or unpleasant. The result showed that participants responded slower when the positive target was primed by an outgroup face than when it was primed by an ingroup face. In contrast, the response to the negative target was slower when it was primed by an ingroup face than when it was primed by an outgroup face. Altogether, the findings from the IAT and the evaluative priming task converge to suggest that the implicit bias toward a different race can be operationally defined as the delayed response times (RT) to a pair of stimuli that have conflicting meanings or relationships, compared to those that have non-conflicting meanings or relationships.

Experimental tasks requiring button presses have been useful in measuring stereotypes (Fazio et al., 1995; Greenwald et al., 1998; Payne, 2001). However, other lines of research that recorded more complex responses suggested that the simple button press tasks might not be sensitive enough to capture the dynamic nature of the behaviors occurring in real-life social interactions. In the study by Marsh, Ambady, and Kleck (2005), participants were instructed to categorize angry faces and fearful faces by moving a joystick toward or away from themselves. The result indicated that participants were faster to categorize angry faces when they had to push the joystick away from themselves compared to when they had to pull the joystick toward themselves. In contrast, when categorizing fearful faces, participants were faster to move the joystick toward themselves compared to when they pushed the joystick away. Based on these findings, Marsh et al. claimed that behavioral and situational contexts could modulate human behavior. That is, an angry face triggered an avoidance response, which, in return, facilitated the movement of pushing the joystick away because the angry face alluded the existence of an immediate threat. They also argued that a fearful face triggered an approach response, and thus facilitated the movement of pulling the joystick inward because the fearful face cued the need for nurturance. In line with this finding, Rinck and Becker (2007) also demonstrated that highly spider-fearful participants were faster to push the joystick away from themselves compared to when they pulled it toward themselves when they had to detect and report whenever a spider was presented. Interestingly, the control group participants showed a reversed pattern in which the pulling behavior was faster than the pushing behavior. Altogether, these results suggested that there are various uncovered factors that have the potential to affect implicit bias, such as behavioral contexts or participant's internal state of mind, which require more sensitive and context-appropriate measures to monitor them.

Another recent approach has allowed for an inspection of the temporal dynamics behind implicit bias mechanisms. Using a computer mouse-tracker program, Freeman et al. (2008) investigated implicit biases against certain gender groups. In this study, participants were presented with faces that differed in the extent to which they had prototypical features of a specific gender. For example, male faces with feminine facial features or female faces with masculine features (atypical faces) were presented on the half of the trials, with male faces with masculine facial features and female faces with feminine features (typical faces) presented on the other half. Participants had to categorize these faces by moving a mouse cursor to either the word "Male" or the word

"Female" presented on the upper side corners of the screen. Freeman et al. found that the mouse trajectory deviated and took a longer path when participants categorized the atypical faces than when they categorized the typical faces. Freeman et al. inferred that the trajectory deviation occurred because the atypical faces had features that can be, in part, described by the opposing nontarget words, thereby distracting participants to a greater extent compared to when they were categorizing the typical faces. Many other mouse-tracking studies have been done, which demonstrated that the mouse-tracking method could be used to make inferences about biases against racial groups (Freeman et al., 2010), as well as about ingroup biases (Lazerus et al., 2016), and to examine the interaction of racial and gender biases (Johnson, Freeman, & Pauker, 2012).

Intersectionality of Gender, Race, and Emotion

Social interactions where implicit biases are triggered, involve online exchanges of behaviors and spoken words between an exhibitor of bias, and a target. On the exhibitor's side, exhibitors of biases demonstrate various behaviors such as approaching, avoiding, giving, and taking. The meaning of these behaviors changes dynamically as they interact with various behavioral contexts. For instance, giving a flower to others is a socially favorable gesture, whereas taking it away from others is unfavorable. In contrast, changing the object given or taken to something unpleasant, like a spider, changes the meaning of the gesture substantially.

On the target's end, the targets usually have multiple identities. A male person we encounter in the street is not just an individual that is male, but for example, might also be a Caucasian American, a graduate student, and a homosexual, each of which is associated with different stereotypes. Despite its diverseness, each stereotype seems to be tied to the common underlying dimension of gender (Johnson et al., 2012; Wiggins, 1991). Black faces, for example, are perceived as more masculine than white faces (Goff et al., 2008). White faces and women are often perceived to have similar levels of competence and warmth compared to other gender or racial groups (Fiske et al., 2002). Asians and women are also perceived as having traits of shy, family-oriented, and soft-spoken, thus are associated closely (Bem, 1974; Devine & Elliot, 1995; Ho & Jackson, 2001). It is also argued that two categories can facilitate or interfere with the social categorization of a face depending on whether two categories are associated with the same gender dimension or not (e.g. black male, Asian female; Johnson et al., 2012). Johnson et al. (2012) demonstrated this by instructing participants to categorize computer-generated faces that were gradually varying along the gender and racial dimensions. Participants used a mouse cursor to click one of the two response options appearing at the upper corners of the screen. The results indicated that participants showed more efficient cursor trajectories and faster response times to the response options when they were categorizing faces of overlapping identities like a black male or an Asian female face, compared when categorizing faces of non-overlapping identities like a black female and an Asian male face. Similarly, Adams et al. (2015) suggested that people implicitly consider males to be powerful and consider females to be powerless. In line with this argument, people are usually faster and more accurate to identify males as angry, angry faces as males, females as happy, and happy faces as females (Becker et al., 2007). In addition to the gender and racial dimensions, targets often express various emotions which are used as a proxy for the motivations that the targets might have (Frank, 1988; Izard, 2013; Plutchik, 1980), which can also affect social categorization.

These identities and emotional expressions are like multiple sides of a dice, which cannot be selectively attended or ignored, and which influence human behaviors as a sum (Becker, 2017; Martin et al., 2015). For example, African American women face racism and sexism because they are black people and women. However, their overall experience of discrimination as a black woman may not be understood if one tried to understand the racial and gender discriminations separately (Crenshaw, 1993). Despite the rich dimensions of identities people use to judge others, traditional research has tended to look at one identity dimension at a time (Johnson et al., 2012) and has therefore defined the implicit bias only in terms of the dimension of interest. However, it is likely that efforts to understand only a single identity dimension and its influence on implicit bias will lead to the failure to capture the complex bias mechanisms.

The Current Dissertation

The previous experimental paradigms indeed have been useful in revealing the influence of implicit bias on human behavior and shedding light on the underlying mechanisms of implicit bias. At the same time, more limitations and confounds had been revealed that need to be addressed. First, converging evidence suggested that studies using a simple button press task might reveal no more than a snapshot of biased behaviors, which can prevent researchers from gaining a better understanding of the implicit bias mechanism. Moreover, such studies lacked ecological validity not only in terms of the bias measurement but also in that they deemed implicit bias as a one-way phenomenon which is mostly dependent on exhibitors of biases, rather than a dyadic phenomenon elicited by the exhibitor-target relationship. Third, little work has been done to perform a fine-grained pinpointing of the cognitive mechanisms that might be responsible for regulating implicit biases, which could help develop strategies to attenuate them. The current dissertation aimed to address these limitations and confounds by 1) using the mouse-tracking method to monitor more complex social behaviors than the behaviors of pressing a button, 2) examining the influence of behavioral contexts and the identity intersectionality effects using the mouse-tracking method, and 3) testing the hypothesis that working memory and cognitive control mechanisms are involved in handling implicit biases.

In Chapters 1 and 2, a mouse-tracking task was integrated into an implicit association task to examine how implicit biases were moderated in different behavioral contexts and as a function of target identities. Chapter 3 laid groundwork for testing working memory as the processing capacity for bias regulation mechanisms. Lastly, in Chapter 4, the mouse-tracking task was integrated into a Stroop task with implicit associations serving as Stroop targets. Particularly, the cognitive mechanisms responsible for regulating implicit bias were sought. Throughout Chapters 1 and 4, arguments are made about important components that have been missing from the previous implicit bias research and how new pieces of information revealed by the novel mouse-tracking method can be leveraged to develop alternative strategies for reducing implicit bias, as well as directions for future studies.

CHAPTER 1

Overview: Experiments 1 & 2

Experiments 1 and 2 in Chapter 1 were conducted to investigate the influence of the gender-race intersectionality on the social categorization of faces. Furthermore, how behavioral contexts like giving and taking affected the social categorization was examined. To this end, a modified version of the IAT was integrated into the mouse-tracking paradigm. In this version of the mouse-tracking paradigm, participants moved a mouse cursor to give and take pleasant or unpleasant objects. This manipulation was intended to simulate motivational behaviors people display during real-life social interactions (Cacioppo & Berntson, 1994; Lang, Bradley, & Cuthbert, 1997). In Experiments 1 and 2, male participants who identified themselves as not being of African American background engaged in tasks that required them to give and take a flower or a spider, to and from, target faces of different genders and races (see Figure 1). Specifically, participants were instructed to move a mouse cursor to one of the two faces that appeared at the upper corners of the screen and retrieve the cursor to the starting location. The shape of the mouse cursor turned into a flower or a spider, cueing participants as to which faces that they should approach. In Experiment 1, participants gave/took a cursor to/from female faces if the cursor turned into a flower, and male faces if it turned into a spider. The same participants then performed an identical task in Experiment 2, with an exception that the decisions were about the race of the faces, rather than the gender. This design was chosen to explore whether the task would replicate previous findings that males and racial outgroups are stereotypically associated with negative emotion whereas females and racial ingroups are associated with positive emotion (Greenwald et al., 1998; Hess, Sabourin & Kleck, 2007). The design goes beyond this, though, by looking at how this is moderated by different kinds of giving and taking.

Method

Participants

Forty male undergraduate students at Arizona State University, Polytechnic campus, who identified themselves as being of European American or Asian background participated in each experiment (35 Caucasian Americans and 5 Middle Easterners; Mean age: 21.1 years). The gender decision task in Experiment 1 required participants to categorize the gender of target faces. Previous research suggested that participants' gender can affect the categorization of target gender (e.g. Johnson et al., 2012). Moreover, the Polytechnic campus had a small population of female students. Therefore, only male participants were recruited in order to test a homogeneous participant sample in terms of gender. All participants had normal or corrected-to-normal vision and were right-handed. All participants were offered 1-course credit for their participation.

Stimuli

Eight faces from the nimStim stimulus set (Tottenham et al., 2009) were used. Two faces were selected from each gender (male, female) and race (Caucasian American, African American) category. Each face had a closed mouth and displayed a neutral emotion. The brightness and contrast level of the images were equated, and the upper and the lower peripheral area of each image was cut, resulting in the images with a size of

506 px by 330 px. All images were then converted into grayscale. Icons of a flower and a spider (32 px by 32 px) were used as the shapes for the cursor, which changed randomly on every trial. The faces and the cursor icons were presented on a computer monitor using a program developed in JavaScript. It should be noted that only eight faces were used as stimuli in Experiments 1 and 2. The size of the stimulus set used here might not have been enough to cancel out the noise effects introduced by outlier faces, which might not fully represent the social categories of interest. However, this was a necessary measure to limit the experiment duration to a maximum of 30 minutes, especially given the rigorous counterbalancing applied in the experiments.

Procedure

Participants in this study performed two tasks in random order: The gender decision task (Experiment 1), and the race decision task (Experiment 2). Upon arrival, participants signed a consent form and were briefed with the instructions for the tasks. In the first half of the study, participants performed two 64-trial blocks of the gender decision task or the race decision task, with 10 practice trials at the beginning of each block, and with a 1-minute break in between the blocks. In the second half of the study, participants went through the same procedure but performed the task which was not assigned in the first half. If the participants performed the gender decision task in the first half, they performed the race decision task in the second half and vice versa. On each trial, participants were required to click a start button located at the bottom-center of the screen to start a trial. After participants clicked the start button, the shape of the mouse cursor changed randomly to a flower or a spider, and two faces appeared at the upper sides of the screen. In the congruent block, participants were asked to identify the shape of the cursor and move the cursor to a stereotypically congruent face (e.g. spider-male/flower-female in the gender task, and spider-black/flower-white in the race task) and retrieve the cursor back to the start button. When participants reached the target faces, a red 'Stop' sign appeared at the location of the start button, which signaled participants to wait before retrieving the cursors. The stop button appeared on the screen for random durations that ranged from 1000 ms to 1750 ms. In the incongruent block, the cursor-face mapping was reversed, guiding participants to reach faces with stereotypically incongruent cursor shapes. The order of the congruent and incongruent blocks was counterbalanced across participants as well as the order of the tasks.

Analysis

Response time and the trajectory deviation were analyzed as dependent variables of interest. First, a total of 212 incorrect trials in the gender decision task was excluded from further analysis, which comprised 4.1 percent of the entire 5120 trials. In the race decision task, 227 incorrect trials were excluded, which comprised 4.4 percent of the entire trials. The remaining trials were then preprocessed to exclude outlier data that were greater than three standard deviations above the mean, or less than three standard deviations below the mean. Moreover, the same criteria were applied again on the number of the cursor trajectory flips (the movement against the direction of the target) to exclude aberrant trajectories. The 3SD criterion resulted in the exclusion of 426 trials (8.3 %) in the gender decision task and 460 trials (9 %) in the race decision task. The RT data were measured by calculating the time that participants took to move a cursor from the start button to a target face (give), and by calculating the time participants took to retrieve the cursor to the start button after the stop sign disappeared (take). The participant's mouse cursor trajectory deviation was calculated using the method used by Freeman et al. (2008). Specifically, all trajectories were resampled to 101 time-steps using linear interpolation. After the resampling, all trajectories were remapped rightward and rescaled to a 1 by 1.5 coordinate plane. The area between each resampled trajectory and a straight trajectory that connected the starting x, y coordinates, and the terminal x, y coordinates of each trajectory, was computed to measure the deviation of the participant's trajectory from an ideal, linear trajectory (AUC).

Using the RT and trajectory data as dependent variables, five-way repeated-measures ANOVAs were performed with face gender (male, female), face race (white, black), congruency (congruent block, incongruent block), behavior (give, take), and distractor identity (same or different race as the target in Experiment 1, gender in Experiment 2) as the within-subjects factors.

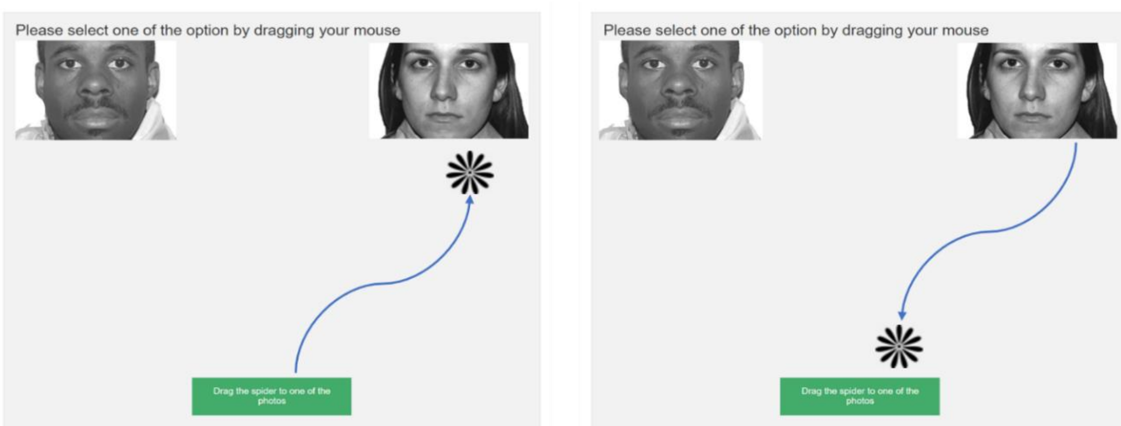

Figure 1. Example of the mouse-tracking task used in Experiments 1 and 2; Left: A giving trial in the gender decision task; Right: A taking trial in the gender decision task.

Experiment 1 Results

RT. Descriptive RT and trajectory deviation data are shown in Figure 2 but refer to Figures 4, 5, and 6 to examine significant interactions that were found in Experiment 1. Main effect of congruency was significant, $F(1, 39) = 7.89$, $p < .01$, $\eta_p^2 = .17$. Participants' responses were slower on incongruent trials ($M = 677$ ms) than on congruent trials ($M = 628$ ms). Main effect of race was significant, $F(1, 39) = 5.49$, $p < .05$, $\eta_p^2 = .12$, because participants' were slower to respond to a black face ($M = 660$ ms) than to a white face ($M = 644$ ms). Main effect of behavior was also significant, $F(1, 39) = 77$, $p < .0001$, $\eta_p^2 = .66$, as participants were about twice slower to give ($M = 870$ ms) than to take ($M = 433$ ms) cursors.

In return, gender interacted with race, $F(1, 39) = 21.3$, $p < .0001$, $\eta_p^2 = .35$. Tukey's HSD test at each gender level suggested that participants responded faster to white female targets compared to black female targets, $p < .0001$, and to black male targets than to white male targets, $p = .05$. Gender also interacted with distractor race, $F(1, 39) = 4.83$, $p < .05$, $\eta_p^2 = .11$. Simple effect analyses revealed that participants were slower to respond to a male target when it was presented with a distractor of a different race (black female) than when it was presented with that of the same race (white female), $p < .05$. In contrast, the mean RTs to female faces did not differ significantly as a function of distractor race, $p = .39$. An interaction between congruency and behavior was obtained, $F(1, 39) = 9.47$, $p < .01$, $\eta_p^2 = .19$, which indicated that congruency effect was present only on giving trials, $p < .01$, but not on taking trials, $p = .49$. An interaction between race and behavior was found, $F(1, 39) = 4.88$, $p < .05$, $\eta_p^2 = .11$. Further analyses revealed that this interaction was driven by facilitated responses to white targets on giving trials compared to black targets, $p < .05$. However, the RT difference on taking trials was nonsignificant between target races, $p = .87$.

A three-way interaction of gender, race, and distractor race was significant, $F(1, 39) = 4.93$, $p < .05$, $\eta_p^2 = .11$. Further analyses at each gender level revealed a marginal two-way interaction of target race and distractor race for male targets, $F(1, 39) = 3.03$, $p = .089$, $\eta_p^2 = .07$, but not for female targets, $F(1, 39) = 1.64$, $p = .21$. The two-way interaction was found on trials with a male target because participants responded faster to white male targets presented with a distractor of the same race (white female), compared to the same targets presented with a distractor of a different race (black female), $p < .01$. In contrast, the RT to black male faces did not differ as a function of distractor race, $p = .83$.

Three-way interaction of gender, congruency and behavior was obtained, $F(1, 39) = 6.77$, $p < .05$, $\eta_p^2 = .15$. Separate two-way ANOVAs at each behavior level revealed a nonsignificant interaction of gender and congruency on giving trials, $F(1, 39) = 1.06$, $p = .31$, and a significant interaction on taking trials, $F(1, 39) =23$, $p < .0001$, $\eta_p^2 = .37$. On giving trials, only the main effect of congruency was significant, $F(1, 39) = 9.08$, $p < .01$, $\eta_p^2 = .19$, suggesting longer RTs on incongruent trials regardless of the target gender. Nevertheless, on taking trials, a reversed congruency effect (longer RTs on congruent trials) was obtained in response to a female target, $p < .0001$, while a congruency effect was obtained in response to a male target, $p < .05$, which contributed to the three-way interaction.

A three-way interaction of gender, race, and behavior was also significant, $F(1, 39) = 34.9$, $p < .0001$, $\eta_p^2 = .47$. Separate two-way analyses at each behavior level replicated the two-way interaction of gender and race on giving trials, $F(1, 39) = 29.7$, $p < .0001$, $\eta_p^2 = .43$, but not on taking trials, $F(1, 39) = 1.31$, $p = .26$. Again, participants were faster to give a cursor to a white female face than to a black female face, $p < .0001$, and to give a cursor to a black male face than to a white male face, $p < .05$. Nevertheless, this difference was not apparent on taking trials.

Lastly, a four-way interaction of gender, race, distractor race, and behavior was present, $F(1, 39) = 7.9$, $p < .01$, $\eta_p^2 = .17$. Two three-way ANOVAs at each behavior level revealed a three-way interaction on giving trials, $F(1, 39) = 6.98$, $p < .05$, $\eta_p^2 = .15$, but not on taking trials, $F(1, 39) = .86$, $p = .36$. Giving trials were examined further using two two-way ANOVAs at each gender level, which revealed a marginal two-way interaction of race and distractor race on female target trials, $F(1, 39) = 3.89$, $p = .058$, $\eta_p^2 = .091$. This two-way interaction occurred because, on giving trials where a white female target appeared, participants took a longer time to give a cursor when the target appeared with a distractor of the same race (white male), than when it appeared with that of a different race (black male), $p < .01$. The race of the distractor did not have a significant effect on the mean RT when the target face was that of a black female, $p = .82$.

AUC. Main effect of behavior was significant, $F(1, 39) = 101$, $p < .0001$, $\eta_p^2 = .72$, because participants showed greater deviations when giving ($M = .23$ unit) than when taking ($M = .018$ unit) cursors.

A two-way interaction of gender and race was obtained, $F(1, 39) = 4.57$, $p < .05$, $\eta_p^2 = .10$. Simple effect analyses at each gender level suggested that participants showed less deviation when they responded to a white female target compared to when they responded to a black female target, $p < .01$. Participants also showed a numerically smaller deviation when they responded to a black male target compared to a white male target, although the difference in the deviation was nonsignificant, $p = .45$. Race interacted with distractor race, $F(1, 39) = 6.87$, $p < .05$, $\eta_p^2 = .15$. When a white face was presented as a target, participants showed greater deviations when it was presented along with a different-race distractor, than when it was presented with a same-race distractor, $p < .05$. When a black face was presented, the deviation did not differ significantly regardless of the race of the distractor face, $p = .19$.

In return, race interacted with gender and congruency, $F(1, 39) = 4.47$, $p < .05$, $\eta_p^2 = .10$. Separate two-way ANOVAs at each gender level yielded a significant two-way interaction on female target trials, $F(1, 39) = 4.79$, $p < .05$, $\eta_p^2 = .11$, but not on male target trials, $F(1, 39) = .23$, $p = .64$. Simple effect analyses were performed on the female target trials only, which indicated that a marginal congruency effect was obtained for black female targets, $p = .10$, but not for white female targets, $p = .43$.

A three-way interaction of race, distractor race, and congruency was significant, $F(1, 39) = 4.2$, $p < .05$, $\eta_p^2 = .097$. Separate two-way ANOVAs at each race level suggested that a two-way interaction of distractor race and congruency was absent on both white target trials, $F(1, 39) = 1.44$, $p = .24$, and black target trials, $F(1, 39) = 2. 23$, $p = .14$. Simple effect analyses were performed to investigate why the three-way interaction was obtained despite the nonsignificant two-way interactions found on both white and black target trials. The result indicated that the congruency effect was absent on every condition, *ps* > . 17. However, a trend of a reversed congruency effect was obtained on white target trials presented with a different-race distractor, $p = .38$, and a trend of positive congruency effect was obtained on black target trials presented with a different-race distractor, $p = .17$, which might have been the reason behind the significant three-way interaction.

A three-way interaction of behavior, gender, and race was obtained, $F(1, 39) = 16.7$, $p < .001$, $\eta_p^2 = .30$, and showed that a two-way interaction of gender and race was significant on giving trials, $F(1, 39) = 15.1$, $p < .001$, $\eta_p^2 = .28$, but not on taking trials, $F(1, 39) = 1.09$, $p = .30$. Further simple effect analyses on giving trials indicated that participants showed smaller deviations when they responded to a white female target than when they responded to a black female target, $p < .001$. Participants also showed a marginally smaller deviation when they responded to a black male target than to a white male target, $p = .08$.

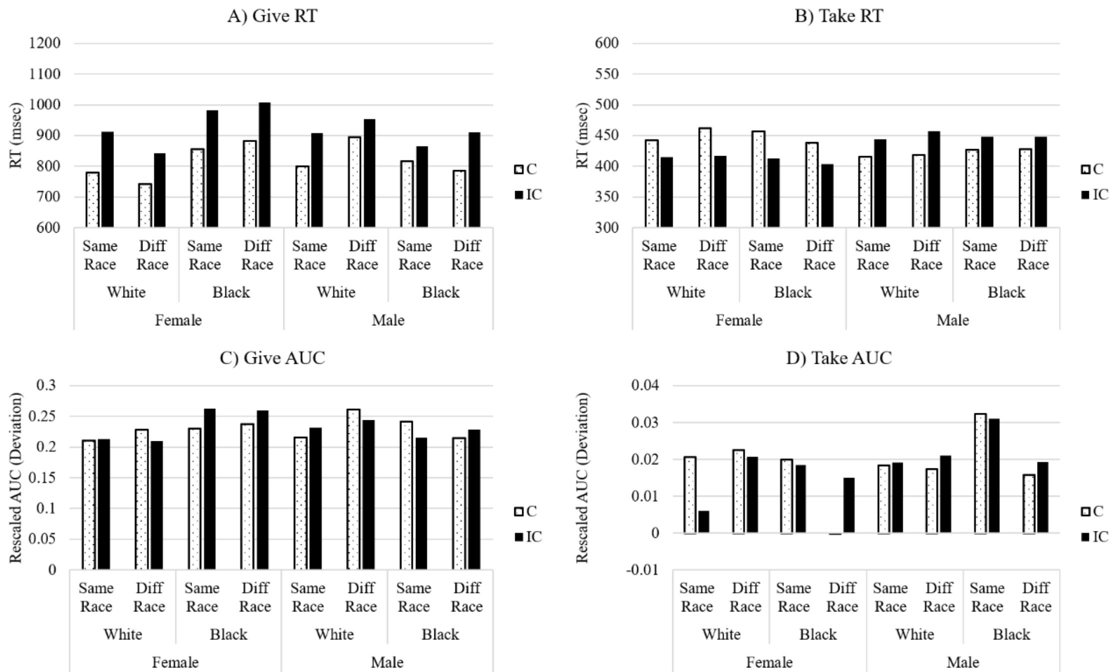

Figure 2. Descriptive data for Experiment 1. A) Give RT. B) Take RT. C) Give AUC. D) Take AUC.

Experiment 2 Results

RT. Descriptive RT and trajectory deviation data are shown in Figure 3 but refer to Figures 7 and 8 to examine significant interactions that were found in Experiment 2. Main effect of behavior was significant, $F(1, 39) = 63.1$, $p < .0001$, $\eta_p^2 = .62$, because mean RT was longer on giving trials (M = 844 ms) than on taking trials (M = 438 ms).

Furthermore, a three-way interaction was obtained between gender, race and distractor gender, $F(1, 39) = 4.76$, $p < .05$, $\eta_p^2 = .11$. Separate two-way analyses at each race level yielded a nonsignificant interaction of gender and distractor gender on white target trials, $F(1, 39) = .75$, $p = .30$, and a significant interaction on black target trials, $F(1, 39) = 5.28$, $p < .05$, $\eta_p^2 = .12$. Further, simple effect analyses on black target trials suggested that participants were slower to respond to a black male target when it was presented with a same-gender distractor (white male) than when it was presented with a different-gender distractor (white female), $p < .05$. However, the RT difference was nonsignificant on trials with a black female, a white male, and a white female targets, *ps* > .24.

Race also interacted with behavior and congruency, $F(1, 39) = 7.19$, $p < .05$, $\eta_p^2 = .16$. Two-way analyses at each behavior level revealed a nonsignificant interaction of gender and congruency on giving trials, $F(1, 39) = 1.64$, $p = .21$, and a significant interaction on taking trials, $F(1, 39) = 29.6$, $p < .0001$, $\eta_p^2 = .43$. Simple effect analyses uncovered a reversed congruency effect (longer RT on congruent trials compared to incongruent trials) on trials where participants were taking a cursor from a white target, $p < .001$, and a positive congruency effect on trials where participants were taking a cursor from a black target, $p < .05$.

AUC. Again, main effect of behavior was significant, $F(1, 39) = 76.6$, $p < .0001$, $\eta_p^2 = .66$, because participants showed greater deviations on giving trials ($M = .21$ unit) than on taking trials ($M = .02$ unit).

A two-way interaction of gender and race was obtained, $F(1, 39) = 4.94$, $p < .05$, $\eta_p^2 = .11$. Similar to Experiment 1, participants showed smaller deviations when responding to a white female target than to a black female target, $p < .05$. Contrarily, the difference in deviation was nonsignificant between different races on male face trials. Lastly, race interacted with congruency, $F(1, 39) = 5.36$, $p < .05$, $\eta_p^2 = .12$. Simple effect analyses at each race level indicated that a marginal reversed congruency effect was obtained on white face trials, $p = .06$, but a nonsignificant effect on black face trials, *p*

= .66. Specifically, people displayed greater deviations when giving and taking a flower cursor to and from a white target, compared to when moving a spider cursor. However, congruency effect failed to reach significance on trials with a black face target.

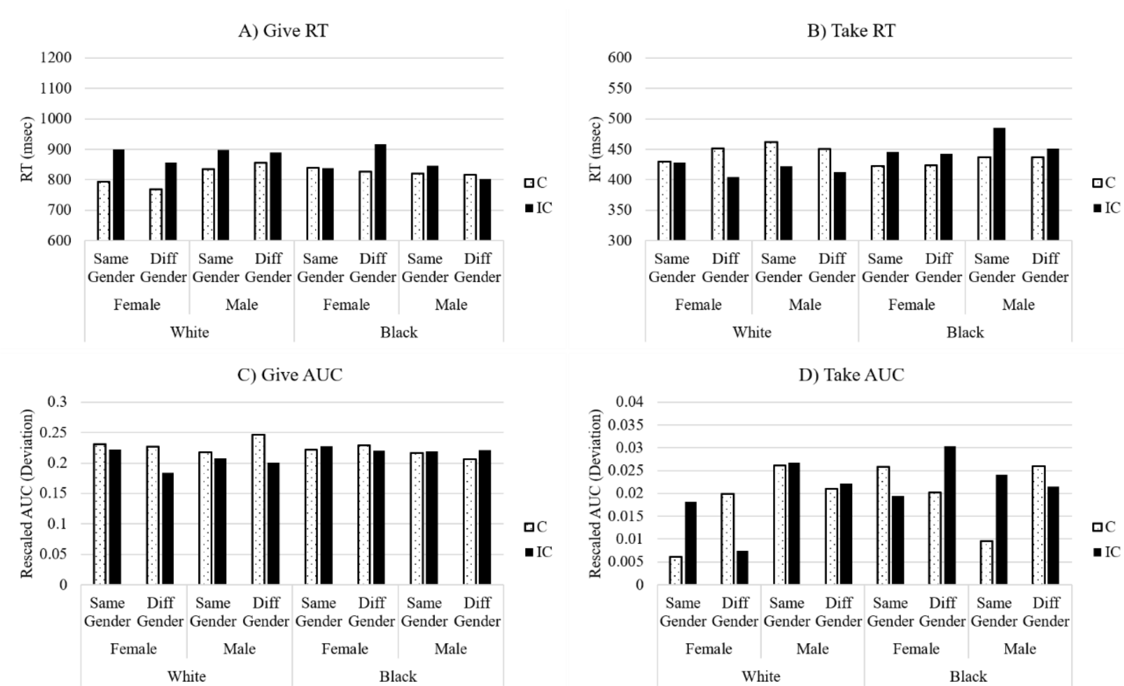

Figure 3. Descriptive data for Experiment 2. A) Give RT. B) Take RT. C) Give AUC. D) Take AUC.

Discussion

Three distinct patterns were unveiled in Experiments 1. First, a congruency effect was present and was moderated by behavioral contexts, as suggested by the three-way interaction of behavior, gender, and congruency (Figure 5B). That is, a congruency effect was obtained on giving trials replicating the implicit association effect. However, a congruency effect was not replicated on taking trials. Regardless of the cursor-target congruency, a slower mean RT was obtained when participants took a flower compared to when they took a spider (Figure 5B). Second, vigilant and efficient responses were observed especially in response to white female and black male targets (Figures 4A & 6A). Participants were faster to give a cursor to a white female, or a black male target, and showed less trajectory deviations when they did, which also suggested that the speed-accuracy trade-off had no significant influence on the observed results. Lastly, the racial facial features seemed to have affected the categorization of the target gender, even though it was not a task-relevant dimension. RT delays were observed on trials with a white male target paired with a black female distractor (Figures 5A), and a white female target paired with a white male distractor (Figure 5D). Moreover, greater trajectory deviations were found for trials with a white target paired with a black target (Figure 6B).

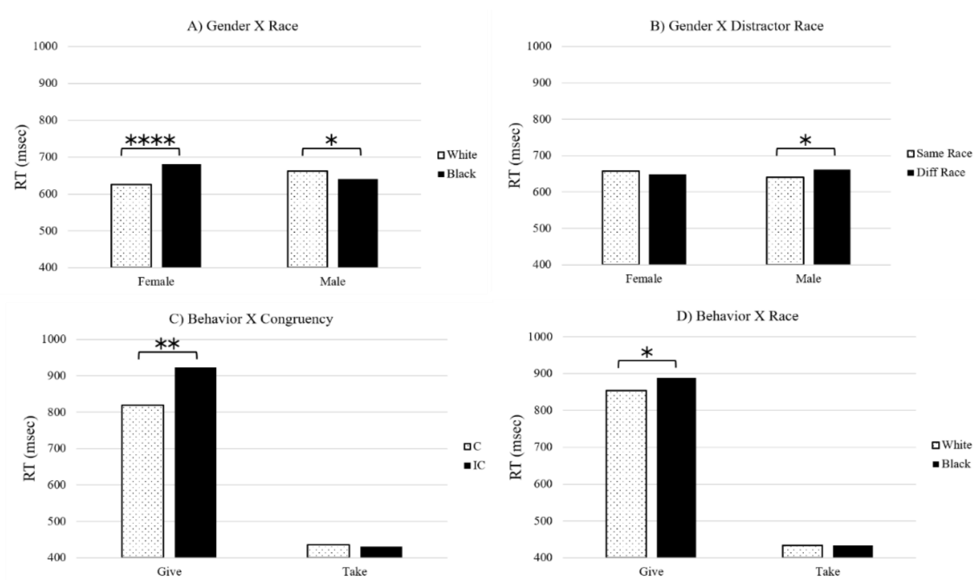

Figure 4. Interaction plots (RT) for Experiment 1. A) Gender by race interaction plot. B) Gender by distractor race plot. C)Behavior by congruency plot. D) Behavior by race plot.

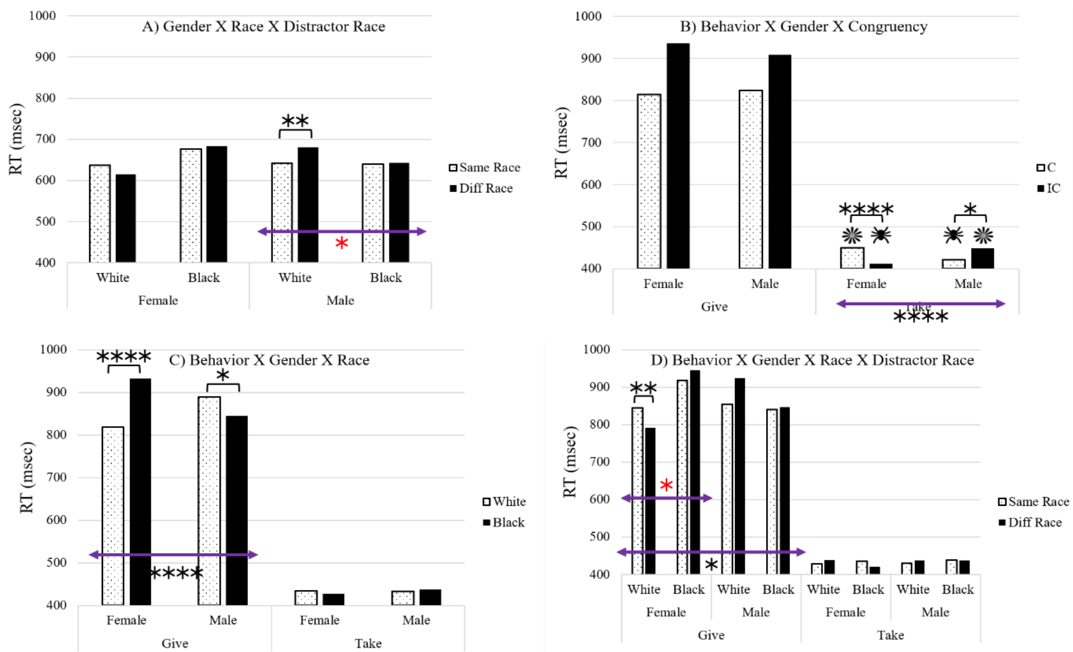

Figure 5. Interaction plots (RT) for Experiment 1. A) Gender by race by distractor race interaction plot. B) Behavior by gender by congruency plot. C) Behavior by gender by race plot. D) Four-way interaction of behavior, gender, race, and distractor race. The red asterisk indicates marginal significance at $0.05 < p < 0.10$.

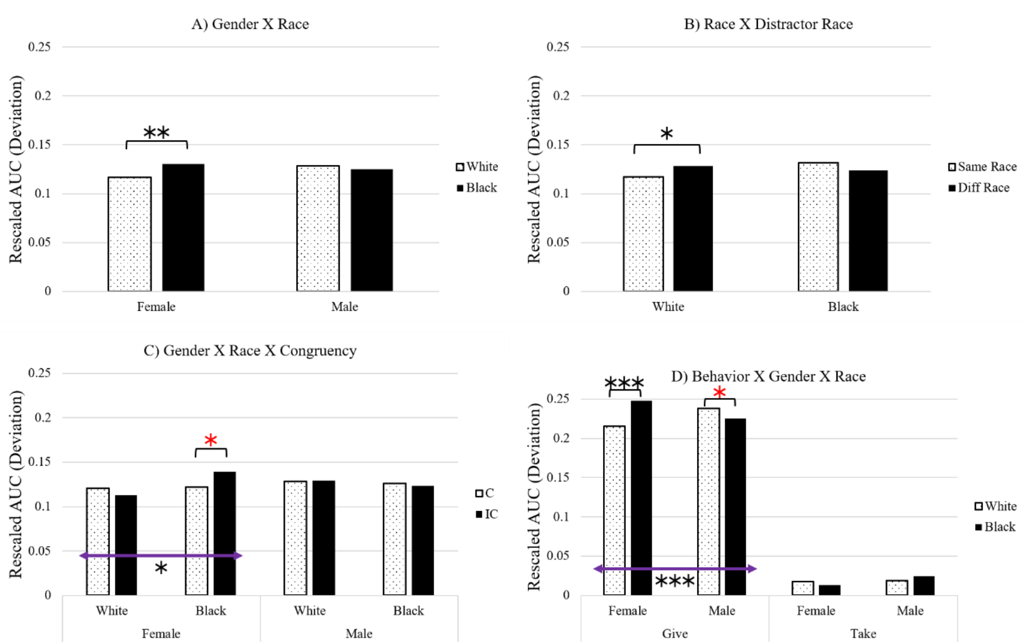

Figure 6. Interaction plots (AUC) for Experiment 1. A) Gender by race interaction plot. B) Race by distractor race plot. C) Gender by race by congruency plot. D) Behavior by gender by race plot. The red asterisk indicates marginal significance at 0.05<p<0.10.

Experiment 2 also revealed results similar to Experiment 1. A marginal congruency effect was observed on giving trials of Experiment 2 (Figure 7B). On the other hand, slower RTs were observed on trials where participants took a flower-shaped cursor than on trials they took a spider-shaped cursor like in Experiment 1 (Figure 7B). Efficient responses to white female targets were also observed, which was reflected through the smaller mean trajectory deviation on white female target trials compared to black female target trials (Figure 8A). Distractor gender seemed to have affected race categorization in Experiment 2. RTs were slowed on trials with a black male target paired with a same-gender distractor (white male) compared to when the same target was paired with a different-gender distractor (white female, Figure 7A).

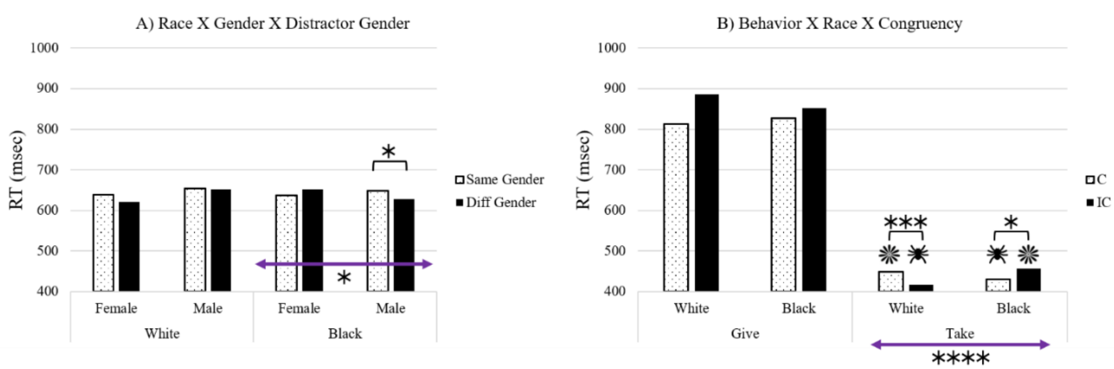

Figure 7. Interaction plots (RT) for Experiment 2. A) Race by gender by distractor gender interaction plot. B) Behavior by race by congruency plot.

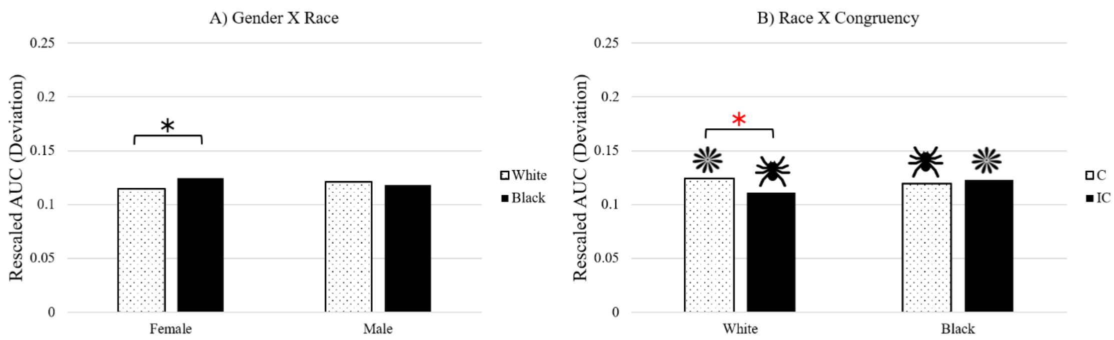

Figure 8. Interaction plots (AUC) for Experiment 2. A) Gender by race interaction plot. B) Race by congruency plot. The red asterisk indicates marginal significance at $0.05 < p < 0.10$.

General Discussion

The IAT was integrated into the mouse-tracking paradigm to capture the dynamic nature of implicit bias mechanisms in different types of behavioral contexts and in response to the faces of different social categories. Experiments 1 and 2 replicated the typical implicit association effect but only on giving trials. Vigilant (faster RT) and efficient (less trajectory deviation) responses to a black male and white female targets were also observed. In addition, the influence of task-irrelevant dimensions of distractors (e.g. race of a distractor in the gender task) was observed, suggesting that social categorization is a dynamic process that is affected by the identity of the target being evaluated, the behavioral context, and the environment that surrounds the target.

Absence of Congruency Effect on Taking Trials

Throughout Experiments 1 and 2, participants were consistently slower to take a flower-shaped cursor, and faster to take a spider-shaped cursor, regardless of the race and gender of the target faces. This observation contrasted with the observation from giving trials in which congruency effects were obtained consistently. Similar to the current experiment, Schouppe et al. (2012) had participants respond to the names of different colors, either inked in a congruent color or an incongruent color, by moving a virtual manikin toward or away from the names. Schouppe et al. found a congruency effect on trials where participants were asked to move the manikin toward (approach) the color names, but not on trials where they were asked to move the manikin away (avoid) from the names. Based on this result, Schouppe et al. concluded that the absence of the congruency effect on avoiding trials was due to the participants' tendency to avoid conflict. According to them, the avoidance behavior becomes a predominant response in the face of conflict, thereby eliminating the congruency effect because the response of avoiding an incongruent stimulus is facilitated. Therefore, it is possible that the absence of implicit association effect (congruency effect) on taking trials might be because the behavior of taking was facilitated when participants took an object counter-stereotypical to the target, as it was a conflict-eliciting behavioral context. Nevertheless, it is unlikely that the presentation of a counter-stereotypical stimulus pair, like a flower and a male face, evoked conflict because participants were consistently slower to take a flower, and faster to take a spider regardless of the target faces. Rather, it seems more likely that it was the act of holding an unpleasant object like a spider over a target face that elicited conflict. Similarly, Cacioppo, Priester, and Berntson (1993) reported that faster responses are observed when a movement involves moving away from aversive stimuli like a spider in the current experiments. However, the question remains: Why was the conflict evoked when holding a spider over a target face, but not when taking the spider back to the participants? According to Markman and Brendle (2005), participants tend to perceive a moving stimulus to be moving around a reference point, which is not necessarily the physical location of the participant. In the current Experiments, participants were given a specific instruction that framed the behavior of giving and taking with reference to the target faces, but not the participants themselves. Accordingly, it seems likely that the taking behavior was perceived analogous to the avoidance behavior, which in turn facilitated the act of taking an aversive object like a spider and slowed the act of taking a pleasant object like a flower from a target face.

Vigilant and Efficient Responses to Black Male and White Female Faces

One explanation for the vigilant and efficient responses to black male and white female targets is that humans have a learned bias that guides them to detect specific combinations of race and gender to achieve the goal of maintaining evolutionary fitness for their ingroups. For example, outgroup males are often considered a threat to resource and safety of one's ingroup (McDonald, Navarrete & Van Vugt, 2012), thereby eliciting vigilant responses (Becker et al., 2014), and are more likely to be perceived as enemies (Becker et al., 2011). Ingroup females are important for the reproductive fitness of an individual's ingroup (McDonald et al., 2012), despite the findings that they are not usually searched efficiently as outgroup males (Becker et al. 2014; Navarette et al. 2009). Previous research has also suggested that beliefs, goals, and concerns people have can tune their attention toward objects or locations that might be relevant to the goals (Folk, Remington, & Johnston, 1992). For example, people often notice more red items than usual when their goal at hand is to find a red item like an apple. This temporary increase in the sensitivity toward specific features in pursuit has helped humankind increase its evolutionary fitness, for example, by allowing a faster search for foods and faster detection of predators (Anderson, 2015). Similarly, it is possible that the human perceptual system was tuned to outgroup males and ingroup females over time because of their significances for the evolutionary fitness of ingroups.

The argument by Johnson et al. (2012) provides a hint on how the tuning might occur. As they suggested, stereotypes and facial features of African Americans converge to those of masculine characteristics, whereas those of Caucasian Americans and Asians converge to feminine characteristics. Therefore, when a face is comprised of two identity features that overlap in the same gender dimension, like a black male face, the categorization of this face is facilitated compared to the face comprised of non-overlapping dimensions like a white male or an Asian male face. If this was the case, it should also be possible to change the sensitivity toward specific identity dimensions by altering stereotypes and beliefs associated with those dimensions, which could be the way the human perceptual and behavioral systems are tuned.

Task-irrelevant Dimensions of a Distractor

In the current experiments, task-irrelevant dimensions of distractors also affected the categorization of the faces. This finding suggests that the distractor faces were also perceived, attended, and possibly went through the categorization process like the target faces. Consider two hypothetical scenarios describing the path of attention in a trial, each of which represents different assumptions on how deeply the target and distractor stimuli were processed. The first scenario would be that focal attention was allocated to the target location first, leaving the distractor unattended. In this case, distraction could arise only when the distractor was perceptually processed by parallel attention, even when participants were attending to the target location and moving a cursor to that location. The perceptual processing of the distractor, in return, should activate a competing response option that would pull the cursor toward the distractor location. However, the parallel processing capacity is known to decrease as a function of stimulus eccentricity (Carrasco et al., 1995). Therefore, it is likely that the distractor located on the opposite side of the screen had a limited impact on pulling the cursor toward the distractor location. The second scenario assumes that the distractor face was attended by the focal attention, regardless of whether it was attended before or after the target was attended. In the second scenario, the time and the detour attention took to visit the distractor location should add a distraction (detour distraction) to the distraction elicited by a perceptually processed distractor (pull distraction), inducing a greater RT and AUC as a linear function of detour frequency. Indeed, researchers using the mouse-tracking method assumed that an experimental task with more than two response alternatives was capable of eliciting response competitions (Dale et al., 2007; Freeman & Ambady, 2009; Spivey et al., 2005). According to this idea, the response competition would result in a bimodal distribution of trajectories comprised of a direct, linear trajectory, and a curved trajectory (Hehman, Stolier, & Freeman, 2015).

Interestingly, the distractor faces that delayed RTs in the current experiments were either black faces, which are perceived as being more masculine than white and Asian faces (Figures 5A, 6B) or white male faces comprised of non-overlapping features (Figure 5D, 7A). These findings suggest that there might be three different sources of distraction, each arising from different stages of visual information processing. The first source of distraction would be the automatic allocation of perceptual or attentional resources to the distractor faces (pull & detour distractions). As previous research pointed out, participants tend to identify males as angry, angry faces as males (Becker et al., 2007), which were the combinations of features that facilitated responses in the current experiments. Therefore, it is possible that the masculine faces (white male & black faces) were perceived as angry because of this tendency, which helped these faces pop-out both when they were targets or distractors. However, contradictory evidence also exists, which suggests that angry faces do not pop-out or capture attention efficiently but instead are only capable of resisting attentional disengagement longer than non-angry faces (Becker et al., 2019). Furthermore, to the author's best knowledge, there is no study which reported that a male face alone popped out of display or showed a greater capacity to capture attention compared to a female face. The second source of distraction can be inferred from the finding that angry faces can resist attentional disengagement (detour distraction; Becker et al., 2019). That is, masculine face distractors capture attention to the extent that other faces do but can resist the disengagement longer because they are closely associated with negative emotions or emotions of anger. Lastly, the third source can be pinpointed to the categorization of the distractor (detour distraction) because the distractor faces with non-overlapping features would take longer times to be categorized, given that they are attended. This possibility is also supported by the data as target faces with non-overlapping features were consistently responded to slower and yielded greater deviations than faces with overlapping features.

Conclusion, and Implication

Experiments 1 and 2 investigated how the behavioral context and exhibitor-target relationships affected implicit bias. It was hypothesized that implicit bias would be eliminated or reversed on taking trials compared to giving trials. Moreover, targets with specific combinations of social categories were expected to be responded more efficiently than other types of targets, possibly because of their evolutionary significance. Indeed, the current study demonstrated that the exhibitor-target relationships and the situational contexts should be considered to accurately evaluate the influence of implicit bias on overt behaviors. Nevertheless, the IAT and its variants lack sensitivity to capture the subtle nuances that exist in various behavioral contexts. In contrast, the mouse-tracking task is capable of monitoring more complex behaviors and providing rich dimensions of information. In return, the mouse-tracking paradigm can be leveraged to enhance the performances in tasks that involve complex behaviors and interactions between individuals of different social categories. Often such tasks can have direct impacts on individuals' well-being like surgical operations or rescue operations, the cases of which can benefit from the accurate measurement of implicit biases.

One limitation of the current study is that the behaviors of giving and taking were behaviors requiring different levels of cognitive resources, which might be one reason behind the significant difference in the overall mean RTs between the two types of behaviors. Specifically, two choice alternatives (a target and a distractor face) were available when participants gave cursors to target faces, whereas only one target (start button) was present when they took the cursors. This design was necessary to capture the behavioral context of actual social interaction because the behavior of giving can often have multiple potential targets in contrast to the behavior of taking, which usually has only one target. Therefore, it is possible that the significant implicit bias effects were observed on giving trials, but not on taking trials because of the difference in the amount of cognitive resources devoted to each behavior. Nevertheless, a couple of findings from the previous study point that it was more than the cognitive resource that contributed to the unique pattern observed on taking trials. For example, as in the study by Schouppe et al. (2012), the implicit bias effect was absent on the taking trials in the current experiments. Accordingly, it is possible that the elimination of the RT bias effect is due to the participants' tendency to avoid conflict.

CHAPTER 2

Overview: Experiments 3 & 4

Experiments 3 and 4 were conducted to explore if the findings from Experiments 1 and 2 could be replicated in situations in which the target's gender and emotional expression intersected. Becker (2017) found an asymmetric relationship between gender and emotional expression dimensions. In his study, participants were instructed to identify either the emotional expression or the gender of target faces. The results indicated that the interference by the gender dimension in the emotional expression decision task was greater compared to the interference by the emotional expression dimension in the gender decision task. In addition, participants suffered from greater interferences when reporting the gender or the emotional expression of angry female faces or happy male faces, which could be considered as faces of non-overlapping features. Based on these findings, Becker suggested that the human emotion recognition system takes advantage of the gender recognition system so that humans can communicate their emotions more efficiently. Therefore, the gender recognition system as a host system can override the emotion recognition system. In light of this finding, Experiments 3 and 4 also sought to examine whether gender and emotional expression showed an intersectionality effect, which contributed to the implicit association effect.

It was anticipated that the results in Experiments 1 and 2 would be replicated in Experiments 3 and 4, if implicit biases are moderated through an interaction between the exhibitor's behaviors and the target's identities. Specifically, the implicit association effect should be moderated by the behavioral contexts of giving and taking. Furthermore, facilitated responses to faces with overlapping features (i.e. angry male, happy female) should be observed. More importantly, an asymmetric pattern of distractor interferences was anticipated: the interference by the task-irrelevant gender dimension on the emotional expression decision would be greater than the interference by the emotional expression on the gender decision. Specifically, the effect sizes of the distractor-related ANOVA interactions should be greater in the emotion decision task than in the gender decision task. The gender, even as a task-irrelevant dimension, should be able to override the categorization of emotional expression and induce greater pull and detour distractions in the emotion decision task, according to Becker (2017).

Method

Participants

Forty-one male undergraduate students who identified themselves as not being of African American background were recruited (21 Caucasian American, 19 Middle Eastern, and 1 East Asian participants; Mean age: 21.3 years). All participants were given 1 course credit in return for their participation.

Stimuli, Procedure & Analysis

Stimuli, procedure, and analysis method were identical to Experiments 1 and 2, with the following exceptions. First, sixteen different faces were selected from the Pictures of Facial Affect Set (Ekman & Friesen, 1976) to prevent outlier faces from distorting statistical estimates and increase the ecological validity of the study. As a result, the number of trials was increased to two blocks of 128 trials (total of 256 trials, twice the trial sizes of Experiments 1 and 2) for each task. Furthermore, participants performed both the gender decision task and the emotional expression decision task, thus performed 512 trials plus 10 practice trials in total. Participants were instructed to move a flower-shaped cursor to a female target, and a spider-shaped cursor to a male target in the congruent block of the gender task. They were instructed to reverse the cursor-target mapping in the incongruent block. In the emotional expression decision task (emotion task), the flower-shaped cursor was associated with a happy face, and the spider-shaped cursor was associated with an angry face in the congruent block, and vice versa in the incongruent block.

As in Experiments 1 and 2, incorrect trials were excluded from the analysis. A total of 136 incorrect trials was excluded in the gender decision task, which comprised 1.3 percent of the entire 10,496 trials. In the emotion decision task, a total of 170 trials (1.6 %) was excluded. In addition, the 3SD criteria filtered 637 outlier trials (6.1 %) in the gender decision task, and 638 trials (6.1 %) in the emotion decision task. Using the RT and trajectory deviation data as dependent variables, five-way repeated-measures ANOVAs were performed with face gender (male, female), emotional expression (emotion; happy, angry), congruency (congruent block, incongruent block), behavior (give, take), and distractor identity (same or different emotion as the target in Experiment 3, gender in Experiment 4) as the within-subjects factors.

Experiment 3 Results

RT. Descriptive RT and trajectory deviation data are shown in Figure 9 but refer to Figures 11 and 12 to examine significant interactions obtained in the current experiment. Main effect of gender was significant, $F(1, 40) = 7.79$, $p < .01$, $\eta_p^2 = .16$, because participants responded slower to male targets (*M* = 952 ms) than to female targets (*M* = 936 ms). Main effect of behavior was also significant, $F(1, 40) = 129$, $p < .0001$, $\eta_p^2 = .76$. Participants were slower to give cursors (*M* = 1175 ms) than to take them (*M* = 713 ms).

Gender interacted with distractor emotion, $F(1, 40) = 4.9$, $p < .05$, $\eta_p^2 = .11$. Simple effect analyses indicated that participants' responses were slowed down when a male target was presented with a distractor displaying the same emotional expression than when it was displaying a different expression, $p < .05$. When a female target was presented, mean RTs were comparable regardless of the emotional expression of the distractors, $p = .31$.

A three-way interaction of gender, congruency, and behavior was also significant, $F(1, 40) = 22.3$, $p < .0001$, $\eta_p^2 = .36$. Separate two-way ANOVAs at each behavior level indicated that two-way interactions of gender and congruency was significant on both giving trials, $F(1, 40) = 12.4$, $p < .01$, $\eta_p^2 = .24$, and taking trials, $F(1, 40) = 23.6$, $p < .0001$, $\eta_p^2 = .37$. On giving trials, a congruency effect was obtained for female target trials, $p < .05$, and a nonsignificant congruency effect was obtained for male target trials, $p = .35$, suggesting that participants were slower to give an incongruent, spider-shaped cursor to a female target than to give a congruent, flower-shaped cursor to the same target. On taking trials, a reversed congruency effect was obtained for female target trials, $p < .05$, and a congruency effect was obtained for male target trials, $p < .01$.

AUC. Main effect of gender was significant, $F(1, 40) = 7.36$, $p < .01$, $\eta_p^2 = .16$. Participants showed greater deviations when responding to a male target (*M* = .11 unit) compared to when responding to a female target (*M* = .096 unit). Main effect of behavior was also significant, $F(1, 40) = 21.3$, $p < .0001$, $\eta_p^2 = .88$, becaues participants showed greater deviations when giving cursors (*M* = .23 unit), than when taking them (*M* = − .026 unit).

A three-way interaction of gender, emotion, and behavior was obtained, $F(1, 40) = 5.12$, $p < .05$, $\eta_p^2 = .11$. Separate two-way analyses revealed a two-way interaction of gender and emotion on giving trials, $F(1, 40) = 4.81$, $p < .05$, $\eta_p^2 = .11$, but not on taking trials, $F(1, 40) = .03$, $p = .86$. On giving trials, participants showed a trend of less deviation when responding to a happy female face than to an angry female face, which was statistically a nonsignificant difference, $p = .26$. Participants also showed less deviation when responding to an angry male face than to a happy male face, although this difference was only marginally significant, $p = .056$.

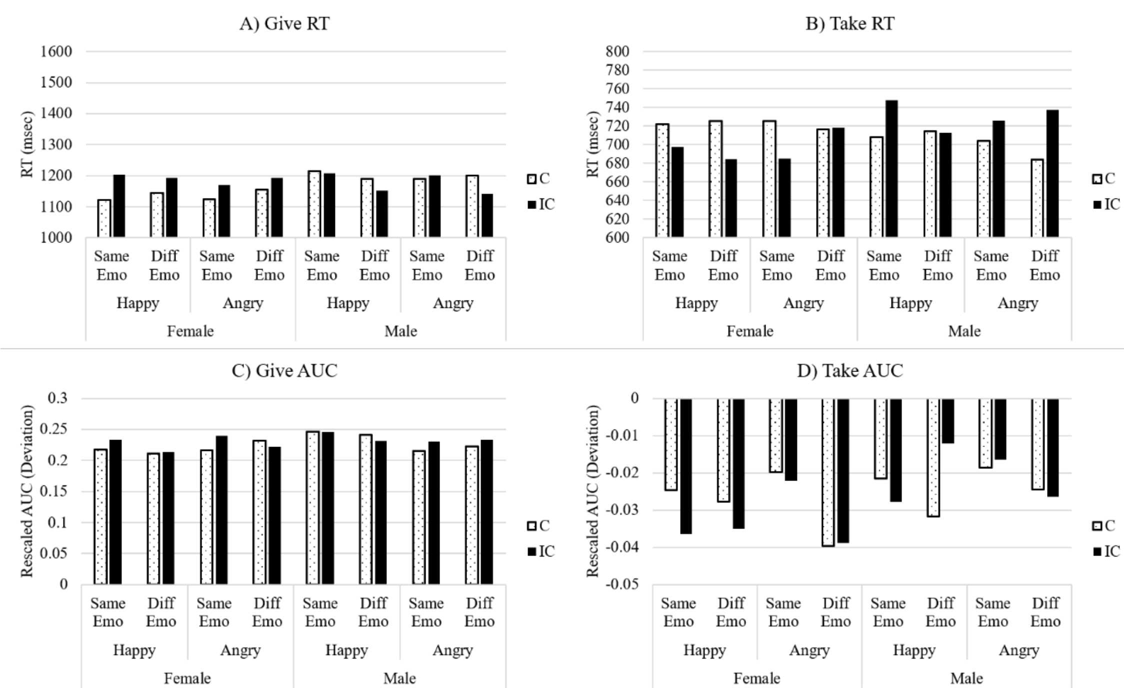

Figure 9. Descriptive data for Experiment 3. A) Give RT. B) Take RT. C) Give AUC. D) Take AUC.

Experiment 4 Results

RT. Descriptive RT and trajectory deviation data are shown in Figure 10 but refer to Figures 13, 14, and 15 to examine significant interactions obtained in Experiment 4. Main effect of congruency was significant, $F(1, 40) = 19.9$, $p < .0001$, $\eta_p^2 = .33$, because RT was longer on incongruent trials (M = 1087 ms) than on congruent trials (M = 1001 ms), confirming a 86-ms congruency effect. Main effect of emotion was significant as well, $F(1, 40) = 12.8$, $p < .001$, $\eta_p^2 = .24$. Participants were slower to respond to an angry target (M = 1061 ms) than to a happy target (M = 1027 ms). Main effect of behavior was significant, $F(1, 40) = 136$, $p < .0001$, $\eta_p^2 = .77$, because participants were slower to give (M = 1372 ms) than to take (M = 717 ms).

A two-way interaction of gender and distractor gender was significant, $F(1, 40) = 10.4$, $p < .01$, $\eta_p^2 = .21$. Simple effect analyses at each gender level indicated that participants were marginally slower to respond to a female target when it was presented with a different-gender distractor (male) than when it was presented with a same-gender distractor (female), $p = .09$. In contrast, participants were slower to respond to a male target when it was presented with a same-gender distractor (male) than when it was presented with a different-gender distractor (female), $p < .001$. Distractor gender also interacted with emotion, $F(1, 40) = 7.73$, $p < .01$, $\eta_p^2 = .16$. Simple effect analyses at each emotion level suggested that participants were slower, in general, to respond to a happy target presented with a same-gender distractor than to the same target presented with a different-gender distractor, $p < .01$. On trials with an angry target, the gender of the distractor face did not affect the RTs to the targets, $p = .15$. A two-way interaction of congruency and behavior was obtained, $F(1, 40) = 25.2$, $p < .0001$, $\eta_p^2 = .39$, which suggested that congruency effect was significant on giving trials, $p < .0001$, but not on taking trials, $p = .99$. Behavior also interacted with emotion, $F(1, 40) = 16.9$, $p < .001$, $\eta_p^2 = .29$. Participants were slower to respond to an angry face than to a happy face on giving trials, $p < .001$, but not on taking trials, $p = .94$.

A three-way interaction of gender, emotion, and distractor gender was obtained, $F(1, 40) = 9.87$, $p < .01$, $\eta_p^2 = .19$. Separate two-way ANOVAs at each gender level revealed a significant interaction of emotion and distractor gender on female target trials, $F(1, 40) = 13.9$, $p < .001$, $\eta_p^2 = .26$, and a nonsignificant interaction on male target trials, $F(1, 40) = .21$, $p = .65$. Specifically, participants were slower when responding to an angry female target presented with a different-gender distractor (male) than when responding to the same target presented with a same-gender distractor (female), $p < .001$. On the other hand, RTs to happy female faces did not vary as a function of distractor gender, $p = .18$.

Congruency interacted with emotion and behavior, $F(1, 40) = 8.29$, $p < .01$, $\eta_p^2 = .17$. Further two two-way ANOVAs at each behavior level were performed to investigate the interaction. Whereas the interaction of emotion and congruency failed to reach significance on giving trials, $F(1, 40) = 2.56$, $p = .12$, it reached significance on taking trials, $F(1, 40) = 14$, $p < .001$, $\eta_p^2 = .26$. The two-way interaction was obtained on taking trials because a reversed congruency effect was obtained when participants took a cursor from a happy face, p = .06, and a congruency effect was obtained when they were taking it from an angry face, p < .01.

A three-way interaction of gender, behavior, and distractor gender was obtained, $F(1, 40) = 7.48$, $p < .01$, $\eta_p^2 = .16$. Two-way ANOVAs at each behavior level were performed, and the two-way interaction of gender and distractor gender was significant on giving trials, $F(1, 40) = 10.3$, $p < .01$, $\eta_p^2 = .20$, but not on taking trials, $F(1, 40) = 1.19$, $p = .28$. On giving trials, participants were slower to give a cursor to a female target when it was presented with a different-gender distractor (male), than when it was presented with a same-gender distractor (female), $p < .05$. Participants were also slower to give a cursor to a male target presented with a same-gender distractor (male), than when it was presented with a different-gender distractor (female), $p < .01$.

A three-way interaction of emotion, behavior, and distractor gender was obtained as well, $F(1, 40) = 4.09$, $p < .05$, $\eta_p^2 = .09$. On giving trials, participants were slower when giving a cursor to a happy target presented with a same-gender distractor, than to the same target presented with a different-gender distractor, though this did not reach criterion, $p = .07$. Contrarily, participants were slower to give a cursor to an angry target presented with a different-gender distractor than to an angry target presented with a same-gender distractor, p .= 07, leading to a significant two-way interaction of emotion and distractor gender, $F(1, 40) = 6.13$, $p < .05$, $\eta_p^2 = .13$. However, such an interaction was not obtained on taking trials, $F(1, 40) = .92$, $p = .34$.

Lastly, a four-way interaction of gender, congruency, emotion, and behavior was significant, $F(1, 40) = 4.51$, $p < .05$, $\eta_p^2 = .10$. Two separate three-way ANOVAs at each behavior level indicated that the three-way interaction of gender, congruency, and emotion was marginally significant on giving trials, $F(1, 40) = 3.62$, $p = .065$, $\eta_p^2 = .083$, but nonsignificant on taking trials, $F(1, 40) = 1.13$, $p = .29$. Additional two two-way ANOVAs were performed at each emotion level on giving trial data only. As a result, the two-way interaction of gender and congruency was marginally significant on happy target trials, $F(1, 40) = 4.06$, $p = .051$, $\eta_p^2 = .092$. Simple effect analyses suggested that this interaction was driven by a significant congruency effect on trials with a happy female target ($M = 245$ ms), $p < .0001$, which was greater than the congruency effect found for trials with a happy male target ($M = 157$ ms), $p < .001$. In contrast, the two-way interaction was nonsigficiant on angry target trials regardless of target gender, $F(1, 40) = .22$, $p = .64$.

AUC. Main effect of emotion was significant, $F(1, 40) = 6.09$, $p < .05$, $\eta_p^2 = .13$, as participants showed greater deviations when responding to an angry face ($M = .12$ unit) than to a happy face ($M = .11$ unit). The main effect of behavior was significant, $F(1, 40) = 410$, $p < .0001$, $\eta_p^2 = .91$. Again, participants showed greater deviations when giving a cursor ($M = .25$ unit) than when taking a cursor ($M = -.026$ unit).

A three-way interaction of gender, emotion, and behavior was obtained, $F(1, 40) = 4.22$, $p < .05$, $\eta_p^2 = .095$. Separate two-way analyses at each behavior level revealed a nonsignificant emotion by gender interaction on giving trials, $F(1, 40) = 1.57$, $p = .22$, and a significant interaction on taking trials, $F(1, 40) = 4.76$, $p < .05$, $\eta_p^2 = .11$. On taking trials, participants showed greater deviations when responding to a happy female face than to an angry female face, p = .22, and when responding to an angry male face than to a happy male face, p = .15, which turned out to be a nonsignificant effect.

Lastly, a three-way interaction of gender, distractor gender, and behavior was significant, $F(1, 40) = 5.89$, $p < .05$, $\eta_p^2 = .13$. Separate ANOVAs were performed on data split by behavior, which revealed a marginal gender by distractor gender interaction on giving trials, $F(1, 40) = 3.41$, $p = .07$, $\eta_p^2 = .079$, but a nonsignificant interaction on taking trials, $F(1, 40) = 2.25$, $p = .14$. Participants showed marginally greater deviations when giving a cursor to a female target presented with a different-gender distractor (male) than to the same target presented with a same-gender distractor (female), p = .056. However, the difference in RTs to male targets was nonsignificant regardless of the distractor gender, p = .75.

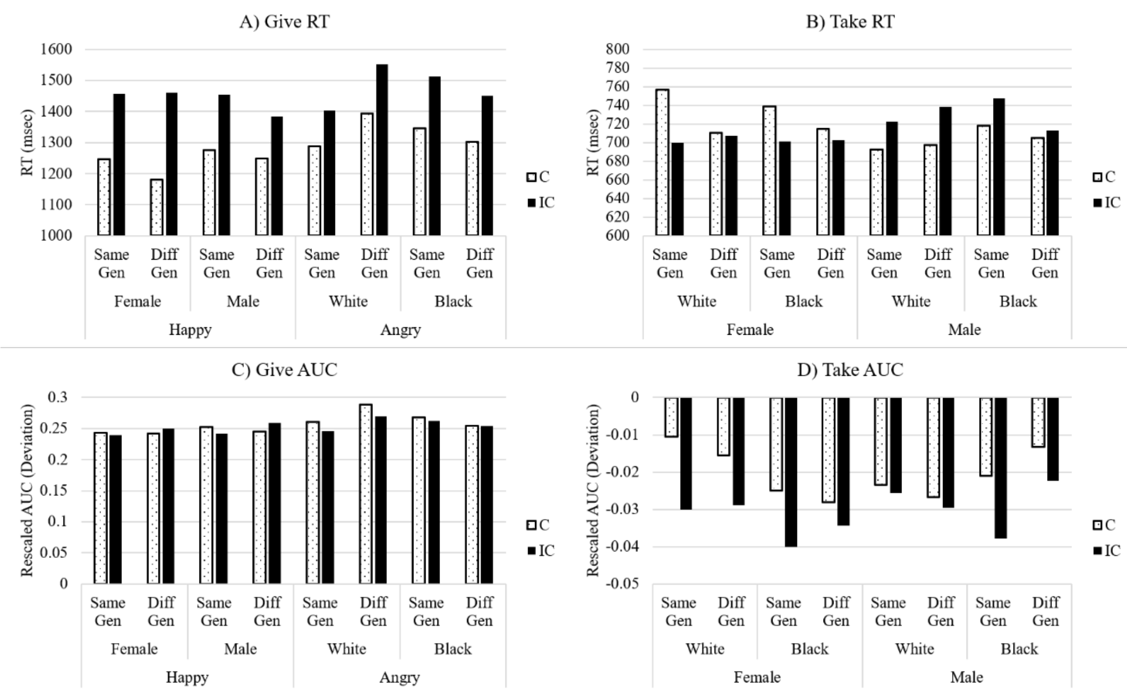

Figure 10. Descriptive data for Experiment 4. A) Give RT. B) Take RT. C) Give AUC. D) Take AUC.

Discussion

Results similar to Experiments 1 and 2 were replicated in Experiment 3. First, a congruency effect was present on giving trials, but only on trials with female targets (Figure 11B). On taking trials, participants were slower to retrieve a flower-shaped cursor, and faster to retrieve a spider-shaped cursor (Figure 11B). Furthermore, a smaller trajectory deviation was observed when giving a cursor to an angry male face, which was a face with overlapping features (Figure 12A). However, the difference in mean deviations between happy female and angry female target trials was nonsignificant although a numerically smaller deviation was found on trials with a happy female target. The task-irrelevant, emotional expression of distractor faces also affected RTs in Experiment 3 (Figure 11A). Participants were slower to give a cursor to a male target presented with a female distractor expressing the same emotion as the target, compared to when the same target was presented with a female distractor expressing a different emotion.

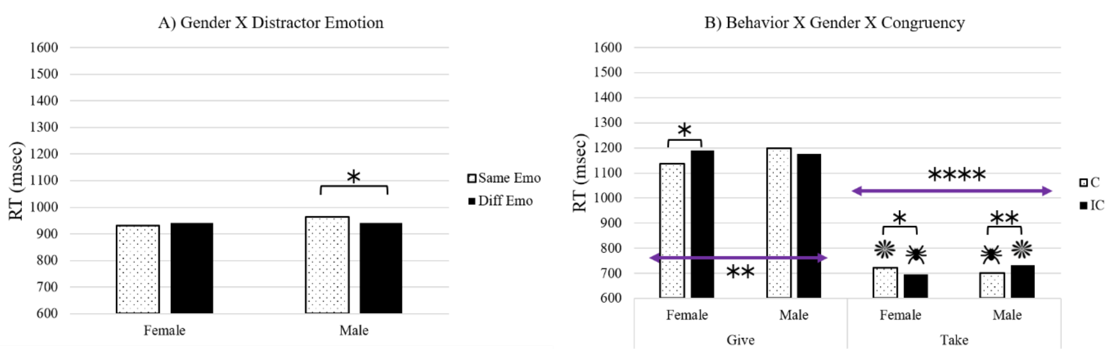

Figure 11. Interaction plots (RT) from Experiment 3. A) Gender by distractor emotion interaction plot. B) Behavior by gender by congruency plot.

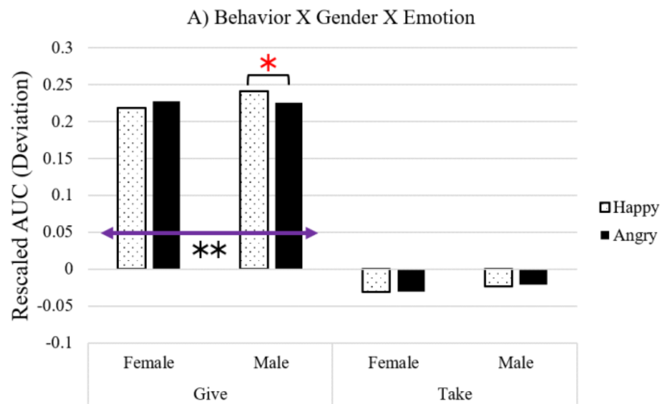

Figure 12. Interaction plots (AUC) from Experiment 3. A) Behavior by gender by emotion interaction plot. The red asterisk indicates marginal significance at 0.05<p<0.10.

Congruency effect was obtained in Experiment 4, again, only on giving trials (Figures 13C, 14E). On taking trials, participants were slower to take a flower-shaped cursor and faster to take a spider-shaped cursor, regardless of the cursor-face congruency (Figure 14B). In contrast to Experiments 1 and 2, vigilant responses to the faces with overlapping features were not observed. Moreover, trajectory deviation was greater when taking a cursor from a happy female face or an angry male face, which was a result contradicting Experiments 1, 2, and 3 (Figure 15A). Distractor gender, which was the task-irrelevant dimension in Experiment 4, also affected participants' responses. Two distinct patterns were identified. First, greater RTs and AUCs were observed when participants responded to a target presented with a male distractor, regardless of the target gender and emotion (Figures 13A, 14C, and 15B). Second, distraction was also found for a happy target paired with an angry distractor of the same gender, and for an angry target paired with a happy distractor of different gender (Figures 13B, 14A, 14D).

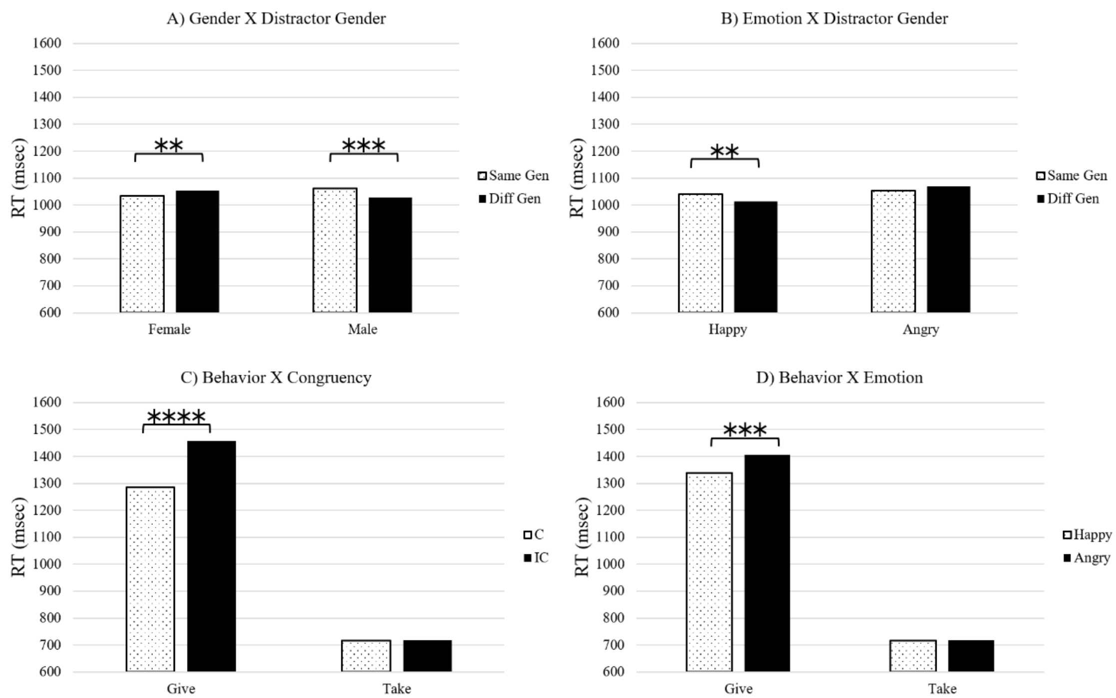

Figure 13. Interaction plots (RT) from Experiment 4. A) Gender by distractor gender interaction plot. B) Emotion by distractor gender plot. C) Behavior by congruency plot. D) Behavior by emotion plot.

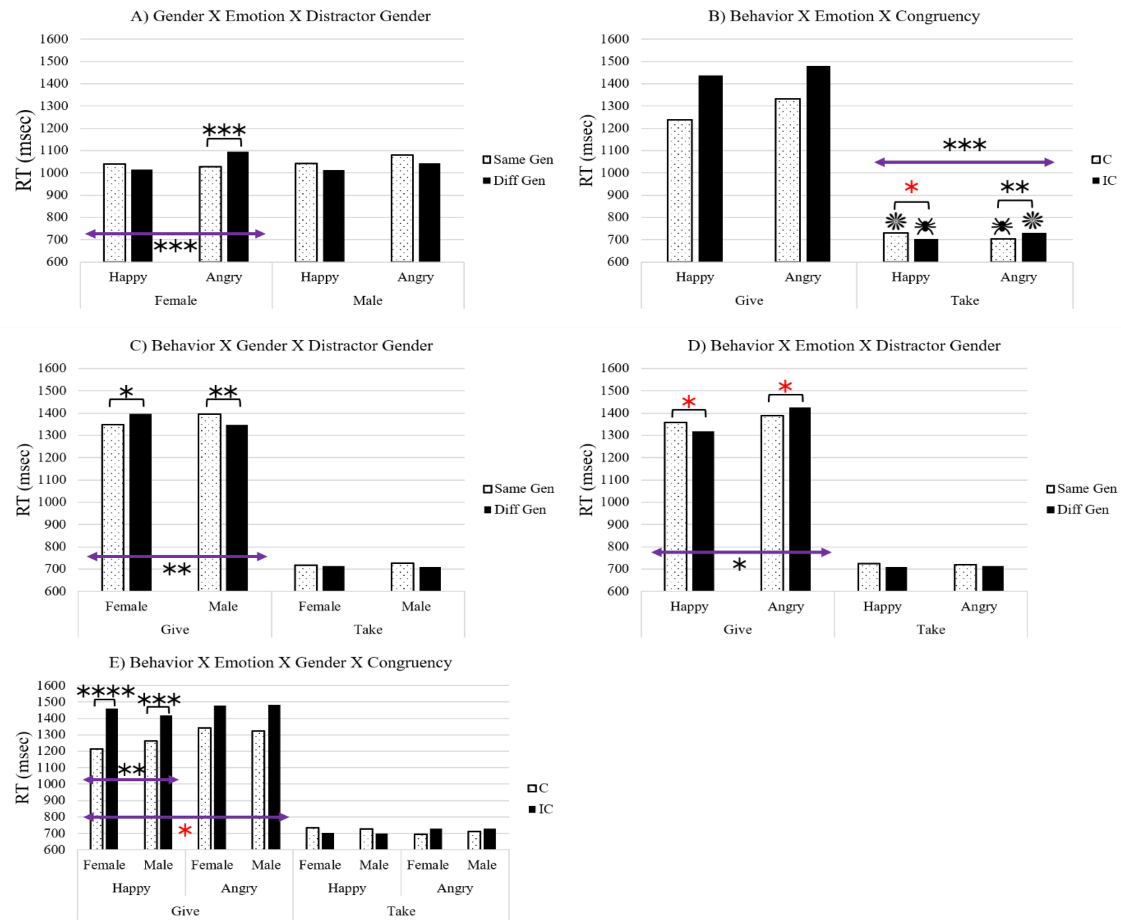

Figure 14. Interaction plots (RT) from Experiment 4. A) Gender by emotion by distractor gender interaction plot. B) Behavior by emotion by congruency plot. C) Behavior by gender by distractor gender plot. D) Behavior by emotion by distractor gender plot. E) Behavior by emotion by gender by congruency plot.

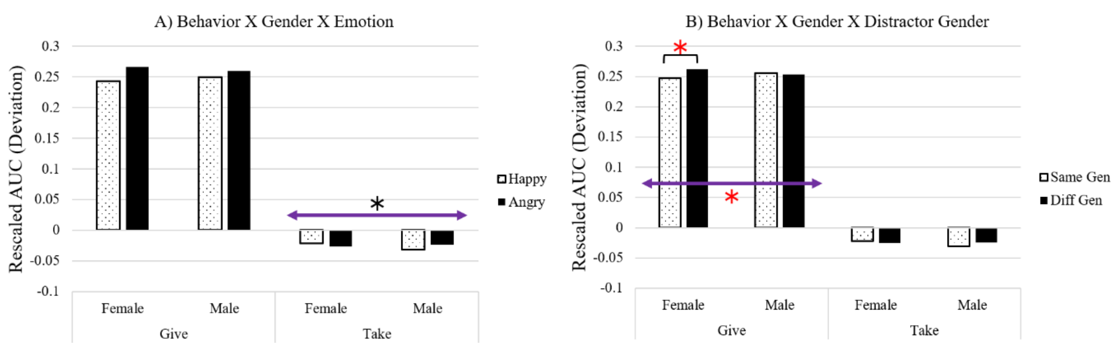

Figure 15. Interaction plots (AUC) from Experiment 4. A) Behavior by gender by emotion interaction plot. B) Behavior by gender by distractor gender plot.

General Discussion

Some of the findings obtained in Experiments 1 and 2 were replicated in Experiments 3 and 4. Regardless of the task at hand, a congruency effect was obtained on giving trials replicating the implicit association effect. However, this effect was absent on taking trials, and instead, a consistent pattern of slowed RTs when taking a flower-shaped cursor, and facilitated RTs when taking a spider-shaped cursor was observed. Smaller trajectory deviation was obtained when participants were reaching for an angry male target compared to when they were reaching for a happy male target in Experiment 3. Participants also displayed smaller trajectory deviation when reaching for a happy female target than when reaching for an angry female target, although this difference was nonsignificant. Moreover, a male face tended to increase distraction when presented as a distractor (Experiment 4; Figures 13A, 14C, and 15B). Interestingly, face gender was the task-irrelevant feature in Experiment 4, because participants were given an emotional expression decision task. This finding contrasted to the result from Experiment 3 in that the distractor emotion displayed relatively less frequent and smaller effects on RT (η_p^2 = .11), compared to the distractor gender in Experiment 4 (η_p^2 = .19~.21). This might be a result supporting Becker's (2017) argument that the gender and the emotion recognition systems are hierarchically structured such that the gender system can override the emotion system occasionally.

Altogether, these results support the previous arguments that positive emotion is stereotypically associated, thus overlap, with feminine features, and that negative emotion is stereotypically associated with masculine features. Experiments 3 and 4 also replicated efficient responses to faces of overlapping features, which corroborates the idea that the intersection of different identity dimensions has an interactive effect on how a face is categorized, hence affect the implicit biases triggered by a face. Lastly, the findings from the current experiments demonstrated that implicit bias could be modulated by the behavioral contexts of giving and taking.

Efficient Responses to Faces of Overlapping Features

In Experiment 4, a smaller deviation was obtained when participants were taking cursors from non-overlapping faces like an angry female or a happy male face (Figure 15A). Note that the AUC values for taking trials in this experiment were negative values, suggesting that the majority of the taking trajectories were convex-shaped, and deviated away from the distractor. Therefore, while a greater AUC value would still suggest that the trajectories curved more toward the distractor location, it does not necessarily indicate a greater absolute deviation from an ideal trajectory, unlike the positive AUC values. When the absolute deviation was considered, the results from Experiment 4 followed that of the previous experiments: greater absolute deviations on trials with faces of non-overlapping features, and smaller deviations on trials with faces of overlapping features. The reason behind the reversed polarity may be because different response tendencies (e.g. which direction to deviate to) were activated for different types of behaviors. In Experiment 4, negative AUCs values were observed only on taking trials, which supported this possibility and corresponded to converging evidence that different response alternatives can be triggered in different contexts (Marsh et al., 2005; Schouppe et al., 2012). Therefore, the reversed polarity observed in Experiment 4 is likely to be an artifact of the behavior type.

Task-irrelevant Dimensions of a Distractor

Moreover, there were findings regarding the task-irrelevant distractor features which could not be explained either by the pop-out and stronger capture possibilities, or the delay in the categorization of distractors. Distractions were observed for a male target paired with a female distractor displaying the same emotion in Experiment 3 (Figure 11A), a happy face paired with an angry distractor of the same gender (Figures 13B, 14D), and an angry target paired with a happy distractor of a different gender in Experiment 4 (Figure 14D). These effects were mostly two-way interaction effects on which either the distractor emotion or the target gender had a little effect. Accordingly, only a limited interpretation is possible for these effects, which is that these distractors were salient, thus were capable of distracting participants' attention away from the target to some extent.

Conclusion and Implication

Experiments 3 and 4 were conducted to see if the findings from Experiments 1 and 2 generalized to cases in which participants relied on gender and emotional expression categorizations. While most of the findings were replicated, a distinct pattern arose in Experiments 3 and 4, which was the asymmetric interference effects between the gender and the emotion decision tasks. The asymmetric interference effects suggest that there is another factor to be considered when assessing the influence of implicit bias: The target-distractor relationship. Indeed, many types of social interactions require an individual to interact with multiple people at the same time, which seems to affect overt behaviors during social interactions. This finding suggests that teammates, coworkers, audiences, or even strangers that appear in social interactions have the potential to affect how implicit bias shapes overt behaviors. Although social interactions can occur randomly and between random individuals, some types of interactions can be expected and even be controlled to some extent. For example, a team of paramedics could be formed based on the social categories of a patient in need of help. The understandings of how biased behaviors are affected by the social memberships of individuals in a given social interaction may help achieve the goal of the interaction.

CHAPTER 3

Overview: Experiments 5 & 6

Experiments 5 and 6 were conducted as preparatory work for demonstrating the relationship between working memory and mouse trajectory in order to pinpoint the specific cognitive mechanism (cognitive control mechanisms) responsible for handling implicit biases.

Working memory is a cognitive function that holds and processes information temporarily before the information is conveyed to long-term memory (Miyake & Shah, 1999). Working memory is thought of as an executive function (McCabe et al., 2010), which is an umbrella term for a collection of cognitive functions that aid goal-driven actions, including the control of goal-driven motor activity (Garavan et al., 2002). Working memory capacity is an essential resource for human cognitive functions (Glisky, 2007). Accordingly, the constraint placed by working memory load has the potential to affect various cognitive functions, even those involved in the suppression of implicit biases. For example, cognitive psychologists have long posited cognitive control mechanisms that engage in resolving response conflicts elicited by multi-dimensional stimuli such as the Stroop stimuli (e.g. the word “blue” printed in red ink). According to this idea, the anterior cingulate cortex (ACC) serves as a conflict detector that signals the dorsolateral prefrontal cortex (DLPFC), a brain area associated with executive functioning, to engage and resolve conflicts elicited by the multi-dimensional stimuli (Botvinick et al., 2004). Moreover, the performance of this cognitive control function decreases when a higher working memory load is imposed (Lavie, 2010; Lavie et al., 2004). In the study by Lavie et al. (2004), participants were asked to report the name of a target letter while ignoring a distractor letter located beside the target letter. At the same time, they were given secondary tasks of memorizing one or six numbers. The result indicated that the interference by the distractor letters was greater when participants had to remember six numbers compared to when they had to remember a single number as the secondary task.

Paralleling the above findings, implicit bias research has suggested that implicit bias might be controlled by executive function (Amodio et al., 2004). An implicit association can also elicit a response competition when the subcomponents of an association (e.g. “happy” and “male” in “happy male”) are associated with different response alternatives (Freeman, Dale, & Farmer, 2011; Greenwald et al., 1998). More than two conflicting response alternatives can elicit cognitive conflict, which will eventually be resolved but after delaying response time (Fazio et al., 1995; Greenwald et al., 1998) or inducing inefficient behaviors such as a trajectory deviation (Freeman et al., 2008). Previous research made efforts to quantify the capacity of the executive function in controlling implicit biases (Payne, 2005; Richeson & Shelton, 2003). In most of these efforts, executive control was operationally defined as the extent to which a subject’s response corresponded to the goal of a given task. Accordingly, the percentage of correct responses on congruent trials minus the percentage of incorrect responses on incongruent trials [*P*(correct response on congruent trials) - *P*(incorrect response on incongruent trials)] was estimated as the index of executive functioning capacity, because this index was considered to reflect only the trials in which the executive function was successful in controlling implicit bias. Using this index, the researchers reported that the capacity of executive function was negatively correlated with biased behaviors (Payne, 2005; Richeson & Shelton, 2003).

Therefore, it would be reasonable to assume that working memory capacity is correlated with an individual's capacity to suppress implicit biases, given that implicit associations are also handled by cognitive control mechanisms. Nevertheless, despite the resemblance between the mechanisms that handle Stroop stimuli and implicit associations, only a few efforts have been made to verify whether the Stroop stimuli and the implicit associations are handled by the same mechanisms (e.g. Amodio, 2010, 2014). For this reason, Experiments 7 and 8, presented in the next chapter, were designed to investigate whether implicit associations were handled by the same mechanisms that handled Stroop stimuli. However, it was first required that the effect of working memory load on mouse trajectory be accurately assessed because noises created by the difficulty of motor tasks (e.g. moving a mouse cursor to an icon that is small and hard to locate) were anticipated.

To this end, Experiments 5 and 6 were conducted, which used a working memory load manipulation to examine how mouse trajectories changed as a function of working memory load. An additional manipulation was introduced in the study, which was the size of the targets. This manipulation was added in order to take account of the theories that concern the human-computer interaction difficulties (motor task difficulty) caused by the physical properties of a computer display layout such as the size of target icons (e.g. Fitts's law; Fitts, 1954). The change in trajectories induced by the motor task difficulty was identified and distinguished from that induced by the working memory load by analyzing how the target size affected hand motion trajectories.

Grimes and Valacich (2015) conducted a study similar to Experiments 5 and 6, which provided an insight into how working memory load would affect mouse trajectories. In their study, participants were required to perform n-back tasks and report the answers to the tasks by clicking one of the two response options appearing at the upper corners of the screen. Participants performed three different n-back tasks that were designed to impose different levels of working memory load. The results indicated that participants were slower and yielded longer cursor trajectories when making responses during the task that imposed the highest level of working memory load. Although Grimes and Valacich did not examine other distraction measures like the trajectory deviation, their study implied that a higher working memory load could introduce a greater distraction or deviation to a mouse trajectory.

Accordingly, the specific goals of the current experiments were to 1) examine if the study by Grimes and Valacich would be replicated, 2) to find out if working memory load effect can be distinguished from motor task difficulty effect, and 3) identify features of hand motion gestures which could be used to infer about the level of working memory load. Mouse trajectories (Experiment 5) and touchscreen gestures (Experiment 6) under different levels of working memory load were recorded to compare the working memory load effects on the use of two different input devices. Furthermore, 39 features of the mouse trajectory and touchscreen gesture were extracted and fed to a machine learning algorithm in order to identify features that best indicated the level of working memory load.

Method

Participants

Forty undergraduate students participated in Experiment 5, and forty-two students participated in Experiment 6 (Experiment 5: 20 male and 20 female participants with a mean age of 22.1 years; Experiment 6: 21 male and 21 female participants with a mean age of 23.3 years). All participants were Korean students at Korea University whose majors were not psychology. One male participant was excluded from analysis in Experiment 6 because this participant reported the same Pass's rating (5: Neither low or high mental effort) for all working memory load conditions. All participants had normal or corrected-to-normal vision and were right-handed. Cash equivalent of about 5 USD was offered for their participation, which lasted up to 25 minutes.

Apparatus and Material

The experimental setting is illustrated in Figure 16. The experiments took place in front of a 22-inch LCD monitor with a screen resolution of 1920 by 1080 pixels (Experiment 5) or a 9.70-inch iPad 2 display with a resolution of 768 by 1024 pixels (Experiment 6). All participants sat approximately 60 cm from the computer monitor or 40 cm from the iPad and made responses by moving a computer mouse cursor or dragging a virtual circle on a touchscreen. The mouse-tracking program was developed using JavaScript and generated x and y coordinate data of the mouse cursor movement at the 70hz sampling rate.

Procedure

Each participant performed six blocks of 27 trials. On each trial, two circles appeared at the same time on the screen, which were the start button and the target circle. Participants performed two different primary tasks: A vertical movement task and a horizontal movement task. In the vertical movement task, the start button appeared at the bottom-center of the screen, and the target circle appeared randomly at the upper-left side of the screen or the upper-right side of the screen. In the horizontal movement task, the start button appeared at the left-center of the screen, and the target circle appeared randomly at the upper-right side of the screen or the lower-right side of the screen. Participants in Experiment 6 were provided with the same display in both the vertical and horizontal tasks, but the iPad device was rotated 90 degrees to the right in the horizontal task so that participants could drag cursors horizontally. Participants were instructed to perform the primary tasks as fast and accurate as possible. The primary tasks required participants to move a circle-shaped cursor from the start button located at the bottom or left-center of the screen to the target circle presented at the upper or right half of the screen. In Experiment 6, participants dragged a circle-shaped cursor for the primary task.

Following Cowan's (1988) model, working memory load was operationally defined as the mental resources occupied by rehearsing and subtracting numbers. Cowan (1988) proposed a working memory model that does not explicitly specify a modular structure, in contrast to Baddeley's model (1986, 1992) which conceptualized working memory as having a modular structure that is comprised of modules, two of which exclusively maintain and process auditory and visuospatial information. Contrarily, according to Cowan, working memory can be defined as an activated part of long-term memory, in which the representations of different sensory information coexist, with a smaller subset of the activated part allocated with attentional resources. Therefore, this model allows more room for crossmodal interference (Cowan, 2014), thus could better explain the cases in which working memory load imposed by an auditory task (i.e. rehearsing numbers) interfered with motor task performance.

During each block, participants performed one of three different secondary tasks: The control task in which participants had to say the number seven aloud on each trial, the low-load task in which participants had to count backward aloud in multiples of one, the high-load task in which participants had to count backward aloud in multiples of seven from a given number (e.g. 771). In addition, they were told not to finish the primary task until they completed the secondary task, which involved subtracting 1 or 7 from the remaining number (from the previous trial). The two different primary tasks and three different secondary tasks resulted in a total of 6 blocks. The order of the blocks was counterbalanced across participants with an exception that participants always performed the control task block first. At the end of each block in Experiment 6, participants were presented with the Paas subjective rating scale (Paas, 1992; Figure 17), which required them to report the amount of mental effort they devoted to the secondary tasks. This scale was inserted to validate whether the secondary tasks were successful in imposing different levels of working memory load. However, the subjective rating scale was not presented to the participants in Experiment 5.

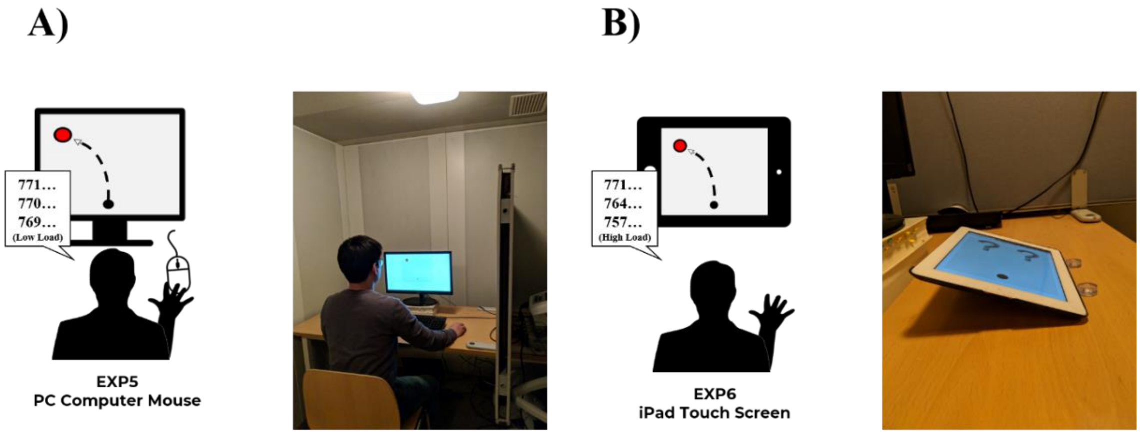

Figure 16. Experimental setting in Experiments 5 and 6. A) Experimental setting in Experiment 5 and the illustration of the low-load secondary task. B) Experimental setting in Experiment 6 and the illustration of the high-load task.

In performing the arithmetic task I invested

1: Very, very low mental effort

2: Very low mental effort

3: Low mental effort

4: Rather low mental effort

5: Neither low nor high mental effort

6: Rather high mental effort

7: High mental effort

8: Very high mental effort

9: Very, very high mental effort

Figure 17. The subjective rating scale created based on Paas (1992).

Analysis

Three independent, within-subject variables were manipulated in this study. First, there were two different primary tasks, which were the vertical movement task and the horizontal movement task (movement orientation: vertical task, horizontal task).

Participants also performed three different secondary tasks (WM load: control task, low-load task, high-load task), which participants had to perform concurrently with the primary tasks. Lastly, the circles used as the targets varied in size to manipulate the difficulty of the primary motor task (target size: small, 25px in diameter; medium, 75px; large, 125px).

Three different types of dependent variables were analyzed in Experiments 5 and 6: Response time (RT), trajectory deviation (AUC), and the time step where the velocity reached maximum (Velocity peak onset). Each trajectory data was processed using the method used in Experiment 1 with the following exceptions. All trials were first preprocessed to exclude outliers that were greater than 2.5 deviations above the mean, or 2.5 standard deviations below the mean. Again, the same criteria were applied to the number of hand motion trajectory flips to exclude aberrant trajectories. Furthermore, the trajectories were not rescaled to fit the 1 by 1.5 coordinate plane. Three-way repeated-measures ANOVAs were performed with the three within-subject variables (movement orientation, WM load, target size) on the three dependent variables (RT, AUC, Velocity peak onset). In addition, velocity and trajectory angle data were unfolded along the 101 time steps within a trial window to visualize how each trajectory developed over time and differed across conditions. This approach was taken additionally because RT and AUC do not provide the temporal information of the trajectory development, thus do not provide information on when differences between experimental conditions occurred if they existed. Velocity at each time step was calculated by dividing the distance traveled between one time step to the next step by the duration of the travel (in ms). The angle at each time-step was estimated by calculating the inverse tangent of the x and y shifts (distances traveled along the x and y axes) from one time step to the next step. Moreover, angle data were converted to complex numbers before averaging by within-subject variables, in order to get the directional average rather than the arithmetic average. The conversion to complex numbers was used because the arithmetic average of angles does not reflect the averaged direction of the trajectory movements.

As the last step, 37 hand motion features other than the RT and AUC were extracted and fed to the Support Vector Machine (SVM) algorithm. The SVM models were trained to predict the secondary task (working memory load) participants performed, and the subjective rating participants reported for the different secondary tasks (Experiment 6). The total 39 features were, 1) response time (RT), 2) time to initiate the first movement (initiation time), 3) absolute area under curve, 4) area under curve (AUC), 5) absolute maximum deviation, 6) maximum deviation, 7) length of the trajectory, 8) mean velocity, 9) maximum velocity, 10) minimum velocity, 11) mean acceleration, 12) maximum acceleration, 13) minimum acceleration, 14) velocity peak onset, 15) onset of the lowest velocity, 16) acceleration peak onset, 17) onset of the lowest acceleration, 18) number of movement flips along the x axis (x flip), 19) movement flips along the y axis (y flip), 20) sample entropy calculated based on the shift in x coordinates (x entropy), 21) sample entropy calculated based on the shift in y coordinates (y entropy), 22) length traveled along the x axis beyond the target (x overshoot), 23) length traveled along the y axis beyond the target (y overshoot), 24) frequency of flips calculated based on the Euclidian distance traveled (2D flip), 25)

Euclidean-distance-based sample entropy (2D entropy), 26) Euclidean-distance-based overshoot (2D overshoot), 27) movement time (RT−Initiation time), 28) mean velocity at quartile 1 (mean velocity between the times steps 1~25, Q1 Velocity), 29) Q2 velocity, 30) Q3 velocity, 31) Q4 velocity, 32) Q1 acceleration, 33) Q2 acceleration, 34) Q3 acceleration, 35) Q4 acceleration, 36) Q1 mean trajectory angle in radian, 37) Q2 angle, 38) Q3 angle, 39) Q4 Angle. See Figure 18 for further descriptions of these features and refer to Appendix G for the full list of the features.

In order to improve the computational efficiency of SVM, avoid overfitting, and select the features that were the most indicative of the level of working memory load, the sequential forward selection method was used with 10-fold cross-validation. As a result, 13 features from Experiment 5, and 10 features from Experiment 6 were identified as effective in predicting the level of working memory load. In addition, 12 features from Experiment 6 were identified as effective in predicting the subjective rating of working memory load reported by the participants. With these selected features, SVM classifiers were trained, with cost and gamma parameters set to 1 and 1/the number of features. As the last step, the performances of the SVM classifiers were evaluated by calculating Spearman's rank correlations between the predicted and the observed values. Furthermore, these correlation coefficients were compared with 1000 correlation coefficients derived from random permutation tests to assess the possibility that the performances of the classifiers were due to random error.

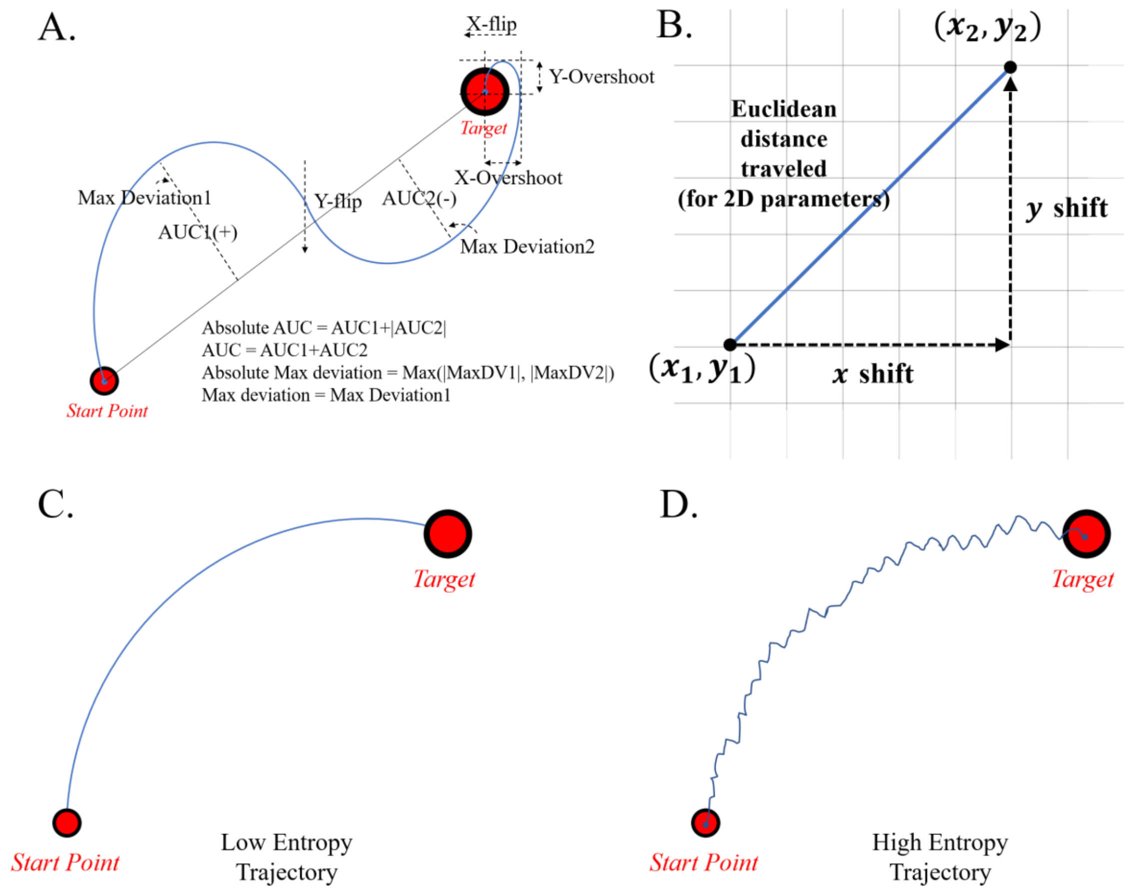

Figure 18. Examples of the hand motion features fed to SVM classifiers. A) Illustrations of AUC, Absolute AUC, flip and overshoot; B) Values used to calculate 1D/2D flips, overshoots, and entropies. The X and Y shifts were used for 1D parameters, whereas the Euclidean distance traveled was used for 2D parameters; C) An example of a low entropy trajectory; D) An example of a high entropy trajectory.

Experiment 5 Result

RT. Descriptive RT and trajectory deviation data are shown in Figures 19 and 20. Three-way repeated-measures ANOVAs on RT data revealed a main effect of working memory load (WM load), $F(2, 78) = 72.4$, $p < .0001$, $\eta_p^2 = .45$. Tukey's Honestly Significant Difference Procedure (Tukey's HSD) was used to perform posthoc tests, which revealed that participants were slower to move a mouse cursor to a target when they were performing the high-load task as the secondary task (*M* = 1818 ms), compared to the low-load task (*M* = 825 ms), $p > .0001$, and the control task, (*M* = 737 ms *),* $p > .0001$. Participants were also slower when they were performing the low-load task compared to when they were performing the control task, $p > .01$. The main effect of target size was also significant, $F(2, 78) = 123$, $p < .0001$, $\eta_p^2 = .76$, because participants were slower when reaching for a small target (*M* = 1259 ms), than when they were reaching for a medium target (*M* = 1073 ms) , $p > .0001$, or a large target (*M* = 1047 ms), $p > .0001$. However, the difference between the RTs to a medium target and a large target was nonsignificant, $p = .19$. In contrast, the main effect of movement orientation was nonsignificant, $F(1, 39) = .11$, $p = .73$.

A significant two-way interaction was obtained between WM load and target size, $F(4, 156) = 2.78$, $p < .05$, $\eta_p^2 = .067$. RT was slower to a small target followed by a medium target and a large target when participants were performing the control task, $F(2, 78) = 369$, $p < .0001$, $\eta_p^2 = .90$, and the low-load task, $F(2, 78) = 143$, $p < .0001$, $\eta_p^2 = .78$, as the secondary tasks. However, on high-load trials, the difference in the RTs to a medium target and a large target was nonsignificant, $p = .68$, while the RT to a small target was significantly slower than both a medium target, $p >.001$, and a large target, $p >.001$, when participants were performing the high-load task, $F(2, 78) = 12.7$, $p < .0001$, $\eta_p^2 = .25$.

The three-way interaction of WM load, target size, and movement orientation was significant, $F(4, 156) = 2.65$, $p < .05$, $\eta_p^2 = .064$. Further two-way ANOVAs at each movement orientation level revealed a significant interaction of WM load and target size on vertical movement trials, *F*(4, 156) = 2.78, $p < .05$, $\eta_p^2 = .067$, as well as on horizontal movement trials, *F*(4, 156) = 2.65, $p < .05$, $\eta_p^2 = .064$. On high-load trials of the vertical movement task, the difference in RTs to large and medium targets was nonsignificant, $p = .43$, while significant differences were found between small and medium target trials, $p = .001$, and small and large target trials, $p = .0001$. The differences in RTs between the trials of different-sized targets were significant on trials where participants performed the control or low-load tasks (*ps* < .05). Similarly, on horizontal trials, no significant RT differences were obtained across target sizes when participants performed the high-load task (*ps* > .37). However, the RT differences were significant across target sizes on both the control and low-load trials (*ps* < .001) with small target trials showing the greatest deviations followed by medium and large target trials.

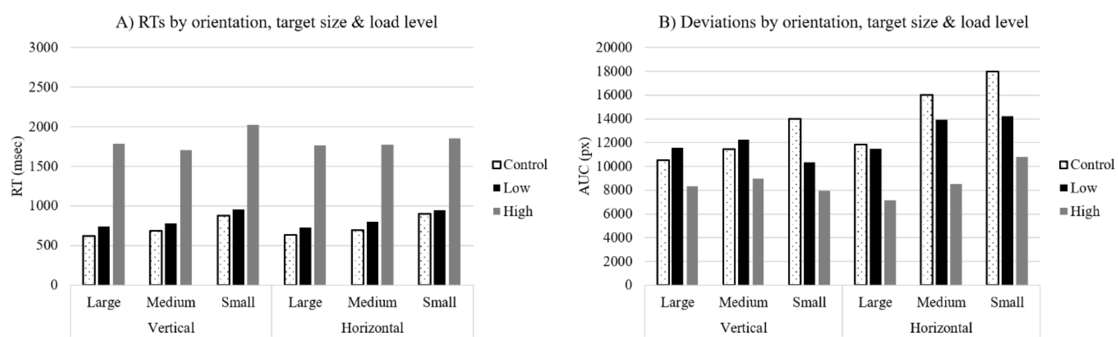

Figure 19. Descriptive data for Experiment 5. A) RT data. B) AUC data.

AUC. Main effect of WM load was significant, *F*(2, 78) = 4.49, $p < .05$, $\eta_p^2 = .10$. Simple effect analyses indicated that a greater deviation was present when participants were performing the control task (*M* = 13,634 px), and the low-load task (*M* = 12,309 px), compared to when they were performing the high-load task (*M* = 8624 px, *ps* < .054). However, the difference in deviations between the control task and the low-load task was nonsignificant, $p = 71$. Main effect of target size was also significant, $F(2, 78) = 7.60$, $p < .01$, $\eta_p^2 = .16$. Tukey's HSD test indicated that participants showed greater deviations when they were reaching for a small target ($M = 12{,}549$ px), and a medium target ($M = 11{,}859$ px), compared to when reaching for a large target ($M = 10{,}159$ px), *ps* $< .05$. However, the difference in deviations between small target trials and medium target trials was nonsignificant, $p = .49$.

WM load interacted with target size, $F(4, 156) = 2.44$, $p < .05$, $\eta_p^2 = .059$. On trials participants performed the control task, target size effect was present, $F(2, 78) = 10.7$, $p < .0001$, $\eta_p^2 = .26$, suggesting that the greatest deviation level was obtained on small target trials. The deviation level on small target trials was marginally greater than that on the medium target trials, $p = .087$, and greater than that of the large target trials, $p < .001$. Contrarily, Target size effect was nonsignificant on trials where participants performed the low-load task, $F(2, 78) = .71$, $p = .49$. Target size effect was marginally significant on high-load trials, $F(2, 78) = 2.62$, $p = .079$, $\eta_p^2 = .063$, because a marginally greater deviation was obtained for small target trials compared to large target trials, $p = .052$. No other differences were significant, *ps* $> .31$. In return, target size interacted with movement orientation, $F(2, 78) = 6.07$, $p < .01$, $\eta_p^2 = .13$, mainly because the target size effect was nonsignificant on vertical trials, $F(2, 78) = .57$, $p = .57$, but significant on horizontal trials, $F(2, 78) = 11.74$, $p < .0001$, $\eta_p^2 = .23$. On horizontal trials, a greater deviation was obtained on small and medium target trials compared to large target trials, *ps* $< .05$. The difference in deviations between the small and medium target trials was nonsignificant, $p = .21$. No other main effects or interactions were significant.

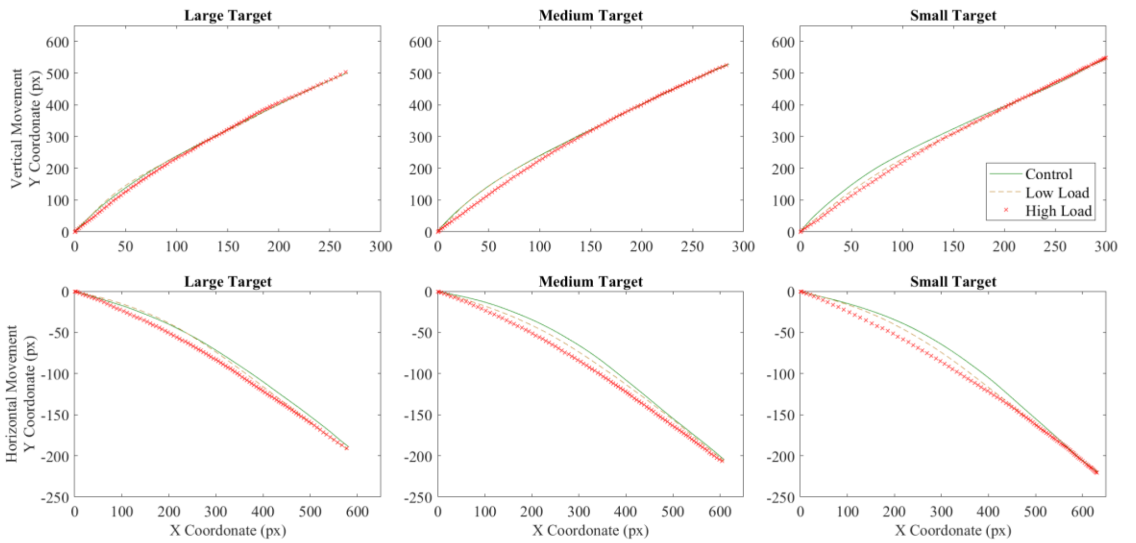

Figure 20. Mean trajectories as a function of movement orientation, WM load, and target size in Experiment 5.

Velocity and angle within a trial window. Figures 21 and 22 indicate mean velocities at each time-step within a trial window. The shaded areas represent within-subject standard errors (Cousineau, 2005), and the gap between the shaded areas larger than the standard error indicates significance at the $p < .05$ level (Cumming, Fidler, & Vaux, 2007). Moreover, Figure 23 indicates velocities and angles averaged by working memory load (Figure 23A & C) and target size conditions (Figure 23B & D). The green bars at the bottom of each plot in Figure 23 indicate ranges of time steps in which velocity and angle differences between conditions were significant at $p < .0005$ (Bonferroni corrected α level). The comparison of velocities across working memory load conditions indicated that the participants were significantly slower in the high-load task throughout the whole trial window than when they were performing the low-load task or the control task (Figure 23A). Mean velocities between target sizes at each time step were compared next. When the target size was small, there was a sharp increase in velocity in the first half of each trial, followed by a sharp decrease in velocity in the latter half of the trial. In contrast, the velocity change was less steep when the target size was medium, followed by when the target size was large, which resulted in crossovers of velocity curves between the 31~36th movement steps (the nonsignificance range in Figure 23B). In addition, the differences in velocities between working memory load conditions were the largest in the early phase of a trial (30~40th steps, Figure 23A), whereas the differences between target sizes were the largest in the later phase of a trial (90~101th steps, Figure 23B), which might suggest that working memory load effect preceded the target size effect. In order to confirm these findings, three-way ANOVAs on the velocity peak onset were performed with WM load, target size, and movement orientation set as within-subject variables. The results revealed a main effect of target size, $F(2, 78) = 424$, $p < .0001$, $\eta_p^2 = .92$. Pairwise comparisons indicated that velocity peaked earliest on the small target trials followed by the medium target trials, $p < .0001$, which in turn peaked earlier than on the large target trials, $p < .0001$. However, the velocity peak time did not vary as a function of working memory load, $F(2, 78) =.84$, $p = .43$. No other main effects or interactions were significant.

On the other hand, the examination of trajectory angles by working memory load (Figures 22, 23C) at each time step showed that participants tended to deviate away from the direction of the target in the early phase (1~19th steps, Figure 23C) when they were performing the control task or the low-load task compared to the high-load task. Target size similarly affected the trajectory angle, but in a slightly later phase (17~56th steps, Figure 23D).

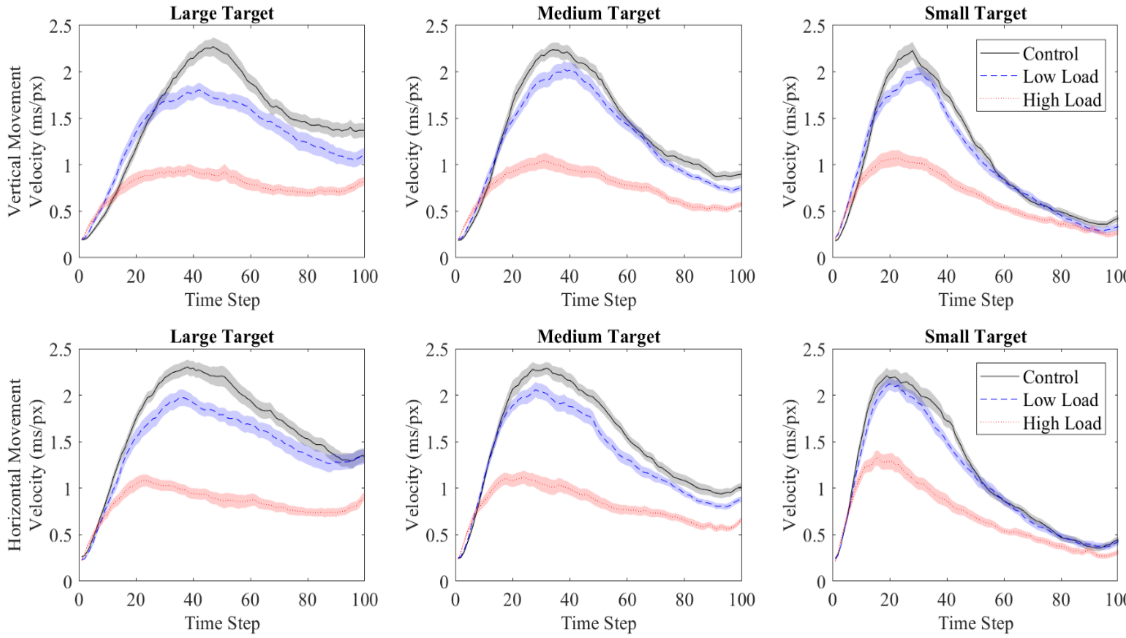

Figure 21. Mean velocities in a trial window as a function of movement orientation, WM load, and target size in Experiment 5.

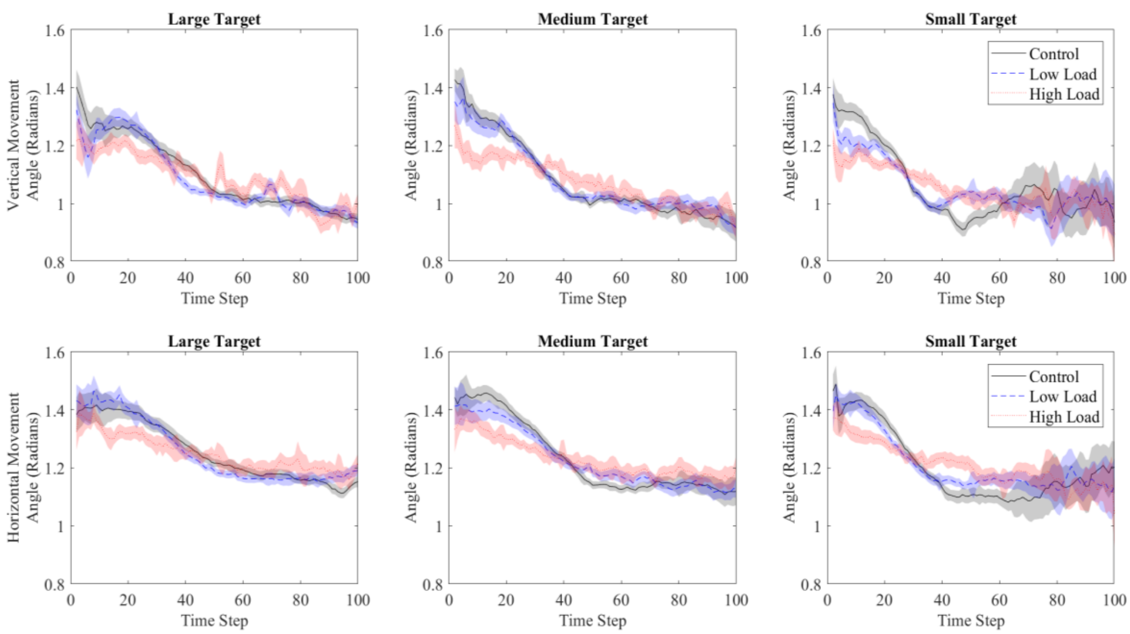

Figure 22. Mean trajectory angles as a function of movement orientation, WM load, and target size in Experiment 5.

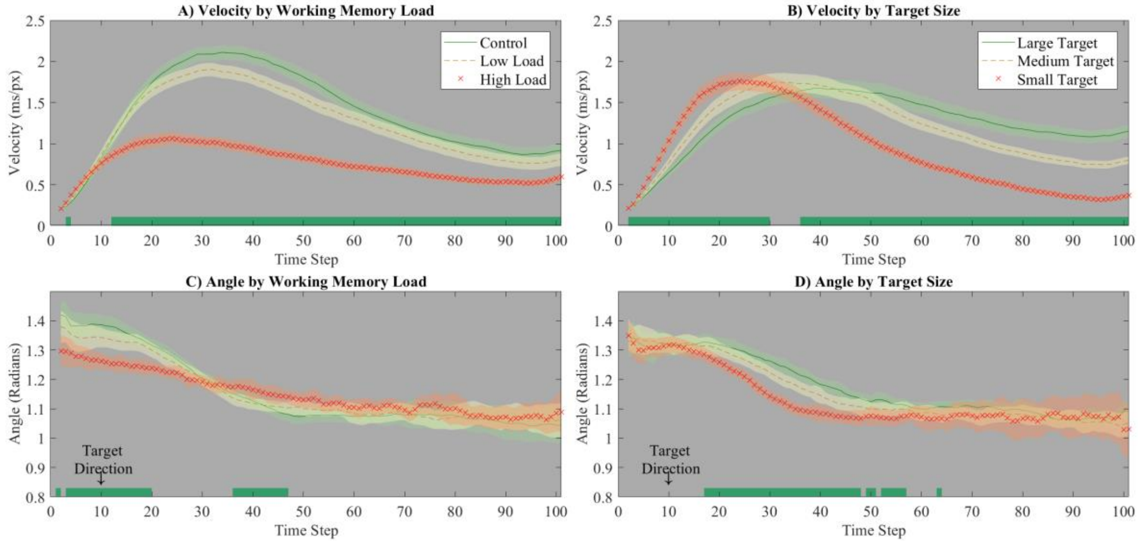

Figure 23. Mean velocities and angles at each time step by the condition. The green bars at the bottom of each plot indicate significant differences across conditions at $p < .0005$. A) Velocity by WM load. B) Velocity by target size. C) Angle by WM load. D) Angle by target size.

Classification. The feature selection procedure identified thirteen effective predictors of WM load which included, RT, initiation time, area under curve (AUC), mean velocity, maximum velocity, velocity peak onset, onset of the lowest acceleration, Y overshoot, mean velocity at Q1, mean acceleration at Q1 and Q2, and mean trajectory angle at Q1, and Q4. These thirteen features showed a mean prediction accuracy of 53.24%. Moreover, Spearman's rank correlation between the predicted and the observed values (secondary task) was $r_s = .45$, $p < .001$. This performance score was slightly better than the model fitted with the entire 39 features. The model fed with all 39 features showed a prediction accuracy of 52.68% and a Spearman's rank correlation of $r_s = .45$, $p < .001$. Lastly, none of the models from the 1000 permutation tests outperformed the model fitted in Experiment 5, suggesting that the observed performance was not due to random error. Altogether, the results from working memory load classification suggested that 13 of the 39 mouse trajectory features could be used to predict the level of working memory load at accuracy about 20% point greater than the chance level (33.3%) and could predict about 20% of the variance of the working memory load.

Experiment 6 Results

Manipulation check. Contrary to Experiment 5, participants were presented the Paas subjective rating scale and instructed to report the amount of mental resources they devoted to the secondary tasks at the end of each block. They were explicitly told not to consider any physical difficulties they experienced when reporting to the scale. Non-parametric, Kruskal-Wallis test was used on the ratings to examine the difference between working memory load conditions. As a result, participants reported a mean rating of 1.66 for the control task (*SD* = 0.73; close to "very low mental effort"), 3.73 for the low-load task (*SD* = 1.32; close to "rather low mental effort"), and 6.15 for the high-load task (*SD* = 1.24; close to "rather high mental effort"), $x^2(2) = 89.6$, $p < .0001$, $\eta^2 = .43$ (Figure 24). Planned pairwise comparisons verified that the mean rating for the high-load task was significantly greater than that of the low-load task, $p < .0001$, which in return, was greater than that of the control task, $p < .0001$. Therefore, this result confirms that the working memory load manipulation used in Experiments 5 and 6 was successful in inducing different levels of WM load because the secondary tasks were identical in both experiments.

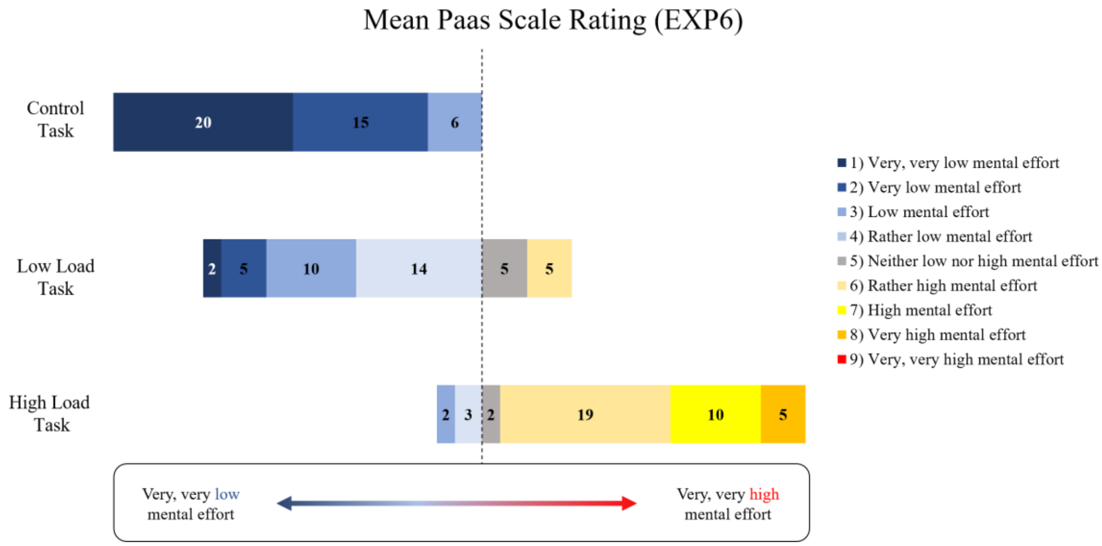

Figure 24. Stacked bar chart showing the relative frequency of subjective mental load ratings, with each bar divided according to the secondary task (working memory load task). For example, 20 out of 41 participants reported that the control task required very, very low mental effort, 15 participants reported that it required very low mental effort, and 6 participants reported that it required low mental effort.

RT. Descriptive RT and trajectory deviation data are shown in Figures 25~26. Three-way repeated-measures ANOVAs on RT data revealed a main effect of WM load, $F(2, 80) = 45.3$, $p < .0001$, $\eta_p^2 = .53$. Simple effect analyses indicated that participants were slower to reach a target when they were performing the high-load task ($M = 2786$ ms), compared to the low-load task ($M = 1775$ ms), $p < .0001$, and the control task ($M = 1952$ ms), $p < .0001$. Participants were also slower when they were performing the control task, compared to the low-load task, $p < .01$. Main effect of target size was significant, $F(2, 80) = 8.85$, $p < .0001$, $\eta_p^2 = .18$, because participants were slower when reaching for a small target ($M = 2252$), than for a medium target ($M = 2125$ ms), $p < .001$, or for a large target ($M = 2135$ ms), $p < .05$. However, RTs to medium and large targets did not differ significantly, $p = .93$. The main effect of movement orientation was also significant,

$F(1, 40) = 17.9$, $p < .0001$, $\eta_p^2 = .31$, which indicated that participants were slower to perform the horizontal movement task (M = 2265 ms) than the vertical movement task, (M = 2077), $p < .0001$. No other main effects or interactions were significant.

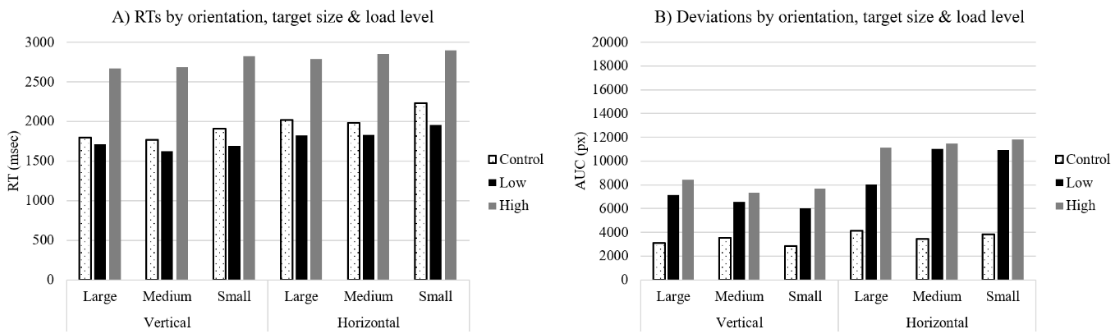

Figure 25. Descriptive data for Experiment 6. A) RT data. B) AUC data.

AUC. Again, main effect of WM load was significant, $F(2, 80) = 18.5$, $p < .0001$, $\eta_p^2 = .32$. Pairwise comparisons indicated that participants showed greater trajectory deviations when they were performing the high-load task (M = 9633 px) than when performing the control task (M = 3485 px), $p < .0001$. Participants showed numerically the second greatest deviation level when performing the low-load task (M = 8274 px), which did not differ significantly from the high-load task, $p = .36$, but was still greater than the control task, $p < .0001$. Main effect of movement orientation was also significant, $F(1, 40) = 7.73$, $p < .01$, $\eta_p^2 = .16$, because participants showed greater deviations when moving horizontally (M = 8410 px), than when moving vertically (M = 5851), $p = .01$.

Two-way interaction of WM load and movement orientation was significant, $F(2, 80) = 3.48$, $p < .05$, $\eta_p^2 = .080$. When participants performed the control task, the difference in deviations observed on the vertical and horizontal movement trials was nonsignificant, p = .40. Nevertheless, the differences were significant on low-load trials, p < .01, and high-load trials, p < .05, with the horizontal movement trials showing greater deviations than the vertical movement trials on both secondary tasks.

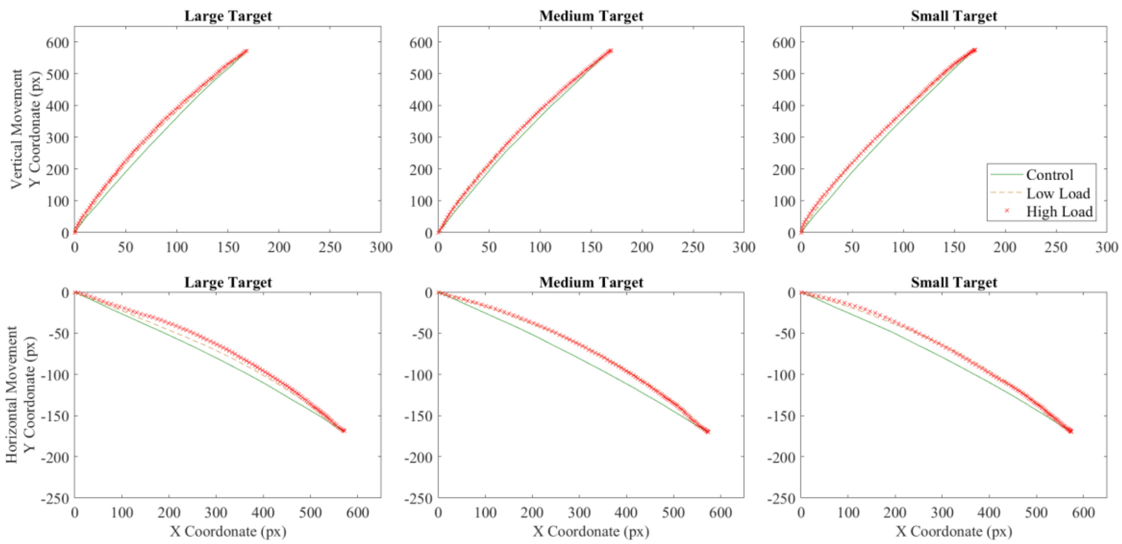

Figure 26. Mean trajectories as a function of movement orientation, WM load, and target size in Experiment 6.

Velocity and angle within a trial window. Figures 27 and 28 indicate mean velocities and angles at each time-step within a trial window. Similar to Experiment 5, participants were significantly slower in the high-load task throughout the whole trial window than in the low-load task or the control task. Moreover, the differences between working memory load conditions were the largest in the early phase of a trial (10~40th steps, Figure 29A), whereas the differences between target sizes were the largest in the later phase of a trial (60~80th steps, Figure 29B), which replicated the findings from Experiment 5. Three-way ANOVAs on the velocity peak time were performed with WM load, target size, and movement orientation set as within-subject variables. The results revealed a main effect of WM load, $F(2, 80) = 10$, $p < .0001$, $\eta_p^2 = .20$. Velocity peaked earliest in both the control and low-load tasks, which did not differ significantly, *p* .=88. In return, the velocity in both tasks peaked earlier than in the high-load task, *ps* < .01. The main effect of target size was significant as well, $F(2, 80) = 37.7$, $p < .0001$, $\eta_p^2 = .49$, with velocity peaking earliest on small target trials, than on medium target trials, $p < .0001$, or large target trials, $p < .0001$. Like in Experiment 5, velocity peaked earlier on medium target trials than on large target trials, $p < .05$. Main effect of orientation was also significant, $F(1, 40) = 4.55$, $p < .05$, $\eta_p^2 = .10$, because velocity reached the maximum earlier on horizontal movement trials, $p < .05$.

In contrast, trajectory angle data by WM load suggested that the deviation in trajectory angles at the initial phase (1~23th steps, Figure 29C) was greater when participants performed the low-load and high-load tasks compared to when they performed the control task. The trajectory also seemed to have deviated more away from the target direction in the early phase of each trial when participants were reaching for a small target compared to when they were reaching for a medium or a large target, although this difference was nonsignificant. Eventually, this led to a mid-flight maneuver on small target trials, possibly, in order to correct the trajectory angle toward the target direction (26~39th steps, Figure 29D), the angle difference of which was significant across target size conditions.

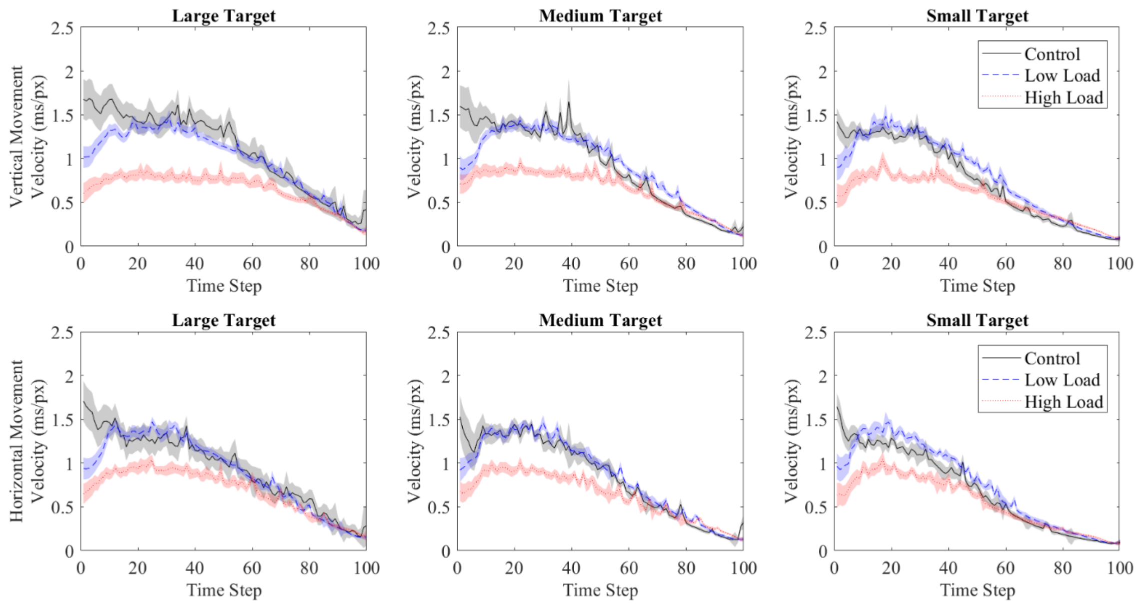

Figure 27. Mean velocities in a trial window as a function of movement orientation, WM load, and target size in Experiment 6.

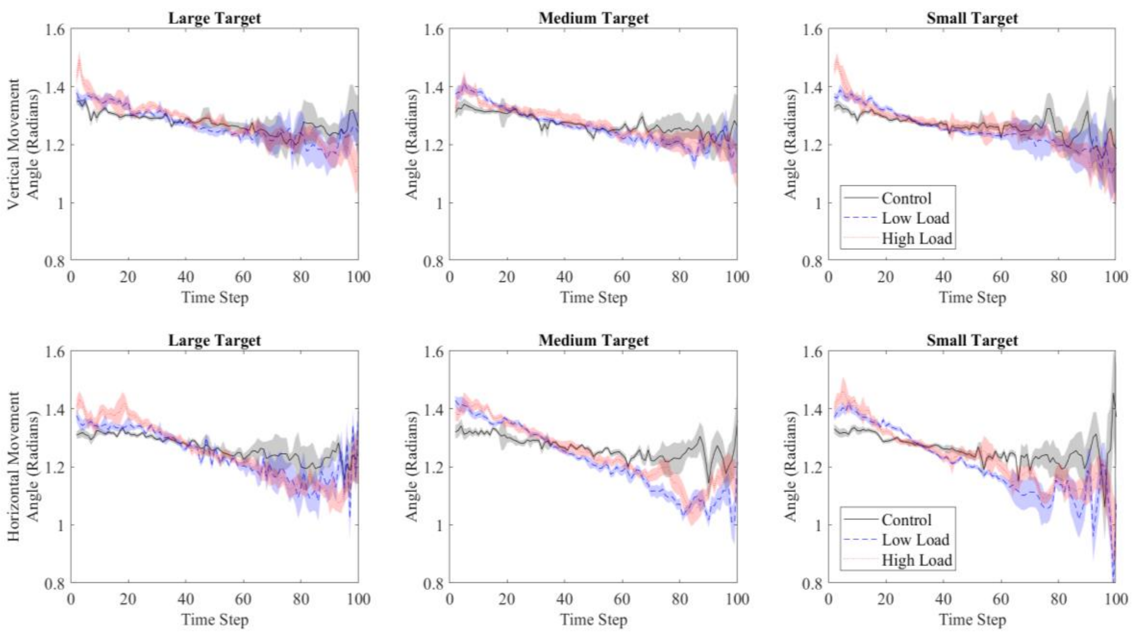

Figure 28. Mean trajectory angles as a function of movement orientation, WM load, and target size in Experiment 6.

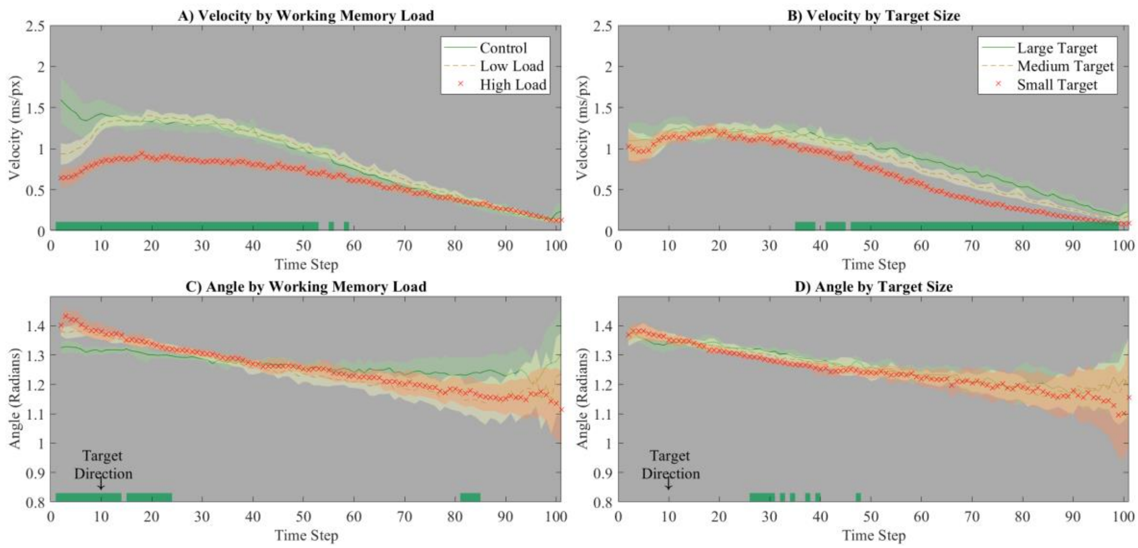

Figure 29. Mean velocities and angles at each time step by the condition. The green bars at the bottom of each plot indicate significant differences across conditions at $p < .0005$. A) Velocity by WM load. B) Velocity by target size. C) Angle by WM load. D) Angle by target size.

Classification. Ten effective predictors were identified in Experiment 6, which included initiation time, absolute max deviation, trajectory length, mean and maximum velocities, velocity peak onset, Y entropy, mean velocities at Q1, Q2, and Q3. These ten features showed a mean prediction accuracy of 54.24% and a Spearman's rank correlation of $r_s = .36$, $p < .001$. Again, the model fed only with the selected features outperformed the model fed with the entire features which showed a prediction accuracy of 53.35% and a Spearman's rank correlation of $r_s = .35$, $p < .001$. Lastly, none of the models from 1000 permutation tests outperformed the model fitted in Experiment 6, suggesting that the observed performances of the classifiers were not due to random error. An additional classification was performed in Experiment 6, which sought to investigate if hand-motion trajectory features could also be used to predict the Pass subjective rating, instead of the working memory tasks participants performed. The feature selection procedure identified twelve effective predictors, including initiation time, absolute max deviation, length of the trajectory, mean and maximum velocities, onsets of the velocity and acceleration peaks, X flip, 2D entropy, mean velocities at Q1, and Q2, and mean angle at Q1. The prediction accuracy was 34.40%, which was greater than the chance level (11.1%) or the accuracy of the model fed with the entire features (33.92%). Spearman's rank correlation for the model of the selected features was, $r_s = .34$, $p < .001$, again outperforming the model of the entire features which showed a correlation of $r_s = .33$, $p < .001$.

Discussion

The results from Experiment 5 showed intriguing tradeoffs. Higher working memory load elicited slower responses, but smaller mouse trajectory deviations. The motor task difficulty manipulation also elicited slower responses but yielded greater trajectory deviations. Examination of cursor's velocity change within a trial window provided further insights on why working memory load and motor task difficulty produced quantitatively different results. When the target size was small, participants showed a faster initial movement, which led to the overshooting of the mouse cursor movement. The overshooting led the cursor to deviate from an ideal trajectory, which the participants had to correct by slowing down, resulting in a greater trajectory deviation and a slower response. In contrast, when the working memory load was high, the overall response was slower throughout a trial window. As a result, the slowed response time did not lead to a greater deviation, which was supported by the less abrupt change in trajectory angle over time. Moreover, the examinations of velocities and trajectory angles at each time step indicated that the influence of working memory load preceded that of the target size.

Experiment 6 revealed mixed results. As in Experiment 5, participants were slower when reaching for a smaller target, or when performing a task imposing a higher level of working memory load. However, a greater mean trajectory deviation was obtained for a task imposing a higher working memory load, contrary to Experiment 5, where participants showed a greater mean deviation in a task imposing a lower working memory load. As in Experiment 5, nevertheless, the working memory load effect preceded the target size effect.

The results from the classification procedures indicated that several features could be used to infer the level of working memory load. Specifically, RT, initiation time, area under curve (AUC), mean velocity, maximum velocity, onset of the maximum velocity, onset of the lowest acceleration, Y overshoot, Q1 velocity, Q1 and Q2 accelerations, and Q1 and Q4 trajectory angles were identified as useful in predicting the overall working memory load of participants interacting with a mouse cursor. For participants interacting with a touchscreen, initiation time, absolute max deviation, trajectory length, mean and maximum velocities, onset of the velocity peak, Y entropy, mean velocities at Q1, Q2, and Q3 were identified as effective. The classifiers from Experiments 5 and 6 showed prediction accuracies that were about 19 ~ 23.3% point greater than the chance level (11~33%) on average and could explain about 12 ~ 20% of the variance of the working memory load. Moreover, none of the models from 1000 random permutation tests outperformed the performances of the models fitted using the real data, suggesting that the observed performances of the classifiers in the current Experiments were not due to random error.

General Discussion

The effects of working memory load and motor task difficulty were investigated in this chapter to be able to dissociate these two effects. While consistent patterns were observed from the temporal features of the hand motion trajectory (i.e. RT & velocity), mixed results were found from the spatial features (i.e. AUC & trajectory angle). Specifically, RTs increased and mean velocities decreased as a function of working memory load both in Experiments 5 and 6, which is a result consistent with the study by Grimes and Valacich (2015). Nevertheless, AUC decreased with working memory load in Experiment 5, whereas it increased with the load in Experiment 6.

Smaller Deviations with Increased Working Memory Load

The different AUC results observed in Experiments 5 and 6 seem to reflect the confounding effect caused by the distinct characteristic of the hand-mouse cursor interaction. A computer mouse cursor movement can be distinguished from a drag-and-drop gesture on a touchscreen, in that the mouse cursor only requires a slight movement of a hand to move the cursor to the target location while the touchscreen gesture requires a user’s finger or hand to be moved to the target location. Moreover, in the current experiments, the drag-and-drop gesture required a finer motor control because participants’ fingers could occlude the parts of the virtual cursor and the target circle. This difference in cursor-hand interaction (movement gain, occlusion) might have contributed to the obtained results by amplifying the overshoot of a trajectory like the ones observed in Experiment 5. Through the examinations of velocities at each time step, it was concluded that a faster initial movement led to a steep decrease in velocity in a later phase, a phenomenon which was coined as the trajectory overshoot. Consider the velocity plot from Experiment 5 (e.g. Figure 23B). This overshoot, elicited by the motor task difficulty (target size), affected velocities of trajectories both in the initial and the later phases in a trial such that the velocity on small target trials was the fastest in the initial phase (3~30th Steps), and the slowest in the later phase (37~101th steps). The initial phase where the overshoot affected velocity also overlapped with the phase in which working memory load affected velocity (13~101th steps, Figure 23A) and trajectory angle (10~46th steps, Figure 23C). Accordingly, it is possible that motor task difficulty and working memory load interacted in the early phase of a mouse cursor movement in Experiment 5 that led to the observation that the trajectory deviation decreased as working memory load increased. This possibility is supported by the two-way interaction of target size and WM load obtained in Experiment 5, which was absent in Experiment 6. Specifically, the differences in AUCs between WM load conditions were the largest on small target trials, which showed the largest magnitude of overshoot (Figure 30). Therefore, the greater deviations on control task trials might have been driven by this effect, rather than by working memory load alone. Moreover, the overall mean RT in Experiment 5 (M = 1127 ms) was about the half of the overall mean RT in Experiment 6 (M = 2171 ms) or the minimum level of RT (M = 1938 ms) in the study by Grimes and Valacich (2015). These findings suggest that the faster, and hasty responses made in Experiment 5 might have left less time to correct the overshot distance, and hence amplified the trajectory deviations on the control and low-load task trials, the mean RTs of which were significantly faster than the mean RT of the high-load task trials.

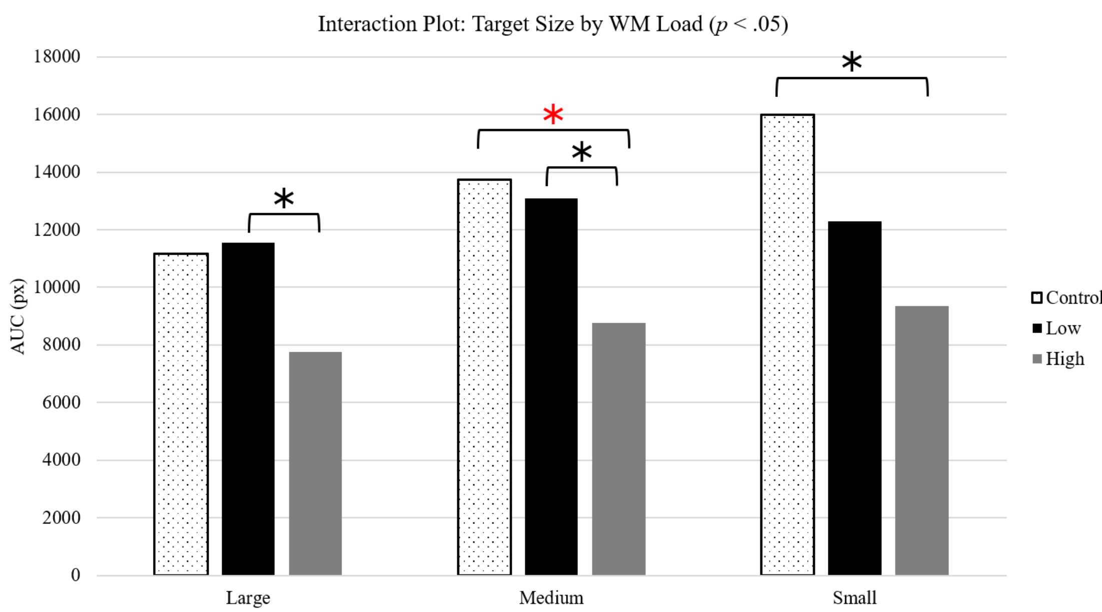

Figure 30. Two-way interaction of target size and WM load in Experiment 5.

Two Phases of the Hand Motion

Another finding worth noting is that the influence of working memory load always preceded the influence of target size or the motor task difficulty. Specifically, working memory load affected the trajectory angle in an initial phase, which contributed to the trajectory deviation (Figures 23C & 29C). In return, hand motions deaccelerated in a later phase as cursors were reaching closer to the target location. Furthermore, the deacceleration was the greatest on small target trials as if participants were carefully coordinating the cursor’s location with the target location (Figures 23B & 29B). This finding is in line with the stochastic optimized submovement (SOS) model (Meyer et al. 1988; Meyer et al. 1990) which asserts that a coordinated motor movement is comprised of an initial phase of inaccurate and fast movements that closes the gap between the current and target locations, followed by a later phase of slow and deliberate movements purposed to correct the errors made in the initial phase. According to the SOS model, it is possible that working memory load served as the noise in the initial phase that deviated hand movements away from the direction of the target. In the later phase, this deviation was corrected during an effort to adjust the trajectory in accordance with the target's accessibility. This would imply that the effect of working memory load can be distinguished from the effect of motor task difficulty by examining the time in which the effects occurred. Additional evidence supporting this finding was obtained from the classifications of working memory load conditions. Only 5 out of 39 hand motion features were unanimously recognized as effective predictors by all three classifications performed throughout Experiments 5 and 6, two of which were the features reflecting the movements made in an initial phase: initiation time and Q1 velocity (see Appendix G).

However, the current findings also warn that hand motions be carefully examined, as working memory load had opposite effects on the use of different input devices in Experiments 5 and 6. Moreover, previous research raises the possibility that different input devices require different levels of working memory load. The elderly population is usually considered as having limited working memory capacity compared to the younger population (Hartman, Bolton & Fehnel, 2001; Hedden & Gabrieli, 2004; Mattay et al., 2006), and performs worse in tasks requiring a computer mouse cursor. However, the performance gap between the elderly and younger populations decreases when a given task involves a touchscreen (Findlater et al., 2013), suggesting that touchscreen gestures are dependent more on the experience with the input-devices, or other cognitive functions than the working memory. That is, a touchscreen may not require excessive mental resources because it takes input from gestures that are frequently practiced in real-life. Nevertheless, the drag-and-drop gesture used in Experiment 6 is often reported as an exception to this pattern as the performance gap does not decrease to the extent that the gaps in other touchscreen gestures decrease (Findlater et al., 2013), implying that the drag-and-drop might be a gesture dependent upon working memory capacity. Accordingly, caution is required when comparing different types of hand motions, because the trajectory deviation is a sensitive measure that is affected by various factors (Kieslich et al., 2019) such as the gain value of a computer mouse cursor.

In conclusion, the results from Experiments 5 and 6 suggested that working memory load affects a hand motion trajectory by introducing motor noises in the early phase of the hand motion, which, in return, induce a greater trajectory deviation and a slower response time.

Conclusion and Implication

So far, Chapters 1 and 2 demonstrated that inferences about an individual's implicit bias could be made by assessing the hand-motions of that individual. Additional efforts were made in the current chapter to evaluate if hand-motions of participants could also be leveraged to make inferences about an individual's capacity to suppress implicit bias (e.g. working memory). It was anticipated that working memory load would impede participants' task performances, defined as the time they took to complete the task, and the deviations in hand-motions. The results indicated that a higher working memory load impeded participants' performances, but the impediment also depended on the type of input devices and the physical layout of the display (i.e. size of targets). These findings also implied that considerations about the motor task difficulty should be made (e.g. Fitts's law) to distinguish the hand-motion features that reflect working memory load. Altogether, the current chapter demonstrated that inferences could also be made about an individual's capacity to suppress implicit bias based on the individual's hand-motion trajectories. Accordingly, the hand motion tracking method seems to be a reliable measurement for both the implicit bias and capacity to suppress the bias.

The hand-motion tracking method has wide applicability and substantial potential to improve performances in tasks involving social interactions. It can be plugged into training systems like the virtual training environment used in the study by Zipp and Craig (2019) and provide rich dimensions of information about the performances of trainees. Furthermore, the hand motion tracking method has a strength in that it can monitor complex body movements in virtual environments. In return, virtual environments, equipped with a hand motion tracking method, can provide opportunities to enhance performances in situations that are often dangerous and difficult to simulate in the real physical world (i.e. triaging, firefighting operation).

CHAPTER 4

Overview: Experiments 7 & 8

Experiments 7 and 8 were conducted to investigate if implicit bias is handled by cognitive control mechanisms and if working memory load modulates the implicit association effect through the cognitive control mechanisms. To this end, a Stroop task was integrated into the mouse-tracking paradigm with a working memory load manipulation. Faces varying in gender, race, and emotional expression served as the Stroop stimuli, and participants were instructed to report the task-relevant identity of these faces. Using this design, three different signatures of cognitive control involvement were sought, which were the Stroop-like effect, conflict adaptation effect, and load modulation of the conflict adaptation effect. In Experiment 7, participants performed a gender decision task and an emotional expression decision task. Faces with overlapping features (Angry male faces & happy female faces) served as congruent stimuli, and faces with non-overlapping features (happy male faces & angry female faces) served as incongruent stimuli. In Experiment 8, participants performed a gender decision task and a race decision task. In this experiment, black male faces and white female faces served as congruent stimuli, whereas white male faces and black female faces served as incongruent stimuli. Therefore, the Stroop-like effect observed in these experiments was considered as equivalent to the implicit association effect. At the same time, participants were given an additional secondary task of memorizing and rehearsing numbers (in the memory task display) presented before target faces on each trial. A single number was presented next, after the offset of the target face display, which required participants to report whether the number was present in the memory task display. The number of digits in the memory task display was either one (low-load) or six (high-load), as in Lavie et al. (2004), thereby manipulating the working memory load level.

Signatures of Cognitive Control Mechanisms

In order to determine whether implicit associations are handled by cognitive control mechanisms, the following signatures were sought. This section describes each signature in detail and provides predictions for Experiments 7 and 8 based on previous literature.

Stroop(-like) effect. One of the main hypotheses of the current dissertation was that the implicit association effect is an effect similar to the Stroop effect which refers to a delayed response to a multidimensional stimulus with conflicting features (e.g. the word "Blue" colored in red ink; Frings et al., 2010). The term "Stroop effect" has been used mainly in cognitive psychology research, whereas the term "implicit association effect" has been used in social psychology research and not interdisciplinarily, which might be one reason why these concepts have not been compared often.

Studies that used some variants of Stroop stimuli often faced criticism that the Stroop-like effects reported in those studies were effects different from the Stroop effects observed in classic Stroop task studies (Algom, Chajut, & Lev, 2004). Emotional Stroop stimuli, for example, are comprised of either emotionally valenced words or neutral words that are inked in different colors. When participants are asked to report the color or the name of the emotional Stroop words, their responses are slowed down when the words are the emotionally-charged words compared to the neutral words, which yields an effect coined the emotional Stroop effect. Nevertheless, Algom et al. (2004) disagreed that the emotional Stroop effect was a type of the Stroop effect, and instead suggested that the slowing down of responses to emotional words reflected the processing of threat-related information that often entailed temporary disruption of mental operations. They suggested five diagnostics of the Stroop effect, which mainly posited that a Stroop-like effect, in order to qualify as a Stroop effect, should change depending on the changes in the features of task-irrelevant dimensions, and the relative salience between the two dimensions (e.g. the name and color of a word) of stimuli tested, and that these patterns should remain unaffected by the slowdown of mental operations elicited by the threat-related information processing.

Therefore, it was important to show, first, that an implicit association could elicit an effect like the Stroop effect, although the Stroop effect is not a direct signature of cognitive control mechanisms. Rather, it is a phenomenon addressed by the mechanisms. Accordingly, it was anticipated that counter-stereotypical associations such as an angry female face or a white male face would slow down participants' responses compared to faces of stereotypical associations (i.e. happy female, black male), leading to a significant Stroop-like effect. If the Stroop stimuli and implicit associations were different in nature, the Stroop-like effect (or congruency effect) should be absent, or be present but fail to elicit a conflict adaptation effect. In the current experiments, the main effect of facial feature congruency (N^{th} trial congruency) should be significant if implicit associations elicited a Stroop-like effect.

Conflict adaptation effect. Another hypothesis of the current study was that implicit biases are handled by the general cognitive control functions. The conflict adaptation effect or the Gratton effect refers to a reduced interference effect (Stroop effect) by task-irrelevant features observed after incongruent trials (Gratton, Coles, & Donchin, 1992; Figure 31A). In a typical Stroop task setting, the Stroop effect on the N^{th} trial is reduced when an incongruent target is presented on the $N\text{-}1^{th}$ trial than when a congruent target is presented on that trial. This decrease in the Stroop effect is considered as a signature that cognitive control functions were engaged to resolve the cognitive conflict elicited by the incongruent target (Botvinick et al., 2004). Therefore, the prediction derived from this hypothesis would be that the faces consisted of counter-stereotypical associations would trigger cognitive control mechanisms to engage. When triggered, the cognitive control mechanisms would reduce the Stroop-like effect or the implicit association effect elicited by the face presented on the next trial. This result should be confirmed by a two-way interaction of $N\text{-}1^{th}$ trial congruency and N^{th} trial congruency.

Working memory load modulation. Additional evidence linking the implicit association effect to cognitive control mechanisms can be found by examining whether the conflict adaptation effect is moderated by working memory load. Previous research has argued that working memory holds task representations (e.g. description of features to be attended or ignored; Braver et al., 1997; Kane & Engle, 2003), which can be interrupted by a high working memory load (Soutschek, Strobach & Schubert, 2013). Accordingly, the cognitive control mechanisms under a high working memory load would leave the Stroop stimuli at the risk of the distraction by task-irrelevant features and decrease the conflict adaptation effect. Soutschek et al. (2013) used a dual-task paradigm in which participants performed different types of working memory tests, while concurrently performing a Stroop task. Their finding was that a higher working memory load led to a reduced conflict adaptation effect (but see their Experiment 3 for contradictory evidence). Therefore, if working memory load interrupts the cognitive control functions, a greater Stroop effect should be observed under a higher working memory load than under a lower load. Furthermore, the conflict adaptation effect should be reduced or eliminated under a higher working memory load (Figure 31B). Lastly, these hypotheses would be confirmed by significant interactions between working memory load and the Stroop effect (Nth trial congruency), and between working memory load, and the conflict adaptation effect (Nth trial congruency * N-1th trial congruency).

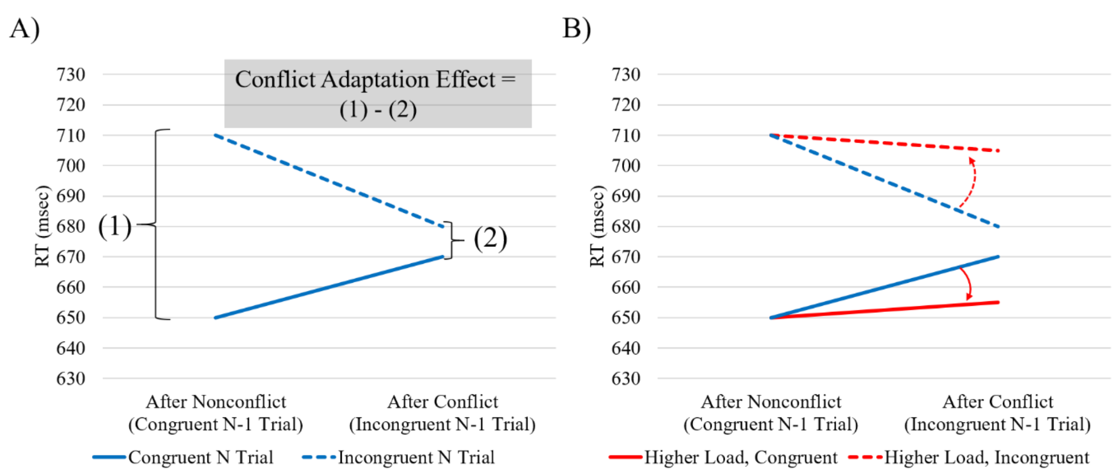

Figure 31. Illustration of a conflict adaptation effect under different levels of working memory load. A) Conflict adaptation effect expected under a low or moderate level of load. B) Anticipated change in the conflict adaptation effect in response to a higher load level. The red lines describe the change introduced by a higher working memory load.

Method

Participants

Forty-eight under undergraduate students who identified themselves as not being of African American background participated in Experiments 7 (25 Caucasian American, 19 East Asian, 3 Middle Eastern, and 1 American Indian participants; Mean age: 20.9 years). Moreover, fifty male undergraduate students participated in Experiments 8 (33 Caucasian American, 8 East Asian, 8 Middle Eastern, and 1 American Indian participants; Mean age: 20.2years). All participants had normal or corrected-to-normal vision and were right-handed. They were offered 1 course credit for their participation, which lasted up to one hour.

Apparatus and Material

The experiment took place in front of a 20-in. LCD monitor with a screen resolution of 1600 by 900 pixels. All participants sat approximately 60 cm from the monitor and made responses by moving a computer mouse cursor and clicking one of the two response boxes that appeared at the upper corners of the screen. The mouse-tracking program was created using the OpenSesame software (Version 3.2.5; Mathôt, Schreij, & Theeuwes, 2012) and the Mousetrap plugin (Kieslich, & Henninger, 2017). The x and y coordinate data and the time data of mouse cursor movements were recorded at the sampling rate of 100hz. Images from the Chicago face database (Ma, Correll, & Wittenbrink, 2015) were used as the target faces. Eight faces from each gender and emotional expression category were selected as the target face in Experiment 7 (2 genders*2 emotional expressions*8 faces = 32 faces), and Eight faces from each gender and race category were selected as the target face in Experiment 8 (2 genders*2 races*8 faces = 32 faces). Note that the 32 faces used in Experiment 7 were from 16 actors (in contrast to 32 actors in Experiment 8), each displaying happy or angry expressions. In addition, the Paas subjective rating scale (Paas, 1992; Figure 17) was used again to check the working memory load manipulation adopted in the current experiments.

Procedure

The experimental sequence is described in Figure 32. Each participant performed four blocks of 66 trials. The working memory load conditions were blocked, and participants performed each block in alternate order. Half of the participants began with a low-load block, and the other half began with a high-load block. The four blocks included two gender decision blocks (gender task) and two emotional expression decision blocks (emotion task) for each working memory load condition in Experiment 7, and two gender decision blocks and two race decision blocks (race task) for each working memory load condition in Experiment 8. At the end of each block, participants were presented with the Paas subjective rating scale asking participants to report the amount of mental effort they devoted to each block. Each trial started with a 500-ms fixation display, followed by a memory task display which was presented for 1500 ms on high-load trials, and 750 ms on low-load trials. The memory task display contained a six-digit number on high-load trials and a single-digit number on low-load trials. Accordingly, the durations for the memory task displays were set differently in order to provide participants sufficient time to read all of the digits on the high-load trials. This design was also identical to that used in the study by Lavie et al. (2004), and the duration for each display was determined based on their pilot testing results. Participants were asked to rehearse the number covertly until the onset of the target probe display. After the offset of the memory task display, a 500-ms fixation display appeared again. The mouse-tracking display was presented next, which required participants to categorize the predefined feature of the target faces presented at the lower center of the screen, by choosing one of the two response options presented at the upper corners of the screen. All participants were instructed to make responses within 2000 ms and the display was automatically replaced by the target probe display 2000 ms after its onset. The target probe display contained a single-digit number and asked participants to report whether the number was present in the memory task display by clicking either the "Yes", or the "No" button, which appeared at the upper corners. Participants were instructed to make a response within 3000 ms because the display was replaced by another fixation display 3000 ms after its onset.

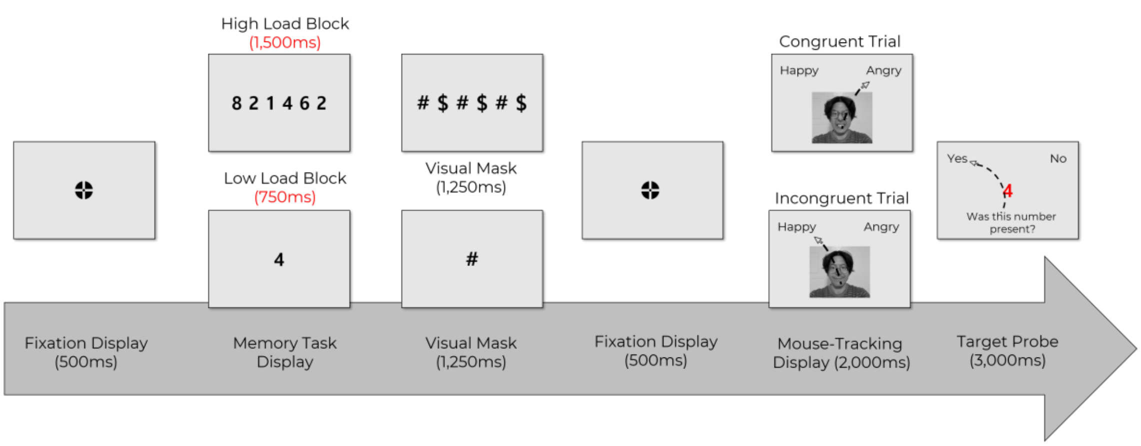

Figure 32. Example of a trial sequence in Experiments 7 and 8.

Analysis

Data were preprocessed to exclude incorrect trials and outlier trials (2.5SD criterion). An incorrect trial was defined as the trial in which a participant incorrectly categorized a target face but not as the trial in which a participant reported a target probe incorrectly. As a result, a total of 210 incorrect trials (1.7 %) was excluded in Experiment 7, and a total of 59 trials (0.4 %) was excluded in Experiment 8. The 2.5SD criterion applied to RT and movement flip also led to the exclusion of 487 additional trials (3.8 %) in Experiment 7 and 504 trials (3.8 %) in Experiment 8.

Four-way repeated measures ANOVAs were performed using N^{th} trial congruency (congruent, incongruent), $N\text{-}1^{th}$ trial congruency (congruent, incongruent), working memory load (WM load: low-load, high-load), and task (gender decision, race decision or emotional expression decision) as within-subject variables on response time (RT), and area under the curve data (AUC). Specifically, the stereotypical congruency between gender and emotional facial features on N^{th} trial was coded as the N^{th} trial congruency, and the congruency of features in the trial right before the N^{th} trial was coded as $N\text{-}1^{th}$ trial congruency. Therefore, a significant main effect of N^{th} trial congruency would indicate an occurrence of the Stroop (-like) effect, a significant two-way interaction of N^{th} trial and $N\text{-}1^{th}$ trial congruencies would indicate an occurrence of the conflict adaptation effect, and a three-way interaction of N^{th} trial and $N\text{-}1^{th}$ trial congruencies and working memory load would indicate that the conflict adaptation effect was moderated by working memory load.

In addition, conflict adaptation effects were calculated and were subject to two additional analyses. First, conflict adaptation effects were calculated by subtracting the Stroop effects observed after nonconflict (after congruent $N\text{-}1^{th}$ trial) from that observed after conflict (after incongruent $N\text{-}1^{th}$ trial). The conflict adaptation effects were then subject to two-way repeated-measures ANOVAs with task and WM load set as within-subject variables. The formula used to calculate the conflict adaptation effect is shown in Equation 1 (also see Figure 31A). Second, the conflict adaptation effect, calculated using velocity and trajectory angle data, were unfolded along the 101 time-steps to examine how the conflict adaptation effect fluctuated within a trial window. This approach was taken additionally to gain insights on when differences between experimental conditions in the conflict adaptation effect occurred if they existed.

$$\textit{Conflict Adaptation} = \text{Congruency Effect}_{\text{(after Nonconflict)}} - \text{Congruency Effect}_{\text{(after conflict)}} \quad (1)$$

Experiment 7 Result

Manipulation check. When performing the working memory tasks, participants showed a mean accuracy of 95.4% on low-load trials, and 86.4% on high-load trials. One-way ANOVA indicated that the difference in accuracies was significant, $F(1, 47) = 35.5$, $p < .0001$, $\eta_p^2 = .42$. Furthermore, the mean Paas subjective rating was 3.44 ($SD = 1.86$; close to "low mental effort") on low-load trials, whereas it was 5.35 ($SD = 1.85$; close to "neither low nor high mental effort") on high-load trials. The difference between the two mean ratings was also significant, $x^2(1) = 21.1$, $p < .0001$, $\eta^2 = .22$.

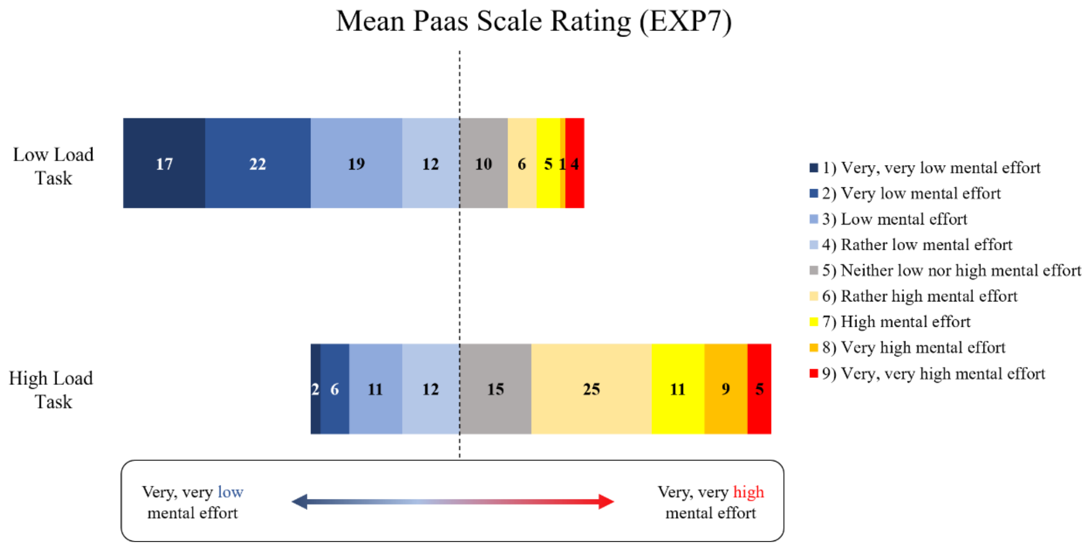

Figure 33. Stacked bar chart showing the relative frequency of subjective mental load ratings, with each bar divided according to the working memory load task.

RT. Figure 34 indicates mean RTs as a function of the independent variables in Experiment 7. Main effect of task was significant, $F(1, 47) = 58.5$, $p < .0001$, $\eta_p^2 = .55$, which indicated that mean RT was longer in the emotion task (M = 1387 ms) than in the gender task (M = 1257 ms). Main effect of Nth trial congruency was also significant, $F(1, 47) = 27.5$, $p < .0001$, $\eta_p^2 = .37$, because RT was longer on incongruent trials (M = 1339 ms) than on congruent trials (M = 1305 ms).

Nth trial congruency interacted with task, $F(1, 47) = 12.4$, $p < .0001$, $\eta_p^2 = .21$. Simple effect analyses at each task level revealed that congruency effect was marginally significant in the gender task (M = 15 ms), $p = .067$, and significant in the emotion task (M = 53 ms), $p < .0001$. Nth trial congruency also interacted with N-1th trial congruency, $F(1, 47) = 9.74$, $p < .01$, $\eta_p^2 = .17$. Further analyses indicated that the interaction was driven by a significant 53-ms congruency effect after nonconflict (N-1th trial congruent),

p <.0001, and a marginal 15-ms congruency effect after conflict (N-1th trial incongruent), p = .073, which confirmed the occurance of the conflict adpatation effect.

The conflict adaptation effect was moderated by task as suggested by the three-way interaction of task, Nth trial congruency, and N-1th trial congruency, $F(1, 47) = 4.81$, $p < .05$, $\eta_p^2 = .092$. Separate two-way ANOVAs at each task level suggested that a conflict adaptation effect (interaction of Nth incongruency and N-1th incongruency) was present in the gender task, $F(1, 47) = 16.8$, $p < .001$, $\eta_p^2 = .26$, but not in the emotion task, $F(1, 47) = .97$, $p = .33$. In the gender task, a 45-ms congruency effect was obtained after nonconflict (N-1th trial congruent), $p < .001$, and the congruency effect was reduced to a nonsignificant −15-ms congruency effect after conflict (N-1th trial incongruent), p =.19, resulting in the conflict adaptation effect of 60 ms.

Lastly, a four-way interaction was sigificant, $F(1, 47) = 15.2$, $p < .0001$, $\eta_p^2 = .24$. Separate three-way ANOVAs at each task level revealed a three-way interaction of WM load, Nth trial congruency, and N-1th trial congruency for the gender task, $F(1, 47) = 13.9$, p < . 001, $\eta_p^2 = .23$, and for the emotion task, $F(1, 47) = 4.4$, $p < .05$, $\eta_p^2 = .09$. The three-way interaction for each task was split again by WM load condition. When WM load was low in the gender task, a two-way interaction of Nth trial congruency, and N-1th trial congruency was significant which suggested the occurrence of conflict adpatation effect, $F(1, 47) = 25.3$, $p < .0001$, $\eta_p^2 = .35$, but the conflict adaptation effect was not found on high-load trials of the gender task, $F(1, 47) = .84$, $p = .36$. Specifically, on low-load trials of the gender task, a 67-ms congruency effect was observed after nonconflict, $p < .0001$, which was reduced to a −36-ms (reversed) congruency effect after conflict, $p < .05$, both of which contributed to the 103-ms conflict adaptation effect. In the emotion task, conflict adaptation effect was absent on low-load trials, $F(1, 47) = .74$, $p = .39$. and marginal on high-load trials, $F(1, 47) = 3.8$, $p = .057$, $\eta_p^2 = .074$.

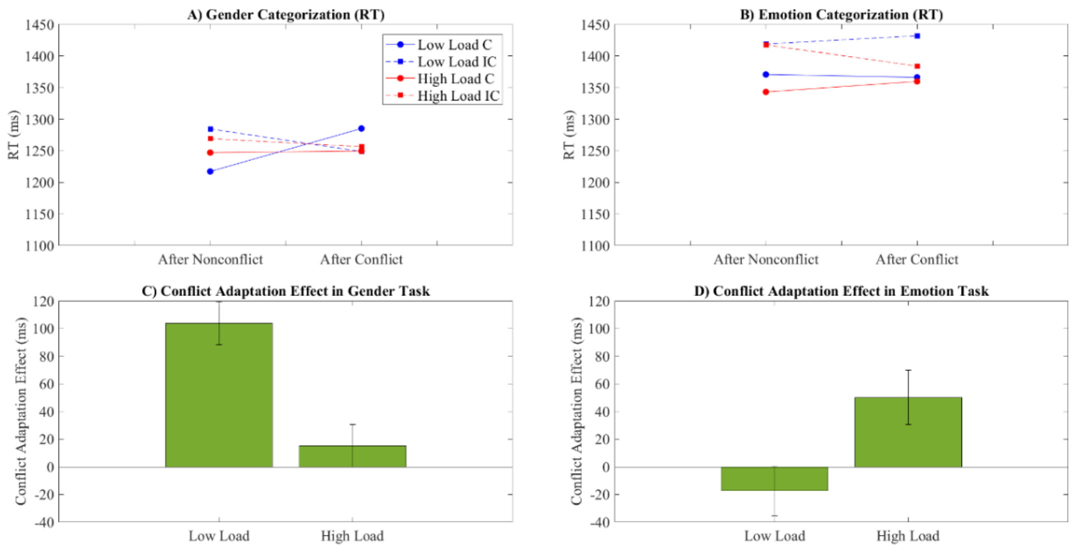

Figure 34. Mean RTs as a function of Task, WM load, N-1 Congruency, and N Congruency. A) Mean RTs in the gender Task. B) Mean RTs in the emotion task. C) Mean conflict adaptation effect in the gender task. D) Mean conflict adaptation effect in the emotion task.

AUC. Figure 35 indicates mean AUCs as a function of independent variables in Experiment 7. Main effect of working memory load (WM load) was significant, $F(1, 47) = 25.1$, $p < .0001$, $\eta_p^2 = .35$, because greater trajectory deviations were found on high-load trials (M = 237,432 px), compared to on low-load trials (M = 219,359 px). Main effect of N^{th} trial congruency also reached significance, $F(1, 47) = 7.05$, $p < .05$, $\eta_p^2 = .13$, with greater deviations found for incongruent trials (M = 236,070 px), than for congruent trials (M = 220,721 px).

In return, WM load interacted with $N\text{-}1^{th}$ trial congruency, $F(1, 47) = 5.45$, $p < .05$, $\eta_p^2 = .10$. Simple effect analyses at each WM load level suggested that the level of deviation did not vary as a function of N-1th trial congruency when WM load was low, $p = .74$, whereas the level of deviation on the Nth trial was attenuated after incongruent trials (after conflict) when WM load was high, $p < .05$. Note that all dependent variables were response indices recorded on Nth trials, which can be sorted by the congruency of the current (Nth) or the previous (N-1th) trials. The two-way interaction between task and Nth trial congruency was obtained, $F(1, 47) = 24.3$, $p < .0001$, $\eta_p^2 = .34$, because congruency effect was nonsignificant in the gender task, $p = .41$, but significant in the emotion task, $p < .0001$. A significant two-way interaction of Nth trial congruency and N-1th trial congruency confirmed the occurrence of the conflict adaptation effect, $F(1, 47) = 15.4$, $p < .0001$, $\eta_p^2 = .25$. Further analyses revealed that a 30,986-px congruency effect was obtained after nonconflict, $p < .001$, while a nonsignificant, −287-px congruency effect was found after conflict, $p = .96$, leading to a 31,273-px conflict adaptation effect.

The conflict adaptation effect, in return, was modulated by task as suggested by the three-way interaction of task, Nth trial congruency, and N-1th trial congruency, $F(1, 47) = 15$, $p < .0001$, $\eta_p^2 = .24$. As in the RT data, a conflict adaptation effect was obtained in the gender task, $F(1, 47) = 32.9$, $p < .0001$, $\eta_p^2 = .41$, but not in the emotion task, $F(1, 47) = .23$, $p = .64$. In the gender task, a significant 21,798-px congruency effect was obtained after nonconflict, and a significant, reversed congruency effect of −35,580 px was obtained after conflict.

A three-way interaction of WM load, task, and Nth trial congruency was obtained, $F(1, 47) = 12.9$, $p < .0001$, $\eta_p^2 = .22$. Separate two-way ANOVAs at each task level showed that the interaction of WM load and Nth trial congruency failed to reach significance in the gender task, $F(1, 47) = 2.14$, $p = .15$, but was significant in the emotion task, $F(1, 47) = 9.65$, $p < .01$, $\eta_p^2 = .17$. In the emotion task, congruency effects were obtained on both low-load trials (M = 53,503 px), $p < .0001$, and high-load trials (M = 21,679 px), $p < .01$, which contributed to the significant interaction effect.

Finally, the four-way interaction was significant, $F(1, 47) = 4.1$, $p < .05$, $\eta_p^2 = .08$. Two three-way ANOVAs at each task level revealed a significant three-way interaction of WM load, N^{th} trial congruency, and N-1^{th} trial congruency for the gender task, $F(1, 47) = 4.29$, $p < .05$, $\eta_p^2 = .083$, but not for the emotion task, $F(1, 47) = 1.17$, $p = .28$. Further two-way ANOVAs at each WM load level were performed on the gender task data, which confirmed significant conflict adaptation effects on both low-load trials, $F(1, 47) = 27.1$, $p < .0001$, $\eta_p^2 = .37$, and high-load trials, $F(1, 47) = 8.94$, $p < .01$, $\eta_p^2 = .16$. On low-load trials of the gender task, a 23,832-px congruency effect was obtained after nonconflict, $p < .05$, and a reversed −53,109-px congruency effect was obtained after conflict, $p < .0001$, resulting in a conflict adpatation effect of 76,941 px. Similarly on high-load trials, a marginally significant 19,764-px congruency effect was obtained after nonconflict, $p = .096$, and a nonsignificant −18,051-px congruency effect was obtained after conflict, $p = .18$, resulting in a 37,815-px conflict adpatation effect.

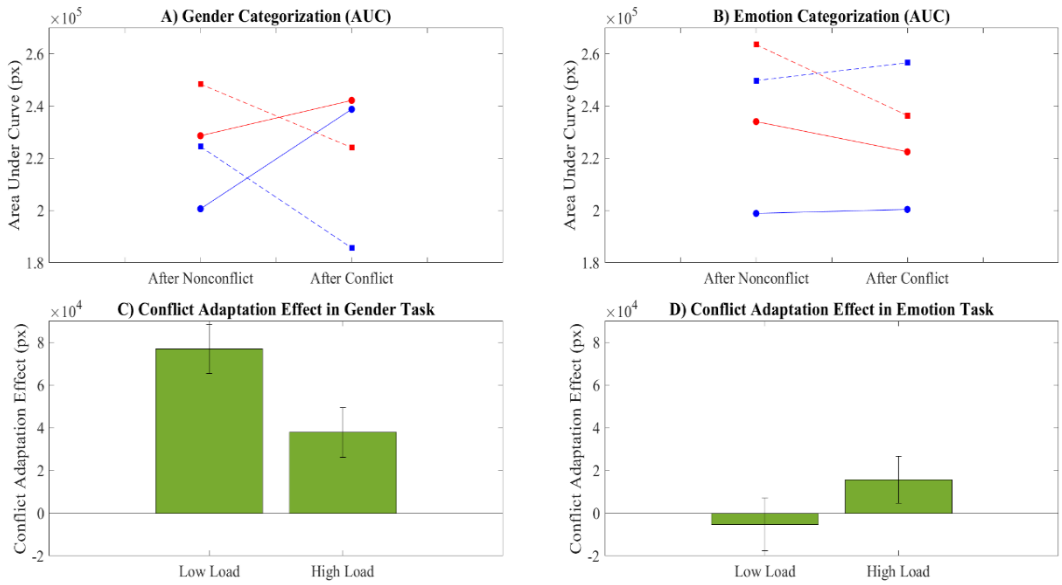

Figure 35. Mean AUCs as a function of Task, WM load, N-1 Congruency, and N Congruency. A) Mean AUCs in the gender Task. B) Mean AUCs in the emotion task. C) Mean conflict adaptation effect in the gender task. D) Mean conflict adaptation effect in the emotion task.

Conflict adaptation effect analysis. Figures 34C and 34D indicate RT conflict adaptation effects, and Figures 35C and 35D indicate AUC conflict adaptation effects. Two-way ANOVAs on the RT data yielded a main effect of task, $F(1, 47) = 4.81$, $p < .05$, $\eta_p^2 = .093$, because the conflict adaptation effect was greater in the gender task ($M = 60$ ms) than in the emotion task ($M = 16$ ms). Task also interacted with WM load, $F(1, 47) = 15.2$, $p < .0001$, $\eta_p^2 = .24$. Simple effect analyses at each task level revealed that the conflict adaptation effect was greater with a lower WM load in the gender task, $p < .001$, whereas it was greater with a higher load in the emotion task, $p < .05$.

Like the RT data, AUC data revealed a main effect of task, $F(1, 47) = 15$, $p < .0001$, $\eta_p^2 = .24$. Again, AUC conflict adaptation effect was greater in the gender task ($M = 57{,}378$ px) than in the emotion task ($M = 5168$ px). Task interacted with WM load,

$F(1, 47) = 4.1$, $p < .05$, $\eta_p^2 = .08$. Further analyses indicated that AUC conflict adaption effect was greater with a lower load in the gender task , $p < .05$, whereas the difference was nonsignificant in the emotion task, $p = .28$.

Figure 36 indicates conflict adaptation effects (CAE) calculated based on the velocity and trajectory angle at each time step. The shaded areas represent within-subject standard errors, and the gap between the areas larger than one standard error for each time step indicates significance at $p < .05$. The colored bars in the middle of each plot in these figures indicate the ranges of time steps in which the conflict adaptation effect was significant against zero levels by one-sample t-tests (Bonferroni corrected α level of 0.0005). Velocity data revealed conflict adaptation effects in the latter half phase (70~80th time steps marked in red) of the mouse trajectories for both gender and emotion tasks (Figures 36A & 36B) but on high load trials only. In contrast, trajectory angle data showed a conflict adaptation effect in the earlier half phase (46th time step marked in blue) only on the low load trials of the gender task.

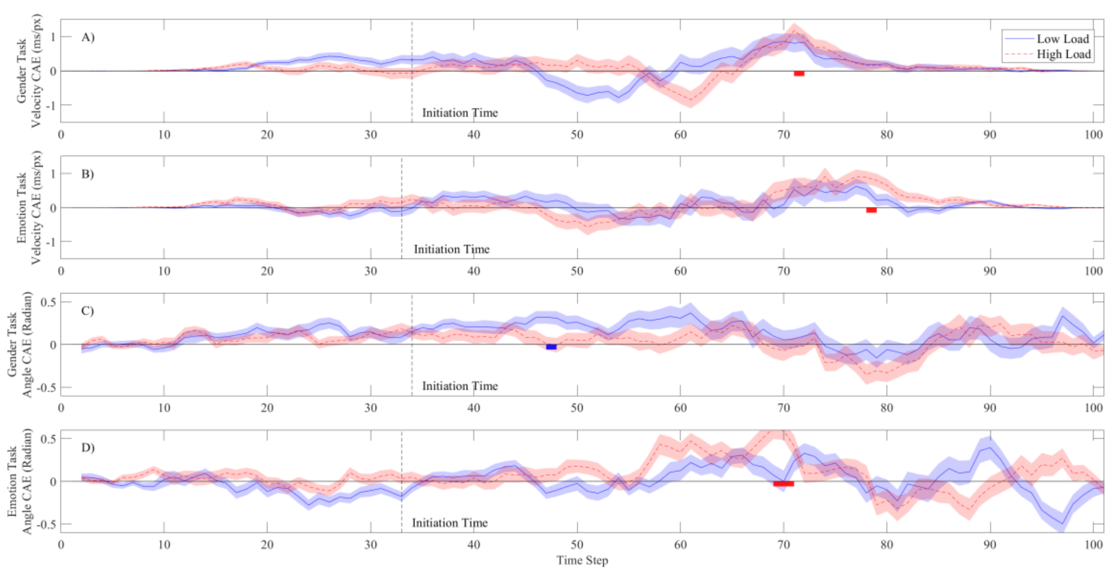

Figure 36. Mean velocity and trajectory angle conflict adaptation effects at each time step. A) Mean velocity effects at each time step in the gender Task. B) Mean velocity effects in the emotion task. C) Mean angle effects in the gender task. D) Mean angle effects in the emotion task.

Experiment 8 Result

Manipulation check. When performing the secondary memory tasks, participants showed a mean accuracy of 96.4% on low-load trials, and 88.8% on high-load trials. One-way ANOVA with WM load set as an independent variable confirmed that this difference was statistically significant, $F(1, 49) = 37.3$, $p < .0001$, $\eta_p^2 = .44$. Moreover, the mean Paas subjective rating was 3.36 (*SD* = 1.92; close to “low mental effort”) on low-load trials, whereas it was 5.2 (*SD* = 1.57; close to “neither low nor high mental effort”) on high-load trials (Figure 33). Again, the difference between these ratings was significant, $x^2(1) = 23.4$, $p < .0001$, $\eta^2 = .23$. Altogether, these results suggest that participants not only perceived the high-load trials to be more mentally taxing but also performed worse on them than on low-load trials. As in Experiment 7, the performance ratings suggest that the working memory manipulation used in Experiment 8 was successful in inducing different levels of working memory load.

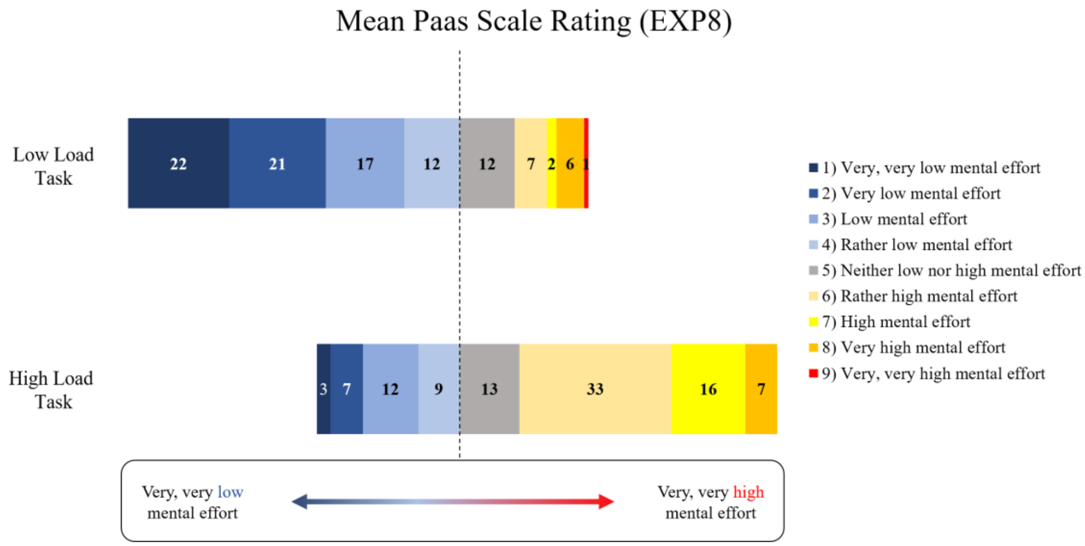

Figure 37. Stacked bar chart showing the relative frequency of subjective mental load ratings, with each bar divided according to the working memory load task.

RT. Figure 38 indicates mean RTs as a function of independent variables in Experiment 8. First, main effect of WM load was significant, $F(1, 49) = 5.15$, $p < .05$, $\eta_p^2 = .10$, because RT was longer on low-load trials (M = 1195 ms) than on high-load trials (M = 1174 ms). Main effect of Nth trial congruency also reached significance, $F(1, 49) = 4.27$, $p < .05$, $\eta_p^2 = .08$. RT was about 10 ms longer on incongruent trials (M = 1189 ms) than on congruent trials (M = 1179 ms), confirming the presence of the Stroop effect.

WM load interacted with N-1th trial congruency, $F(1, 49) = 4.14$, $p < .05$, $\eta_p^2 = .078$. Simple effect analyses at each WM load level indicated that mean RT was longer for high-load trials that followed a congruent trial (after nonconflict) than those that followed an incongruent trial (after conflict), $p < .01$. Contrarily, RTs on low-load trials did not vary as a function of N-1th trial congruency, $p = .82$. A two-way interaction of Nth trial congruency and N-1th trial congruency was obtained again, $F(1, 49) = 10.9$, $p < .01$, $\eta_p^2 = .18$, which suggested the presence of the conflict adaptation effect. Further analyses at each N-1th trial congruency level revealed a 24-ms congruency effect after nonconflict, p = .001, and a nonsignificant −5-ms congruency effect after conflict, p = .47, all summing up to the 29-ms conflict adaptation effect. No other main effects or interactions were significant.

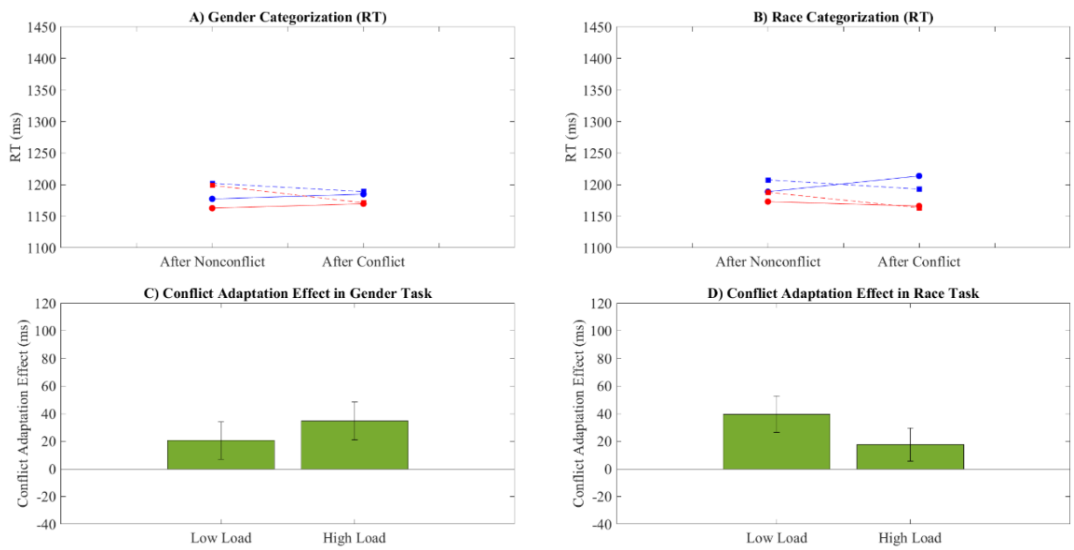

Figure 38. Mean RTs as a function of Task, WM load, N-1 Congruency, and N Congruency. A) Mean RTs in the gender Task. B) Mean RTs in the race task. C) Mean conflict adaptation effect in the gender task. D) Mean conflict adaptation effect in the race task.

AUC. Figure 39 indicates mean AUCs as a function of independent variables in Experiment 8. No main effect was observed from the AUC data. However, WM load interacted with task, $F(1, 49) = 4.76$, $p < .05$, $\eta_p^2 = .089$. A detailed inspection of the two-way interaction indicated that the high-load trials (M = 218,638 px) in the gender task yielded greater deviations compared to the low-load trials (M = 202,985 px), $p < .05$. However, such a difference was not observed in the race task data, p = .76. Task interacted marginally with Nth trial congruency, $F(1, 49) = 3.18$, $p = .08$, $\eta_p^2 = .061$, driven by a marginally significant congruency effect obtained in the gender task, p

= .063, and a nonsignificant congruency effect in the race task, p = .96. N^{th} trial congruency also showed a marginal interaction with N-1^{th} trial congruency, F(1, 49) = 3.13, p = .08, η_p^2 = .06, indicating the occurrence of a conflict adpatation effect. Simple effects analyses at each N-1^{th} trial congruency level showed that there was a 13,197-px congruency effect after nonconflict, p < .05, and a nonsignificant -1283-px congruency effect after conflict, which altogether accounted for the conflict adaptation effect of 14,480-ms.

Lastly, a four-way interaction of WM load, task, N^{th} trial congruency, and N-1^{th} trial congruency was significant, F(1, 49) = 6.37, p < . 05, η_p^2 = .12. Separate three-way ANOVAs at each task level revealed a marginal three-way interaction of WM load, N^{th} trial congruency and N-1^{th} trial congruency for the gender task, F(1, 49) = 2.93, p = .09, η_p^2 = .06, and a nonsignificant three-way interaction for the race task, F(1, 49) = 2, p = .16. The three-way interaction observed from the gender task data was further examined by performing two, two-way ANOVAs at each WM load level. The result indicated that the interaction of N^{th} trial congruency and N-1^{th} trial congruency was nonsignificant on low-load trials, F(1, 49) =.01, p = .94, but was significant on high-load trials, F(1, 49) = 4.05, p < .05, η_p^2 = .08. Specifically, on high-load trials of the gender task, a 35,092-px congruency effect was obtained after nonconflict, and a nonsignificant 1989-px congruency effect was obtained after conflict, resulting in a 37,081-px conflict adaptation effect.

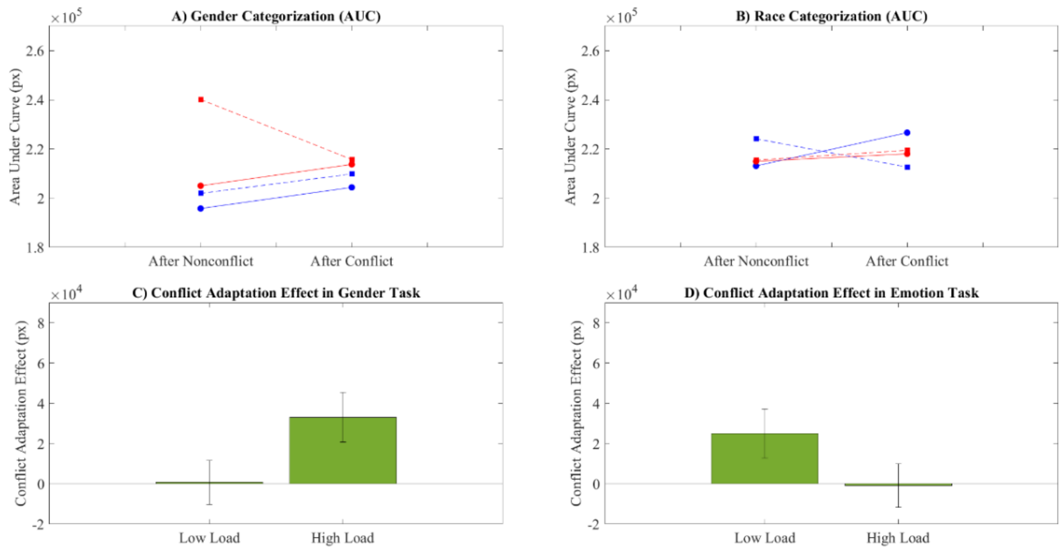

Figure 39. Mean AUCs as a function of Task, WM load, N-1 Congruency, and N Congruency. A) Mean AUCs in the gender Task. B) Mean AUCs in the race task. C) Mean conflict adaptation effect in the gender task. D) Mean conflict adaptation effect in the race task.

Conflict adaptation effect analysis. Figures 38C and 38D indicate RT conflict adaptation effects, and Figures 39C and 39D indicate AUC conflict adaptation effects. Figure 40 indicates velocity and angle conflict adaptation effects at each time step. Two-way repeated-measures ANOVAs were performed on the RT and AUC conflict adaptation effects with WM load and task set as within-subject variables. When the RT effect was examined, no main effect or interaction was obtained, *Fs* < 1.45, *ps* > .23. However, the AUC data revealed a significant interaction of WM load and task, $F(1, 49) = 6.37$, $p < .05$, $\eta_p^2 = .12$. Further analyses suggested that the conflict adaptation effect increased marginally with a higher WM load in the gender task, $p = .093$, whereas it decreased with a higher load in the race task, although the decrease was statistically nonsignificant, $p = .16$.

The examinations of velocities and trajectory angles at each time step confirmed these findings. A significant conflict adaptation effect was found only on high load trials of the gender task (68~74th steps in Figure 40A marked in red). The conflict adaptation effect failed to reach significance on other types of trials, in both RT and AUC data.

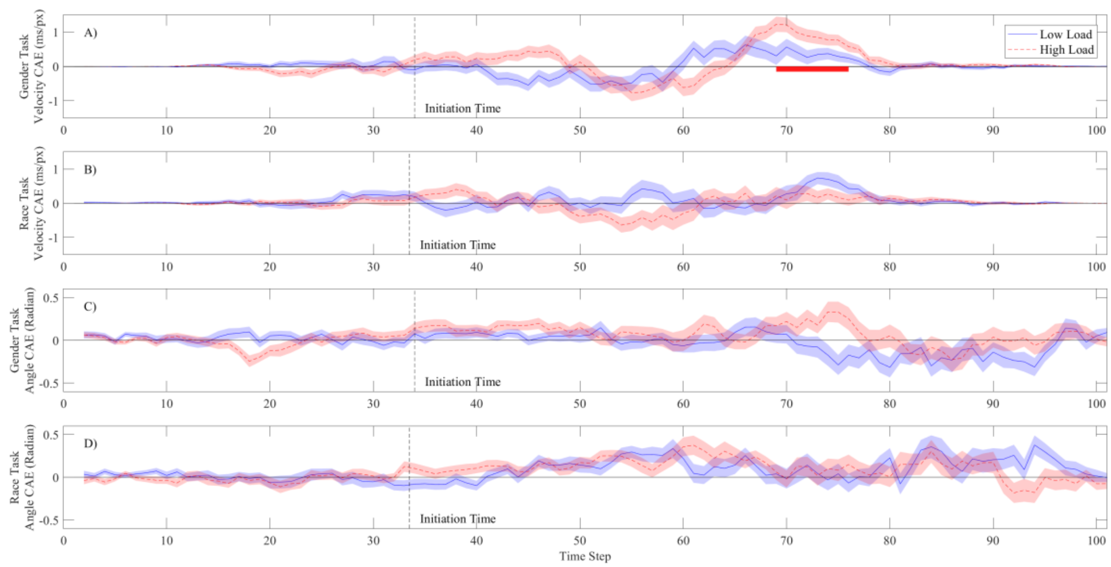

Figure 40. Mean velocity and trajectory angle conflict adaptation effects at each time step. A) Mean velocity effects at each time step in the gender Task. B) Mean velocity effects in the emotion task. C) Mean angle effects in the gender task. D) Mean angle effects in the emotion task.

Discussion

Congruency effects were present in most types of data except in Experiment 8's AUC data, which suggested that implicit associations were capable of eliciting the Stroop-like effects. The interference by the task-irrelevant features was observed and was moderated depending on which two of the gender, emotional expression and race were used as task-relevant and -irrelevant features. Specifically, RT congruency effect was larger in the emotion task of Experiment 7, whereas AUC congruency effect was larger in the gender task of Experiment 8. These results seem to suggest that gender features were more salient than emotional expression features but less salient than racial features, which might be one reason behind the asymmetric interference effects. Taken together, the congruency effects observed in the current Experiments seem to be in accordance with most of the diagnostics suggested by Algom et al. (2004), hence is a phenomenon similar to the Stroop effect. However, it should be noted that no systematic manipulation was made to the task-irrelevant features (i.e. size, salience of features), which made it difficult to test if congruency effects varied as a function of changes in task-irrelevant features. Future study is required to fully examine if the congruency effect obtained in the current Experiments is comparable to the Stroop effect.

In Experiment 7, all of the Stroop-like effect, conflict adaptation effect, and working memory modulation of the conflict adaptation effect were present but only in the gender task. In the gender task, the Stroop-like effect was present, although it was smaller in magnitude compared to the emotion task. The conflict adaptation effect was also present and was reduced to a nonsignificant effect under a higher load (Figures 34C, 35C, & 36C) as predicted by the cognitive control theories. In contrast, unexpected results were revealed in the emotion task. As in the gender task, the Stroop-like effect was significant. However, the Stroop-like effect was greater on low-load trials (M = 53,503 px) than on high-load trials (M = 21,679 px), which is a result inconsistent with the previous literature. Furthermore, the conflict adaptation effect was presented and moderated by working memory load such that the effect was significant only under a higher load (Figures 34D, 35D, 36B, & 36D), which contradicted the prediction by the cognitive control theories.

In Experiment 8, some signatures of the cognitive control mechanisms were found but in unexpected directions. Overall Stroop-like effect and conflict adaptation effect were found in the RT data but were not moderated by working memory load or task. AUC data provided insights on why this result was obtained. The AUC data suggested that the Stroop-like effect was obtained only in the gender task (Figures 39A & 3C). Moreover, the conflict adaptation effect was also found only on high-load trials of the gender task. The examination of the velocity conflict adaptation effects within a trial window confirmed these findings, as the effect was significant only in the latter half phase of the gender task's high-load trials (Figure 40A). Therefore, these results suggest the possibility that the significant Stroop-like and conflict adaptation effects obtained in Experiment 8 might not reflect the cognitive control mechanisms.

General Discussion

Asymmetries Between Identity Dimensions

Throughout Experiments 7 and 8, mixed results were found, which highlight the difference between the implicit association effect and the Stroop effect. Only the gender task in Experiment 7 revealed all three signatures of the cognitive control mechanisms. Interestingly in Experiment 8, the Stroop effect was obtained regardless of the task and the conflict adaptation effect was obtained only on the high-load trials of the gender task.

The confounded signal hypothesis (CSH; Becker et al., 2007) suggests that emotion recognition systems might take advantage of the older gender recognition systems in order to communicate threats or opportunities more efficiently. As a result, the emotion recognition systems can often be overridden by gender recognition systems

(Becker, 2017). Indeed, in Experiment 7's gender task, the modulation of the conflict adaptation effect by working memory load provided evidence that the newer emotion recognition systems could be sidelined by the older gender recognition systems. Consistent with the CSH, the interference by the task-irrelevant emotional expression features was diminished after it conflicted with the gender features, leading to a conflict adaptation effect. In addition, the cognitive control performance was hampered under a higher load, suggesting that the cognitive control mechanisms might be the mechanisms responsible for coordinating attentional operations during face categorization processes. Along with the argument that working memory maintains task representations (Braver et al., 1997; Kane & Engle, 2003), these findings explain why the inhibition of the task-irrelevant dimensions was weakened under higher load. On the other hand, the Stroop effect was consistently obtained in Experiment 7's emotion task, regardless of N-1th trial congruency. This result suggests that the task representation tuned to the emotional expression dimension was not successful in suppressing the task-irrelevant gender dimension, which replicates the finding by Becker (2017). However, it seems unlikely that the conflict elicited by incongruent faces in the emotion task was recognized as the conflict to be resolved by the cognitive control mechanisms because such processes should have produced a conflict adaptation effect only on the low-load trials. Instead, the Stroop effect observed in the emotion task might reflect the implicit association effect in its default form, which is only capable of introducing a delay or a distraction.

Furthermore, the CHS provides one possible explanation of why cognitive control signatures were not found in Experiment 8. The absence of cognitive control may reflect the failure to maintain the top-down task representation if the race categorization was prioritized similarly to the gender categorization. That is, the task-irrelevant facial dimension might have led to the inhibition of attentional allocation to the task-relevant dimension when the task-irrelevant dimension had an equally high value as the relevant dimension. Consistent with this possibility, the presence of highly valuable but task-irrelevant stimuli, like those implicitly associated with monetary rewards, can slow down visual search for a target (Anderson, Laurent, & Yantis, 2011; Anderson & Yantis, 2013). Therefore, it is also plausible that a certain social category can have long-term effects on the categorization of other categories if it has been frequently associated with important values (e.g., coalitional cue), especially if the associations were learned throughout an individual's lifetime.

However, this possibility only holds if the priorities for the gender and race categories are comparable. Against the literature which points gender, race and age as the three superordinate social categories (e.g. Fiske, 1998; Fiske & Neuberg, 1990; Stangor et al., 1992), developmental psychology research has suggested that the gender of an interaction partner starts to guide the social preference of children earlier than the race of the partner (Aboud, 2003; LaFreniere, Strayer, Gauthier, 1984; Ruble, Martin, & Berenbaum, 2006), despite their ability to attend to multiple social categories concurrently (Kinzler, Shutts, & Correll, 2010). In a similar vein, an evolutionary perspective suggests that the race categorization system might have been of less use, thus was not developed until mankind was able to travel to locations distant enough to interact with the groups of other races (Cosmides et al., 2003; Kurzban et al., 2001). Supporting this idea, people tend to categorize others by their race, only when the race is the coalitional cue required to distinguish in-group members from out-group members. When there is an alternative coalitional cue that is salient enough (e.g., colors of the t-shirts targets wore), the race categorization can quickly be overridden by the alternative-cue categorization (Kurzban et al., 2001). On the other hand, Hastie and Park (1986; also see Ito, & Urland, 2003) argued that the inconsistent findings surrounding the relative priority between the three superordinate categories might be arising from the fact that different studies tapped into different stages of face categorization. Accordingly, further studies are required to examine if the current findings can be generalized to other categorization processes and social categories.

Increase of Conflict Adaptation Effect with Working Memory Load

Unexpected results were also obtained in the current experiments, which were the presence of the conflict adaptation effects exclusive to high-load trials (e.g. emotion task in Experiment 7 RT data, gender task in Experiment 8 AUC data). These effects could have been random effects, which is plausible because these effects turned out to be only marginal. Alternatively, it might also be possible that there was just a limited involvement of the cognitive control mechanisms. This possibility suggests that cognitive control mechanisms might have been in action but were unable to attenuate congruency effects on some occasions, because of the environments limiting their capacity. The most limited environments for cognitive control mechanisms to operate, thus yield the greatest congruency effect, would be where no conflict was detected in the previous trial, where working memory load was high (Lavie, 2010; Lavie et al., 2004; Soutschek et al., 2013) and where task-irrelevant features had substantial potential to capture attention compared to task-relevant features (Becker, 2017). Therefore, in Experiments 7 and 8, the greatest congruency effect was anticipated on trials that followed a congruent trial and when participants were performing the emotion task (in Experiment 7; gender task in Experiment 8) under higher load. Further simple effect analyses were performed and reported here to validate this possibility (Table 1; also refer to Appendices H & I). Indeed, the results confirmed that congruency effects were the greatest on trials following a congruent trial (nonconflict), under high-load in Experiment 7's emotion task, and Experiment 8's gender task (Table 1). In return, this sudden increase in the congruency effect might have contributed to the pattern that looked similar to the conflict adaptation effect on high-load trials. However, the conflict adaptation effect was not obtained on low-load trials, because on low-load trials, the cognitive control mechanisms would have been in a greater control of the interference effects both after nonconflict and conflict.

It is not clear at this point why the sudden increase in the Stroop effect was obtained on high-load-trials of Experiment 2's gender task. Ito and Urland (2003) previously demonstrated that the early perceptual processing of gender-specific facial features was preceded by the processing of race-specific features, regardless of the task-relevant dimensions, and even after the visual salience of faces across different races was equated. Their explanation for the finding was that the racial category may affect the earlier perceptual and attentional processing stages, whereas the gender category may affect the later social judgment. They also suggested a possibility that the racial dimension became more salient because participants in their study had less experience categorizing race than gender, all of which might be the reasons behind our finding as well.

Table 1.
Comparisons of Congruency Effects in the Emotion and Gender Tasks.

	Working Memory load	N-1th Trial Congruency	Congruency Effect	Standard Error	*p*	Lower limits 95% C.I.	Upper limits 95% C.I.
Exp 7 Emotion Task (RT)	High	Congruent	74.44 ms	19.99	$p < .001$	34.23	114.66
		Incongruent	24.22 ms	16.968	$p = .16$	-9.9132	58.357
	Low	Congruent	48.24 ms	14.119	$p < .01$	19.841	76.65
		Incongruent	65.65 ms	16.009	$p < .001$	33.452	97.865
Exp 8 Gender Task (AUC)	High	Congruent	35092 px	12500	$p < .01$	9973.3	60211
		Incongruent	1989.6 px	11956	$p = .87$	-22036	26016
	Low	Congruent	6140.8 px	8114.3	$p = .45$	-10166	22447
		Incongruent	5452.4 px	9925.2	$p = .59$	-14493	25398

Conclusion and Implication

The final chapter investigated if conflicts elicited by implicit associations and implicit biases are handled by cognitive control mechanisms that handle Stroop stimuli. To this end, a mouse-tracking paradigm was integrated into a Stroop task with implicit associations serving as the Stroop stimuli. Using this new paradigm, the signatures of cognitive control mechanisms, widely reported in classical Stroop task studies, were sought.

The signatures obtained in this chapter shadowed the signatures of cognitive control mechanisms. While the evidence is not entirely conclusive, the results suggest that implicit associations and the conflicts elicited by them were handled by the cognitive mechanisms that, at the least, overlap with cognitive control mechanisms. In return, these findings indicate that approaches other than the traditional bias reduction strategies may also be effective in reducing implicit biases. It still remains an open question, whether which aspect of the cognitive control mechanisms (e.g. conflict monitoring, deployment of attentional resources) is directly associated with the reduction of implicit bias. A more detailed pinpointing of the bias regulation mechanisms may help increase the effectiveness of the current bias reduction strategies.

CHAPTER 5

FURTHER DISCUSSIONS AND CONCLUSIONS

Significance of the Study and Application

So far, previous approaches to implicit bias have been one-dimensional in that only a single identity dimension of a target was tested at a time (Johnson et al., 2012). Moreover, the current measurements of implicit bias lacked ecological validity. As recent findings suggested (Marsh et al., 2005), the button press tasks such as the IAT are not sensitive enough to capture the interactive nature underlying the social behaviors. These limitations not only prevented researchers from understanding the implicit bias mechanisms in detail but also placed constraints on the bias reduction strategies. Another limitation of the previous approaches is that they have been pinpointing the executive function as the cognitive mechanism responsible for moderating implicit biases. This can be problematic in that the "executive function" is a broad, umbrella term referring to various cognitive functions such as attentional control, cognitive flexibility, and response inhibition (Logue, & Gould, 2014). The broadly defined bias regulation mechanisms might be one reason why the current bias reduction strategies are effortful and time-consuming. The current dissertation addressed these limitations and demonstrated that multiple identity dimensions of an individual interacted and affected the implicit biases that were targeted to the individual. Furthermore, the current dissertation suggested cognitive control mechanisms, a finer definition than the executive functions, as the processor of the implicit biases, which opens up new possibilities for bias reduction strategies. For example, researchers reported that several genres of computer games that were not designed with an aim to enhance the cognitive control capacity could also enhance cognitive control capacity (Dobrowolski et al., 2015). Accordingly, the findings reported here might allow a new approach that is context-appropriate, less time- and money-consuming, but enjoyable, which could, in turn, offer solutions that are more adaptive and effective in the environment where collaboration and harmony are more important than ever.

Table 2.
Summary of the Dissertation.

		Hypotheses, Findings, and Implications
Chapters 1 & 2	Hypotheses	Implicit bias effect will be moderated by behavioral contexts and social categories of targets if implicit bias is a phenomenon determined by various external factors other than the internal state of the exhibitor of bias.
	Findings & Implications	Chapters 1 and 2 demonstrated that the manifestation of implicit biases depended on the behavioral context as well as the distinctive identity created by the combinations of different social categories of an individual. This finding implies that a more sensitive measure of the implicit bias than a simple button-press task and an experimental design factoring in the behavioral contexts are required to assess how implicit biases affect human behavior.
Chapter 3	Hypotheses	The speed and deviation of hand motion trajectories will vary as a function of working memory load if hand motions reflect an individual's capacity to suppress implicit bias.
	Findings & Implications	Chapter 3 identified several hand-motion indices (e.g. Mean velocity, velocity peak onset) of working memory load that could be used to infer the capacity of an individual to suppress the influence of implicit bias. Moreover, indices reflecting hand motions in an early phase (e.g. Initiation time, Q1 velocity) were identified as strong predictors of working memory load or the capacity to suppress implicit bias.

Chapter 4	Hypotheses	A Stroop-like effect, a conflict adaptation effect, and a working memory load modulation of the CAE will be observed in response to counter-stereotypical implicit associations, if implicit biases are handled by mechanisms overlapping with cognitive control mechanisms.
	Findings & Implications	The implicit associations produced various signatures indicating the cognitive control involvement, including the conflict adaptation effect, just like the Stroop targets, which suggested that implicit associations and Stroop stimuli are handled by overlapping cognitive mechanisms. The findings may allow alternative approaches for reducing biases that are more adaptive and effective.

Limitations

However, there were also several limitations that should be acknowledged. First, participant samples recruited for the current dissertation were comprised of participants who reported themselves as not being of African American background. This raises a concern that the validity of the experiments was jeopardized, especially for experiments in which the task was to categorize the race of target faces. In those experiments (Experiments 1, 2, & 8), only the Caucasian and African American faces were used as target faces, which might have elicited weaker implicit bias effects from participants who were not Caucasian American participants. That is, participants from another background, an Asian participant, for example, might have judged both the Caucasian and African American faces as outgroup faces, thus activated outgroup biases whenever the target faces were presented. The data also seems to support this possibility as the sizes of the congruency effects in Experiments 3, 4, and 7 tended to be numerically greater than those in Experiments 1, 2, and 8 (Appendices A, B, C, D, H & I). However, other implicit bias studies also used heterogenous participant samples (Payne, 2001, 2005;

Rudman & Ashmore, 2007), and still managed to obtain reliable and consistent results. Further research with a homogeneous participant sample and corresponding face stimulus set is required to investigate whether a semi-heterogeneous sample would produce results different from a homogeneous sample.

Another limitation of the current experiments (Experiments 7 and 8) is that they failed to replicate the results from Experiment 3 of the study by Soutschek et al. (2013), in which a similar working memory task was used. In their Experiments 1 and 2, an arithmetic task and an n-back task were used as working memory manipulations. In these experiments, the modulation of conflict adaptation effect by working memory load was observed just like the Experiments 7 and 8 of the current dissertation. However, in their third experiment, they used a memory task similar to the task used in Experiments 7 and 8 but failed to obtain the working memory modulation. Soutschek et al. suggested a couple of explanations for the absence of the working memory load modulation effect in their Experiment 3. According to them, the target probe task in their Experiment 3 only required participants to maintain up to six items in their working memory. However, when the working memory tasks were the arithmetic task or n-back task, participants had to update the contents in working memory, the operation which the authors suspected of sharing resources with the cognitive control functions.

However, there were three critical differences between the designs of the current Experiments and Experiment 3 in Soutschek et al. (2013). First, only fifteen participants participated in their study, thus Soutschek et al. might not have had sufficient statistical power to uncover the interaction between the working memory load and the conflict adaptation effect. Moreover, 75 percent of the trials in their Experiment 3 were congruent trials (25% incongruent trials), as opposed to 50 percent in the current experiments. The design with more frequent congruent trials than incongruent trials is often reported to show greater Stroop effects (Logan & Zbrodoff, 1979; Lowe & Mitterer, 1982; Mordkoff , 2012; West & Baylis, 1998), as more frequent exposure to congruent trials can motivate participants to strategically allocate more attentional resources to the task-irrelevant features of Stroop stimuli (Crump, Gong & Milliken, 2006). This is because when there are more congruent trials than the chance level, task-irrelevant features start to contain information on which response should be made, thus guide participants to attend to them intentionally. For example, when the word "BLUE" inked in blue color is presented more frequently than the word "BLUE" inked in red color, it is highly likely that a word inked in blue is a congruent target. Therefore, it seems likely that an additional factor, expectancy or participant strategy, had a confounding effect on the observed results, hence does not necessarily support the argument by Soutschek et al. that the maintenance demands in working memory do not interfere with the cognitive control mechanisms. Lastly, the task used in their study required participants to type the specific numbers that they memorized, meaning that they had to type only one number on low-load trials, and six numbers on high-load trials. This design provides room for other cognitive operations during later response selection stages, which limits the generalizability of the interpretations by Soutschek et al.

There are also skeptical views on whether the conflict adaptation effect reflects cognitive control mechanisms. The theories on the conflict adaptation effect gained popularity after the seminal work by Gratton et al. (1992), but other researchers also suggested that the conflict adaptation effect might be an artifact of the mechanisms independent of the cognitive control mechanisms (Hommel et al., 2004; Mayr, Awh & Laurey, 2003; Schmidt, 2013). According to the repetition priming account (Mayr et al., 2003), responses to congruent trials after nonconflict (CC), and incongruent trials after conflict (II) are more likely to be facilitated compared to congruent trials after conflict (IC), and incongruent trials after nonconflict (CI). The reason behind this argument was that, in a typical conflict task like the Stroop task, 50% of the CC and II trials involve a stimulus that is presented on both N-1th and Nth trials, whereas none of the IC and CI trials involve such a stimulus repetition (Figure 41A). Because of the repetition of the same stimulus, CC and II trials benefit from a priming effect, thus imitate what it seems like a conflict adaptation effect. Mayr et al. (2003) demonstrated this by showing that the conflict adaptation effect was eliminated when they excluded trial sequences in which an identical stimulus was presented consecutively.

Similarly, the feature-binding account provides an alternative possibility, which argues that CC and II trials are facilitated for a different reason. According to this account, when a stimulus is presented, which requires a response, the stimulus and the response are integrated into an event-file in episodic memory (Kahneman, Treisman & Gibbs, 1992). This event-file, in return, facilitates the same response to the same stimulus next time it is presented (e.g. response to the target A with the response option “A”). However, when the stimulus-response mapping should be changed, for example, by an instruction (e.g. respond to the target A with the response option “B”), the response to the stimulus is delayed because it does not match the existing event-file (Target A-Response A). Now, consider a Stroop task with the instruction to report the color of a target word by pressing a specific key (Figure 41B). Although the color-response key mapping does not change, the word-response key mapping changes constantly because the word is a task-irrelevant dimension. In this design, 50% of the CC and II trials would be comprised of two consecutive trials that have the same word-response key mapping, thus facilitate responses to those trials, whereas CI and IC trials would involve no such trials, thus require an existing event-file to be updated and delay responses.

Lastly, the contingency learning account posits that the conflict adaptation effect reflects the strategic allocation of attentional resources, rather than reflecting attentional inhibition or cognitive control (Schmidt, 2013). As mentioned earlier, a higher proportion of congruent trials guide participants to allocate more attentional resources to task-irrelevant features as they become more informative about the desired responses. The contingency learning camp has been taking advantage of this finding and tested the conflict adaptation effect using experimental designs with different ratios of congruent-incongruent trials. Consider an experiment in which congruent trials outnumber incongruent trials. This design would facilitate responses on congruent trials, increasing the RT gap between the congruent and incongruent trials in the process. In contrast, with a higher proportion of incongruent trials, the task-irrelevant feature would now contain information that would benefit the recognition of incongruent trials. As a result, the RT gap between the two types of trials would be reduced. This decrease and increase in the

RT gaps were what Schmidt (2013) claimed as the reason behind the pattern of results, which has been reported as conflict adaption effects.

Whether the conflict adaptation effect is an artifact or not has important implications for the current dissertation. Specifically, the skeptical views of conflict adaptation effect would suggest that the effects observed in the current dissertation are the artifacts of episodic memory or the strategic allocation of attention to task-irrelevant features, which are both independent of cognitive control mechanisms. That is, these artifacts are independent of a task goal or task representation based on which cognitive control mechanisms operate. Rather, they seem to reflect the processes in which participants adapt to the design of experimental tasks in pursuit of optimizing task performance. However, measures were taken in the experiments reported here to address the concerns above. First, a large stimulus set was used in Experiments 7 and 8 in the current dissertation (2 genders * 2 races/emotions * 8 faces = 32 faces) which is larger than most of the traditional conflict tasks. Moreover, the orders of the trials were coordinated such that no identical target face was presented in consecutive trials. The larger stimulus set should have minimized the influence of the stimulus repetition effect. Moreover, the locations of the response boxes in the current experiments changed randomly on a trial-by-trial basis, unlike traditional conflict tasks in which response keys are fixed throughout a block or an entire experiment. This design should have prevented event-files from facilitating CC and II trials, and instead introduced the same levels of delay for all types of trials. Lastly, the current design prevented the strategic allocation of attention to task-irrelevant features by maintaining equal numbers of congruent and incongruent trials. Conflict adaptation effects were obtained even after controlling for most of the artifacts. Therefore, it seems that the effects observed in the current experiments reflect the cognitive control mechanisms, except those obtained under higher load.

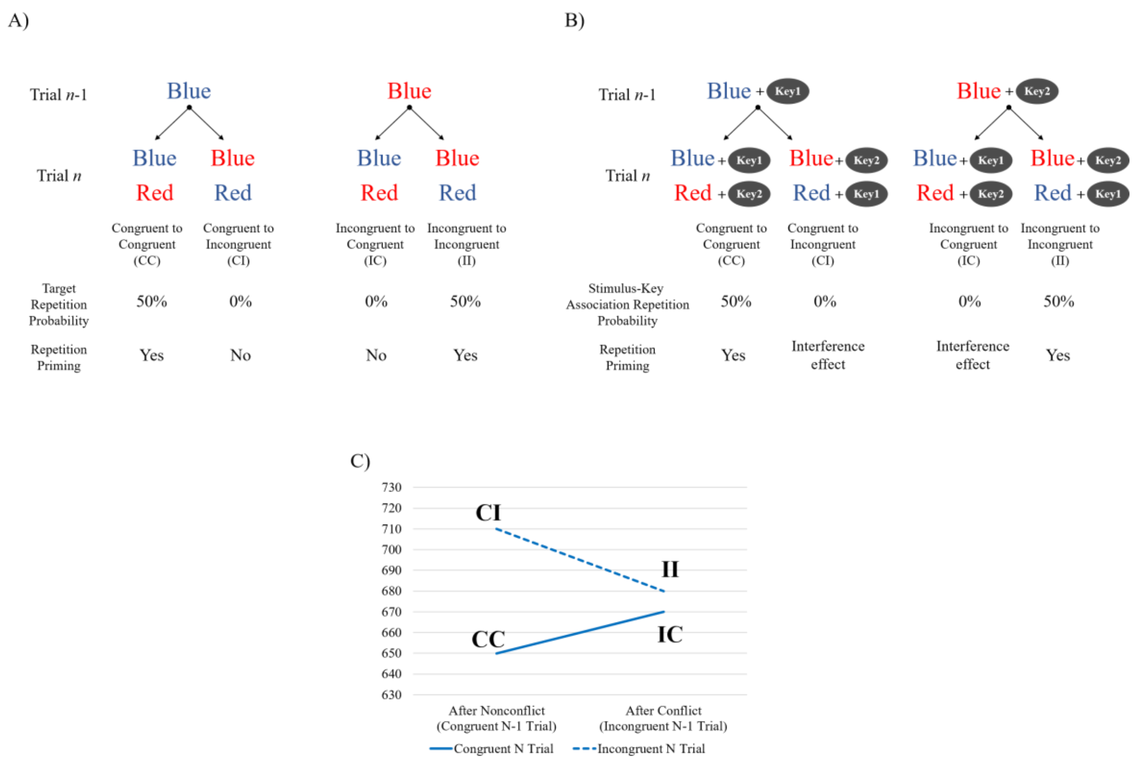

Figure 41. Anatomy of the conflict adaptation effect explained by the alternative accounts. A) Repetition priming account's explanation of the conflict adaptation effect. B) Feature-binding account's explanation. This hypothetical example assumes that participants responded to a target word inked in blue color with the key 1, and to a target word inked in red color with the key 2. C) Illustration of the conflict adaptation effect.

Possible Extensions of the Study

Previously, biased behavior has been conceptualized as a mixed consequence of the automatically triggered implicit bias, and the explicit control of the implicit bias (Devine, 1989; Fazio, 1990; Greenwald & Banaji, 1995; Jacoby, 1991; Payne, 2005). It was also suggested that the executive functions serve the role of explicitly controlling the implicit bias (Payne, 2005; Payne et al., 2005). Based on this conceptualization, researchers have suggested several strategies that focused on developing a rival response that will offset the implicit bias. Devine and her colleagues (2012) suggested five strategies for reducing biased behaviors induced by implicit biases, which included stereotype replacement, counter-stereotypic imaging, individuation, perspective-taking, and increasing opportunities for contact. Stereotype replacement involves recognizing a biased response as biased and replacing it with an unbiased response. Counter-stereotypic imaging involves imagining a counter-stereotypic figure that lives the life contradicting the stereotypes about his or her group (e.g. a female miner). Individuation involves learning specific information about a person from an out-group. Perspective-taking involves empathizing with an out-group member by taking the first-person perspective of that member. Lastly, increasing opportunities for contact requires participants to seek actively for opportunities to meet an out-group member. Devine her colleagues (2012) also tested these strategies to examine if their effects were sustained in the long term (for 8 weeks) and found a significant reduction in the level of the implicit bias during, and after the study.

One practical issue to consider is, to what extent is an individual willing to voluntarily sign up and go through a series of bias reduction training that is sometimes effortful and time-consuming? The effectiveness of the bias reduction strategies can be affected by an individual's attitude toward diversity in workplace (Ellis, 1994; Paluck, & Green, 2009), and motivation to reduce the prejudice (Devine & Monteith, 1993; Plant &

Devine 2009), which can be a luxury to some individuals strongly motivated to succeed in work. Therefore, whether the bias reduction training can be implemented successfully and naturally depends on the extent to which the training method can maintain participants' engagement and the ability to execute the training without causing inconvenience during the participants' daily life.

Although the strategies based on the executive function model were found effective and long-lasting (Devine et al., 2012), additional questions emerge in the current dissertation. Specifically, it was revealed that implicit biases might be handled by cognitive control mechanisms and that working memory load can facilitate or interfere with cognitive control functions. These findings helped narrow down the range of the bias regulation mechanisms set by the previous study, thus might contribute to the development of a new bias reduction strategy that is less effortful, time-consuming, and still long-lasting. However, there are prerequisites to the development of an alternative strategy. That is, should one try to enhance working memory capacity, cognitive control functions, or both in order to reduce the influences of implicit biases? This question also connects to the question of which specific training tasks should be applied to accomplish the goal of reducing implicit biases. These research questions could be explored by conducting a longitudinal study with a pretest-posttest design. For example, participants could be invited to perform various training games over a predetermined period. Experiments 1 through 8 could serve as the pretest-posttest measures of implicit biases that will help evaluate the progress of the participants in detail. By using these evaluation tasks, one will be able to find out which specific training task is capable of reducing biased behaviors, in a given behavioral context and toward a target of specific social categories.

In Experiments 7 and 8, it was reported that gender, assumed as a superordinate social category, attenuated the interference effect by the subordinate emotional expression through the conflict adaptation effect. However, the attenuation of the interference effect was absent when gender and race, the two social categories of comparable priorities, were paired together. While this finding was not unexpected, there is also a possibility that this finding does not reflect the processes in which cognitive control mechanisms assess the priorities of the perceived social categories and implement the control based on the assessment. For example, the attenuation of the race interference effect by the gender recognition systems should have been present according to the perspectives that the gender recognition systems develop earlier in life compared to the race recognition systems (e.g. Cosmides et al., 2003; Kurzban et al., 2001; LaFreniere, Strayer, Gauthier, 1984). On the other hand, there also exists a view that the racial features of a face are perceived earlier than the gender features (e.g. Ito & Urland, 2003), which allows for a prediction that the gender interference effect will be attenuated by the race recognition systems. One reason behind the finding could be that the current study only used black and white faces. That is, the race category might have been subordinate to the gender category but gained more attentional weight when black faces were contrasted to white faces, which made the racial features more salient cues for face categorization. Therefore, one potential extension would be to use faces of lower contrast (e.g. Asian and Caucasian faces) to see if the findings from the current dissertation generalize when faces of races other than black and white faces are contrasted.

Another question worth exploring is, will the findings from the data obtained from a lab setting generalize to real-world situations? Imagine a scenario in which a doctor is taking out an organ from a donor and transplanting it to his patient, both of whom may be from the doctor's ingroup or outgroup. Will the efficiency of the doctor's operation be affected by these contexts? Zipp and Craig (2019) addressed this question in their study, where participants engaged in a virtual training game that required them to triage virtual patients. Their findings were that participants took a longer time to initiate the triage procedure and made more errors while triaging dark-skinned agents compared to when triaging light-skinned agents. These results are in accordance with the results from the current dissertation if one interprets the triaging behavior as the behavior of providing a pleasant object. However, would the participants' tendency of slower and error-prone responses be eliminated if the operation involved taking something pleasant or good from the virtual patients, as suggested by the results from the current dissertation? The use of the virtual training environment in addressing this question may lead to the development of strategies to minimize error and increase the efficiency of the operations (behaviors) of interest, some of which are being carried out in life-or-death situations.

Lastly, the three accounts questioning the conflict adaptation effect (i.e. The repetition priming account, the feature-binding account, and the contingency learning account) were introduced to consider the possibility that cognitive control mechanisms might not handle implicit associations or biases. However, these accounts rather suggest an opportunity for the current dissertation than discrediting the arguments made. In conventional conflict tasks, only a few simple-shaped stimuli are presented repeatedly to minimize confounding effects. As the repetition priming account has suggested, this design can be vulnerable to the confounding repetition effect caused when two identical stimuli are presented consecutively. When researchers addressed this problem by increasing the stimulus set, the issue of contingency learning arose (Duthoo et al., 2014). For example, using the words “yellow” and “green” as target words in addition to the “blue” and “red” would decrease the repetition priming effect. However, at the same time, it becomes more difficult to maintain equal numbers of congruent and incongruent trials for all target words, thereby leaving room for contingency learning. Contrarily, the experimental design applied in this dissertation bypassed these complications and confounding effects by using multiple faces as target stimuli to prevent stimulus repetition, while maintaining equal numbers of congruent and incongruent trials. Moreover, the locations of response boxes changed randomly on each trial, which controlled for the confounding effect induced by the event-file priming. Therefore, it might be possible to use this design to investigate whether the conflict adaptation effect reflects the involvement of cognitive control mechanisms. In return, the new findings about the mechanisms would strengthen the claim that cognitive control mechanisms also serve the role of detecting and resolving conflicts elicited by implicit biases.

Summary and Conclusion

Throughout Experiments 1 to 8, a novel mouse-tracking approach was applied in order to address the limitations of the current implicit bias research and expand understanding of the cognitive mechanisms responsible for regulating implicit biases.

In Experiments 1 and 2, the implicit association test was integrated into the mouse-tracking paradigm to examine the behaviors of giving and taking objects to and from faces varying in gender and race. A typical implicit association effect was replicated in these experiments, but they also showed that the implicit association effects could be affected by target identities and behavioral contexts. Specifically, stereotypically congruent identity pairs, such as a black male or a white male face, elicited faster and more efficient responses. Moreover, the implicit association effect disappeared in the behavior of taking, which suggested that social categorization is a dynamic process affected by various factors.

Further efforts were made in Experiments 3 and 4 to see if these findings could be generalized if target faces varying in gender and emotional expression were used instead. Most of the results were replicated in Experiments 3 and 4. Implicit association effect was obtained again on giving trials, which was eliminated on taking trials. Stereotypically congruent pairs like angry male and happy female faces elicited faster and more efficient responses. In addition, participants suffered greater from the interference by the task-irrelevant distractor features when the task was to categorize the emotional expression of the faces varying in gender. These findings also implied that there might be a hierarchical relationship between the gender and race recognition systems, as suggested by Becker (2017) and the CSH (Becker et al., 2007).

Based on the prior evidence, which suggested that the implicit bias mechanisms might be linked to the cognitive control mechanisms, the influence of working memory, an essential function, and a resource for the cognitive control mechanisms, on hand motions was evaluated in Experiments 5 and 6. Another focus was placed on distinguishing the working memory load effect on hand motions from that of the motor task difficulty. In these Experiments, participants moved a cursor from a starting location to a target varying in size, while performing three different secondary tasks that were designed to impose different levels of working memory load. Although the trajectory deviation measure (AUC) was suspected to be sensitive to various factors such as the gain value of a computer mouse cursor, a working memory load effect was present consistently in an early phase of a hand movement, which was followed by the motor task difficulty effect. A machine learning technique was used with a feature selection procedure to identify the effective predictors of working memory load, which revealed that only 10 to 13 hand motion features out of 39 tested features were useful in predicting the level of working memory load. The five most frequently identified features were initiation time, mean velocity, maximum velocity, velocity peak onset, and mean velocity at the first quartile of a trial.

Finally, Experiments 7 and 8 investigated whether the three signatures hinting the involvement of cognitive control mechanisms were found in reaction to the conflicts induced by stereotypically incongruent implicit associations. Participants were asked to categorize the faces varying in either emotional expression or gender (Experiment 7) and varying in race or gender (Experiment 8). In all experiments, the main effects of N^{th} trial congruency was significant (except for Experiment 8's AUC data), which replicated the findings from studies that used the Stroop task (e.g. Stroop, 1935), and the variants of the implicit association task (e.g. Fazio et al., 1995; Greenwald et al., 1998). However, the other two signatures—the conflict adaptation effect and the modulation by working memory load—were observed only in the gender task when the emotional expression was task-irrelevant (Experiment 7 gender task), but not in other tasks when the gender was task-irrelevant or was pitted against the race category. Based on these findings, it was concluded that faces with stereotypically incongruent facial features could elicit effects similar to the Stroop effect. Nevertheless, unlike the Stroop stimuli, it also seemed that mere exposures to faces did not necessarily lead to the involvement of cognitive control mechanisms. Instead, it seemed that the attentional priority assigned to different social categories affected whether cognitive control was activated to address the response conflicts elicited by implicit associations.

Altogether, the findings from Experiments 1 to 8 demonstrated how implicit biases affect overt human behaviors depending on the distinct nuances created by the layers of target identities, and the behavioral contexts in which the implicit biases operate. The current dissertation extended the findings by pinpointing the cognitive control mechanisms as one of the major processors of implicit biases, and that these processors, too, are affected by the intersection of social categories.

REFERENCES

Aboud, F. E. (2003). The formation of in-group favoritism and out-group prejudice in young children: Are they distinct attitudes? *Developmental Psychology, 39*(1), 48-60.

Adams, R. B., Hess, U., & Kleck, R. E. (2015). The intersection of gender-related facial appearance and facial displays of emotion. *Emotion Review, 7*(1), 5-13.

Algom, D., Chajut, E., & Lev, S. (2004). A rational look at the emotional stroop phenomenon: a generic slowdown, not a stroop effect. *Journal of Experimental Psychology: General, 133*(3), 323-338.

Allport, G. W. (1954). *The nature of prejudice*. Reading, MA: Addison-Wesley.

Amodio, D. M. (2010). Coordinated roles of motivation and perception in the regulation of intergroup responses: Frontal cortical asymmetry effects on the P2 event-related potential and behavior. *Journal of Cognitive Neuroscience, 22*(11), 2609-2617.

Amodio, D. M. (2014). The neuroscience of prejudice and stereotyping. *Nature Reviews Neuroscience, 15*(10), 670-682.

Amodio, D. M., & Devine, P. G. (2006). Stereotyping and evaluation in implicit race bias: evidence for independent constructs and unique effects on behavior. *Journal of Personality and Social Psychology, 91*(4), 652-661.

Amodio, D. M., Harmon-Jones, E., Devine, P. G., Curtin, J. J., Hartley, S. L., & Covert, A. E. (2004). Neural signals for the detection of unintentional race bias. *Psychological Science, 15*(2), 88-93.

Anderson, B. A. (2015). Value-driven attentional capture is modulated by spatial context. *Visual Cognition, 23*(1-2), 67-81.

Anderson, B. A., Laurent, P. A., & Yantis, S. (2011). Value-driven attentional capture. *Proceedings of the National Academy of Sciences, 108*(25), 10367-10371.

Anderson, B. A., & Yantis, S. (2013). Persistence of value-driven attentional capture. *Journal of Experimental Psychology: Human Perception and Performance, 39*(1), 6-9.

Baddeley, A. D. (1986). *Working memory*. Oxford, England: Oxford University Press

Baddeley, A. D. (1992). Working memory. *Science, 255*(5044), 556-559.

Becker, D. V. (2017). Facial gender interferes with decisions about facial expressions of anger and happiness. *Journal of Experimental Psychology: General, 146*(4), 457-463.

Becker, D. V., Kenrick, D. T., Neuberg, S. L., Blackwell, K. C., & Smith, D. M. (2007). The confounded nature of angry men and happy women. *Journal of Personality and Social Psychology*, *92*(2), 179-190.

Becker, D. V., Mortensen, C. R., Ackerman, J. M., Shapiro, J. R., Anderson, U. S., Sasaki, T., ... & Kenrick, D. T. (2011). Signal detection on the battlefield: Priming self-protection vs. revenge-mindedness differentially modulates the detection of enemies and allies. *PloS One*, *6*(9), e23929.

Becker, D. V., Mortensen, C. R., Anderson, U. S., & Sasaki, T. (2014). Out of sight but not out of mind: memory scanning is attuned to threatening faces. *Evolutionary Psychology, 12*(5), 901-902.

Becker, D., Rheem, H., Pick, C. M., Ko, A., & Lafko, S. R. (2019). Angry faces hold attention: Evidence of attentional adhesion in two paradigms. In *Progress in Brain Research* (Vol. 247, pp. 89-110). Elsevier BV

Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, *42*(2), 155-162.

Billig, M. (1985). Prejudice, categorization and particularization: From a perceptual to a rhetorical approach. *European Journal of Social Psychology*, *15*(1), 79-103.

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends in Cognitive Sciences, 8*(12), 539-546.

Branscombe, N. R., & Smith, E. R. (1990). Gender and racial stereotypes in impression formation and social decision-making processes. *Sex Roles, 22*(9-10), 627-647.

Braver, T. S., Cohen, J. D., Nystrom, L. E., Jonides, J., Smith, E. E., & Noll, D. C. (1997). A parametric study of prefrontal cortex involvement in human working memory. *Neuroimage, 5*(1), 49-62.

Cacioppo, J. T., & Berntson, G. G. (1994). Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates. *Psychological Bulletin, 115*(3), 401-423.

Cacioppo, J. T., Priester, J. R., & Berntson, G. G. (1993). Rudimentary determinants of attitudes: II. Arm flexion and extension have differential effects on attitudes. *Journal of Personality and Social Psychology, 65*(1), 5-17.

Carrasco, M., Evert, D. L., Chang, I., & Katz, S. M. (1995). The eccentricity effect: Target eccentricity affects performance on conjunction searches. *Perception & Psychophysics, 57*(8), 1241-1261.

Cavalcanti, T., & Tavares, J. (2016). The Output Cost of Gender Discrimination: A Model-based Macroeconomics Estimate. *The Economic Journal, 126*(590), 109-134.

Cohn, L. (2016, November 10). The U.S. Made Zero Progress in Adding Women to Congress. *Fortune*, Retrieved from https://fortune.com/2016/11/10/election-results-women-in-congress/

Cosmides, L., Tooby, J., & Kurzban, R. (2003). Perceptions of race. *Trends in Cognitive Sciences*, 7(4), 173-179.

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology, 1*(1), 42-45.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin, 104*(2), 163-191.

Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educational Psychology Review, 26*(2), 197-223.

Crenshaw, K. (1993). Demarginalizing the intersection of race and sex: A Black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. In D. K. Weisbert (Ed.), *Feminist legal theory: Foundations* (pp. 383–395). Philadelphia: Temple University Press. (Original work published 1989)

Crump, M. J., Gong, Z., & Milliken, B. (2006). The context-specific proportion congruent Stroop effect: Location as a contextual cue. *Psychonomic Bulletin & Review*, *13*(2), 316-321.

Cumming, G., Fidler, F., & Vaux, D. L. (2007). Error bars in experimental biology. *The Journal of Cell Biology*, *177*(1), 7-11.

Dale, R., Kehoe, C., & Spivey, M. J. (2007). Graded motor responses in the time course of categorizing atypical exemplars. *Memory & Cognition*, *35*(1), 15-28.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56*(1), 5-18.

Devine, P. G., & Elliot, A. J. (1995). Are racial stereotypes really fading? The Princeton trilogy revisited. *Personality and Social Psychology Bulletin, 21*(11), 1139-1150.

Devine, P. G., Forscher, P. S., Austin, A. J., & Cox, W. T. (2012). Long-term reduction in implicit race bias: A prejudice habit-breaking intervention. *Journal of Experimental Social Psychology, 48*(6), 1267-1278.

Devine, P. G., & Monteith, M. J. (1993). The role of discrepancy-associated affect in prejudice reduction. In *Affect, Cognition and Stereotyping* (pp. 317-344), New York, NY: Academic Press.

Dobrowolski, P., Hanusz, K., Sobczyk, B., Skorko, M., & Wiatrow, A. (2015). Cognitive enhancement in video game players: The role of video game genre. *Computers in Human Behavior, 44*, 59-63.

Duthoo, W., Abrahamse, E. L., Braem, S., Boehler, C. N., & Notebaert, W. (2014). The congruency sequence effect 3.0: a critical test of conflict adaptation. *PloS One, 9*(10), e110462.

Ehrlich, H. J. (1973). *The social psychology of prejudice*. New York: Wiley.

Ekman, P., & Friesen, W. (1976). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.

Ellis, C. (1994). Diverse approaches to managing diversity. *Human Resource Management, 33*(1), 79-109.

Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. *Advances in Experimental Social Psychology, 23*, 75-109.

Fazio, R. H., Jackson, J. R., Dunton, B. C. & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: a bona fide pipeline? *Journal of Personality and Social Psychology, 69*, 1013-1027.

Findlater, L., Froehlich, J. E., Fattal, K., Wobbrock, J. O., & Dastyar, T. (2013, April). Age-related differences in performance with touchscreens compared to traditional mouse input. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 343-346). New York, NY, US: ACM.

Fiske, S. T. (1998). Stereotyping, prejudice, and discrimination. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology: Vols. 1 and 2* (4th ed. pp. 357–411). New York, NY: McGraw-Hill.

Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology, 82*(6), 878-902.

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 23, pp. 1–74), New York, NY: Academic Press.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology, 47*(6), 381-391.

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance, 18*(4), 1030-1044.

Frank, R. H. (1988). *Passions within reason: the strategic role of the emotions*. New York: Norton.

Freeman, J. B., & Ambady, N. (2009). Motions of the hand expose the partial and parallel activation of stereotypes. *Psychological Science, 20*(10), 1183-1188.

Freeman, J. B., Ambady, N., Rule, N. O., & Johnson, K. L. (2008). Will a category cue attract you? Motor output reveals dynamic competition across person construal. *Journal of Experimental Psychology: General, 137*(4), 673-690.

Freeman, J., Dale, R., & Farmer, T. (2011). Hand in motion reveals mind in motion. *Frontiers in Psychology, 2*, 59.

Freeman, J. B., Pauker, K., Apfelbaum, E. P., & Ambady, N. (2010). Continuous dynamics in the real-time perception of race. *Journal of Experimental Social Psychology, 46*(1), 179-185.

Frings, C., Englert, J., Wentura, D., & Bermeitinger, C. (2010). Decomposing the emotional Stroop effect. *The Quarterly Journal of Experimental Psychology, 63*(1), 42-49.

Garavan, H., Ross, T. J., Murphy, K., Roche, R. A. P., & Stein, E. A. (2002). Dissociable executive functions in the dynamic control of behavior: inhibition, error detection, and correction. *Neuroimage, 17*(4), 1820-1829.

Geggel, L. (2017, March 4). 100 Years of Women in Politics: How They've Served. *Live Science,* Retrieved from https://www.livescience.com/58118-100th-anniversary-of-women-in-congress.html

Glisky, E. L. (2007). Changes in cognitive function in human aging. In R. Riddle (Ed.), *Brain aging: Models, methods, and mechanisms* (pp. 3-20), Boca Raton, FL: CRC Press.

Goff, P. A., Thomas, M. A., & Jackson, M. C. (2008). "Ain't I a woman?": Towards an intersectional approach to person perception and group-based harms. *Sex Roles, 59*(5-6), 392-403.

Gratton, G., Coles, M. G., & Donchin, E. (1992). Optimizing the use of information: strategic control of activation of responses. *Journal of Experimental Psychology: General*, *121*(4), 480-506.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*(1), 4-27.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of Personality and Social Psychology, 74*(6), 1464-1489.

Grimes, M., & Valacich, J. (2015, December). *Mind Over Mouse: The Effect of Cognitive Load on Mouse Movement Behavior*. Paper presented at the *2015 International Conference on Information Systems: Exploring the Information Frontier*, Fort Worth, Texas. Retrieved from http://aisel.aisnet.org/icis2015/proceedings/ HCI/9

Hartman, M., Bolton, E., & Fehnel, S. E. (2001). Accounting for age differences on the Wisconsin Card Sorting Test: Decreased working memory, not inflexibility. *Psychology and Aging, 16*(3), 385-399.

Hastie, R., & Park, B. (1986). The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line. *Psychological Review*, *93*(3), 258-268.

Hedden, T., & Gabrieli, J. D. (2004). Insights into the ageing mind: a view from cognitive neuroscience. *Nature Reviews Neuroscience, 5*(2), 87-96.

Hehman, E., Stolier, R. M., & Freeman, J. B. (2015). Advanced mouse-tracking analytic techniques for enhancing psychological science. *Group Processes & Intergroup Relations, 18*(3), 384-401.

Hess, U., Sabourin, G., & Kleck, R. E. (2007). Postauricular and eyeblink startle responses to facial expressions. *Psychophysiology*, *44*(3), 431-435.

Ho, C., & Jackson, J. W. (2001). Attitude toward asian americans: Theory and measurement. *Journal of Applied Social Psychology*, *31*(8), 1553-1581.

Hommel, B., Proctor, R. W., & Vu, K. P. L. (2004). A feature-integration account of sequential effects in the Simon task. *Psychological Research*, *68*(1), 1-17.

Hsieh, C. T., Hurst, E., Jones, C. I., & Klenow, P. J. (2013). *The allocation of talent and us economic growth*. Unpublished manuscript, Stanford University, Palo Alto, CA.

Izard, C. E. (2013). *Patterns of emotions: A new analysis of anxiety and depression*. New York: Academic Press.

Ito, T. A., & Urland, G. R. (2003). Race and gender on the brain: electrocortical measures of attention to the race and gender of multiply categorizable individuals. *Journal of Personality and Social Psychology, 85*(4), 616-626.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*(5), 513-541.

Johnson, K. L., Freeman, J. B., & Pauker, K. (2012). Race is gendered: How covarying phenotypes and stereotypes bias sex categorization. *Journal of Personality and Social Psychology*, *102*(1), 116-131.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, *24*(2), 175-219.

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: the contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, *132*(1), 47-70.

Kieslich, P. J., & Henninger, F. (2017). Mousetrap: An integrated, open-source mouse-tracking package. *Behavior Research Methods, 49*(5), 1652-1667.

Kieslich, P. J., Schoemann, M., Grage, T., Hepp, J., & Scherbaum, S. (2019). Design factors in mouse-tracking: What makes a difference?. *Behavior Research Methods,* 1-25.

Kinzler, K. D., Shutts, K., & Correll, J. (2010). Priorities in social categories. *European Journal of Social Psychology, 40*(4), 581-592.

Kurzban, R., Tooby, J., & Cosmides, L. (2001). Can race be erased? Coalitional computation and social categorization. *Proceedings of the National Academy of Sciences, 98*(26), 15387-15392.

LaFreniere, P., Strayer, F., & Gauthier, R. (1984). The emergence of same-sex affiliative preferences among preschool peers: A developmental/ethological perspective. *Child Development*, *55*(5), 1958–1965.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). Motivated attention: Affect, activation, and action. In P. J. Lang & R. F. Simons (Eds.), *Attention and orienting: Sensory and motivational processes* (pp. 97–135). Mahwah, NJ: Erlbaum.

Lavie, N. (2010). Attention, distraction, and cognitive control under load. *Current Directions in Psychological Science, 19*(3), 143-148.

Lavie, N., Hirst, A., De Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General, 133*(3), 339-354.

Lazerus, T., Ingbretsen, Z. A., Stolier, R. M., Freeman, J. B., & Cikara, M. (2016). Positivity bias in judging ingroup members' emotional expressions. *Emotion, 16*(8), 1117-1125.

Logan, G. D., & Zbrodoff, N. J. (1979). When it helps to be misled: Facilitative effects of increasing the frequency of conflicting stimuli in a Stroop-like task. *Memory & Cognition*, *7*(3), 166-174.

Logue, S. F., & Gould, T. J. (2014). The neural and genetic basis of executive function: attention, cognitive flexibility, and response inhibition. *Pharmacology Biochemistry and Behavior, 123*, 45-54.

Lowe, D. G., & Mitterer, J. O. (1982). Selective and divided Attention in a Stroop task. *Canadian Journal of Psychology, 36*(4), 684-700.

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*(4), 1122-1135.

Markman, A. B., & Brendl, C. M. (2005). Constraining theories of embodied cognition. *Psychological Science, 16*(1), 6-10.

Marsh, A. A., Ambady, N., & Kleck, R. E. (2005). The effects of fear and anger facial expressions on approach-and avoidance-related behaviors. *Emotion, 5*(1), 119-124.

Martin, D., Swainson, R., Slessor, G., Hutchison, J., Marosi, D., & Cunningham, S. J. (2015). The simultaneous extraction of multiple social categories from unfamiliar faces. *Journal of Experimental Social Psychology, 60*, 51-58.

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods, 44*(2), 314-324.

Mattay, V. S., Fera, F., Tessitore, A., Hariri, A. R., Berman, K. F., Das, S., ... & Weinberger, D. R. (2006). Neurophysiological correlates of age-related changes in working memory capacity. *Neuroscience Letters, 392*(1-2), 32-37.

Mayr, U., Awh, E., & Laurey, P. (2003). Conflict adaptation effects in the absence of executive control. *Nature Neuroscience*, *6*(5), 450-452.

McCabe, D. P., Roediger III, H. L., McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: evidence for a common executive attention construct. *Neuropsychology*, *24*(2), 222-243.

McDonald, M. M., Navarrete, C. D., & Van Vugt, M. (2012). Evolution and the psychology of intergroup conflict: The male warrior hypothesis. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367*, 670 – 679.

Meyer, D. E., Abrams, R. A., Kornblum, S., Wright, C. E., & Keith Smith, J. E. (1988). Optimality in human motor performance: ideal control of rapid aimed movements. *Psychological Review, 95*(3), 340-370.

Meyer, D. E., Keith-Smith, J. E., Kornblum, S., Abrams, R. A., & Wright, C. E. (1990). Speed-accuracy tradeoffs in aimed movements: Toward a theory of rapid voluntary action. In M. Jeannerod (Ed.), *Attention and performance 13: Motor representation and control* (pp. 173-226). Hillsdale, NJ, US: Lawrence Erlbaum Associates Inc.

Miyake, A., & Shah, P. (1999). *Models of working memory: Mechanisms of active maintenance and executive control.* Cambridge, England: Cambridge University Press.

Mordkoff, J. T. (2012). Observation: Three reasons to avoid having half of the trials be congruent in a four-alternative forced-choice experiment on sequential modulation. *Psychonomic Bulletin & Review*, *19*(4), 750-757.

Navarrete, C. D., Olsson, A., Ho, A. K., Mendes, W. B., Thomsen, L., & Sidanius, J. (2009). Fear extinction to an out-group face: The role of target gender. *Psychological Science, 20*(2), 155-158.

Olson, M. A., & Fazio, R. H. (2004). Trait inferences as a function of automatically activated racial attitudes and motivation to control prejudiced reactions. *Basic and Applied Social Psychology, 26*(1), 1-11.

Paas, F. G. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology, 84*(4), 429-434

Paluck, E. L., & Green, D. P. (2009). Prejudice reduction: What works? A review and assessment of research and practice. *Annual Review of Psychology, 60*, 339-367.

Payne, B. 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, 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.

Payne, B. K., Jacoby, L. L., & Lambert, A. J. (2005). *Attitudes as accessibility bias: Dissociating automatic and controlled components*. In R. A. Hassin, J. S. Uleman, & J. A. Bargh, (Eds.),The New Unconscious. New York: Oxford University Press.

Plant, E. A., & Devine, P. G. (2009). The active control of prejudice: Unpacking the intentions guiding control efforts. *Journal of Personality and Social Psychology, 96*(3), 640-652.

Plutchik, R. (1980). *Emotion: A psychoevolutionary synthesis*. New York: Harper & Row.

Preble, E. (2017, March 5). Addressing the gender wage gap: Career center and Women's Center team up to promote awareness. *The Brown and White*, Retrieved from https://thebrownandwhite.com/2017/03/05/womens-pay-gap/

Richeson, J. A., & Shelton, J. N. (2003). When prejudice does not pay effects of interracial contact on executive function. *Psychological Science*, *14*(3), 287-290.

Rinck, M., & Becker, E. S. (2007). Approach and avoidance in fear of spiders. *Journal of Behavior Therapy and Experimental Psychiatry, 38*(2), 105-120.

Ruble, D. N., Martin, C. L., & Berenbaum, S. A. (2006). Gender development. In N. Eisenberg, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology: Vol. 3, Social, emotional, and personality development* (6th ed., pp. 858–932). Hoboken, NJ, US: John Wiley & Sons Inc.

Rudman, L. A., & Ashmore, R. D. (2007). Discrimination and the implicit association test. *Group Processes & Intergroup Relations, 10*(3), 359-372.

Schmidt, J. R. (2013). Questioning conflict adaptation: proportion congruent and Gratton effects reconsidered. *Psychonomic Bulletin & Review*, *20*(4), 615-630.

Schouppe, N., De Houwer, J., Richard Ridderinkhof, K., & Notebaert, W. (2012). Conflict: Run! Reduced Stroop interference with avoidance responses. *Quarterly Journal of Experimental Psychology, 65*(6), 1052-1058.

Smith, E. R., & Branscombe, N. R. (1985). *Stereotype traits can be processed automatically*. Unpublished manuscript, Purdue University, West Lafayette, IN.

Soutschek, A., Strobach, T., & Schubert, T. (2013). Working memory demands modulate cognitive control in the Stroop paradigm. *Psychological Research, 77*(3), 333-347.

Spivey, M. J., Grosjean, M., & Knoblich, G. (2005). Continuous attraction toward phonological competitors. *Proceedings of the National Academy of Sciences, 102*(29), 10393-10398.

Stangor, C., Lynch, L., Duan, C., & Glas, B. (1992). Categorization of individuals on the basis of multiple social features. *Journal of Personality and Social Psychology, 62*(2), 207-218.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*(6), 643-662.

Tajfel, H. (1981). *Human groups and social categories: Studies in social psychology*. Cambridge, England: Cambridge University Press.

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., ... & Nelson, C. (2009). The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry Research, 168*(3), 242-249.

West, R., & Baylis, G. C. (1998). Effects of increased response dominance and contextual disintegration on the Stroop interference effect in older adults. *Psychology and Aging, 13*(2), 206-217.

Wiggins, J. S. (1991). Agency and communion as conceptual coordinates for the understanding and measurement of interpersonal behavior. In W. Grove & D. Cicchetti (Eds.), *Thinking clearly about psychology: Essays in honor of Paul Meehl* (Vol. 2, pp. 89–113). Minneapolis, MN: University of Minnesota Press.

Zipp, S. A., & Craig, S. D. (2019). The impact of a user's biases on interactions with virtual humans and learning during virtual emergency management training. *Educational Technology Research and Development*, 1-20.

APPENDIX A

CONGRUENCY EFFECTS IN EXPERIMENT 1

CONGRUENCY EFFECTS IN EXPERIMENT 1.

Data Type	Behavior	Target Feature 1	Target Feature 2	Distractor Feature	Congruency Effect	Standard Error	*p*	Lower limits 95% C.I.	Upper limits 95% C.I.
RT (msec)	Give	Female	White	Same Race	126.25	50.292	0.0163	24.525	227.97
				Different Race	124.53	63.403	0.0567	-3.717	252.77
			Black	Same Race	134.01	46.556	0.0065	39.845	228.18
				Different Race	100.74	46.533	0.0366	6.6161	194.86
		Male	White	Same Race	48.227	41.484	0.2521	-35.683	132.14
				Different Race	126.04	48.804	0.0137	27.322	224.75
			Black	Same Race	109.61	56.602	0.0601	-4.8735	224.1
				Different Race	57.323	39.917	0.1590	-23.418	138.06
	Take	Female	White	Same Race	-44.495	13.609	0.0023	-72.023	-16.96
				Different Race	-34.506	14.967	0.0266	-64.78	-4.231
			Black	Same Race	-26.613	15.075	0.0853	-57.105	3.879
				Different Race	-44.937	16.35	0.0090	-78.008	-11.866
		Male	White	Same Race	21.404	14.959	0.1604	-8.8533	51.662
				Different Race	19.325	15.413	0.2174	-11.851	50.502
			Black	Same Race	27.907	17.766	0.1243	-8.0291	63.843
				Different Race	39.176	19.348	0.0498	0.03976	78.311
AUC (Unit)	Give	Female	White	Same Race	0.0333	0.020667	0.1152	-0.0085	0.0751
				Different Race	0.022931	0.020968	0.2808	-0.0194	0.0653
			Black	Same Race	0.002606	0.017488	0.8823	-0.0327	0.0379
				Different Race	-0.01764	0.017806	0.3281	-0.0536	0.0183
		Male	White	Same Race	-0.02661	0.016867	0.1228	-0.0607	0.0075
				Different Race	0.013547	0.014993	0.3718	-0.0167	0.0438
			Black	Same Race	0.015894	0.012082	0.1960	-0.0085	0.0403
				Different Race	-0.01744	0.014945	0.2504	-0.0476	0.012
	Take	Female	White	Same Race	-0.00134	0.01006	0.8946	-0.0216	0.0190
				Different Race	0.015382	0.012414	0.2227	-0.0097	0.0404
			Black	Same Race	-0.01462	0.008982	0.1116	-0.0327	0.0035
				Different Race	-0.00172	0.0125	0.8911	-0.0270	0.0235
		Male	White	Same Race	-0.00132	0.011426	0.9087	-0.0244	0.0217
				Different Race	0.003518	0.012538	0.7805	-0.0218	0.0288
			Black	Same Race	0.000819	0.012531	0.9482	-0.0245	0.0261
				Different Race	0.003822	0.010576	0.7198	-0.0175	0.0252

APPENDIX B

CONGRUENCY EFFECTS IN EXPERIMENT 2.

CONGRUENCY EFFECTS IN EXPERIMENT 2.

Data Type	Behavior	Target Feature 1	Target Feature 2	Distractor Feature	Congruency Effect	Standard Error	*p*	Lower limits 95% C.I.	Upper limits 95% C.I.
RT (msec)	Give	White	Female	Same Gender	-0.9058	47.82	0.9850	-97.63	95.82
				Different Gender	89.013	67.764	0.1967	-48.05	226.08
			Male	Same Gender	26.333	39.064	0.5042	-52.68	105.35
				Different Gender	-13.515	43.783	0.7592	-102.07	75.045
		Black	Female	Same Gender	106.82	56.943	0.0682	-8.363	221.99
				Different Gender	88.268	42.675	0.0453	1.948	174.59
			Male	Same Gender	62.526	55.223	0.2644	-49.17	174.22
				Different Gender	35.217	58.163	0.5484	-82.42	152.86
	Take	White	Female	Same Gender	23.816	15.617	0.1353	-7.772	55.404
				Different Gender	19.374	15.965	0.2322	-12.91	51.666
			Male	Same Gender	48.361	17.814	0.0098	12.32	84.393
				Different Gender	14.248	15.664	0.3686	-17.43	45.933
		Black	Female	Same Gender	-1.2596	12.686	0.9214	-26.91	24.4
				Different Gender	-47.679	16.875	0.0074	-81.81	-13.54
			Male	Same Gender	-40.221	15.636	0.0140	-71.84	-8.594
				Different Gender	-37.758	17.128	0.0335	-72.40	-3.113
AUC (Unit)	Give	White	Female	Same Gender	0.005017	0.019071	0.7939	-0.033	0.0435
				Different Gender	-0.00875	0.014784	0.5575	-0.038	0.0211
			Male	Same Gender	0.002809	0.017601	0.8740	-0.032	0.0384
				Different Gender	0.015558	0.018912	0.4157	-0.022	0.0538
		Black	Female	Same Gender	-0.00877	0.016459	0.5972	-0.042	0.0245
				Different Gender	-0.04263	0.019637	0.0361	-0.082	-0.002
			Male	Same Gender	-0.01	0.016518	0.5484	-0.043	0.0234
				Different Gender	-0.04631	0.014881	0.0035	-0.076	-0.016
	Take	White	Female	Same Gender	-0.00631	0.011838	0.5968	-0.030	0.0176
				Different Gender	0.010065	0.010525	0.3448	-0.011	0.0313
			Male	Same Gender	0.014506	0.016932	0.3969	-0.019	0.0487
				Different Gender	-0.00447	0.014438	0.7587	-0.033	0.0247
		Black	Female	Same Gender	0.01214	0.014849	0.4186	-0.017	0.0421
				Different Gender	-0.01251	0.012008	0.3041	-0.036	0.0117
			Male	Same Gender	0.000544	0.0106	0.9594	-0.0209	0.0219
				Different Gender	0.001187	0.010523	0.9108	-0.0201	0.0224

APPENDIX C

CONGRUENCY EFFECTS IN EXPERIMENT 3

CONGRUENCY EFFECTS IN EXPERIMENT 3.

Data Type	Behavior	Target Feature 1	Target Feature 2	Distractor Feature	Congruency Effect	Standard Error	*p*	Lower limits 95% C.I.	Upper limits 95% C.I.
RT (msec)	Give	Male	Happy	Same Emotion	46.924	35.482	0.1935	-24.787	118.63
RT (msec)	Give	Male	Happy	Different Emotion	37.6	35.712	0.2987	-34.577	109.78
RT (msec)	Give	Male	Angry	Same Emotion	82.678	32.84	0.0159	16.306	149.05
RT (msec)	Give	Male	Angry	Different Emotion	47.356	30.721	0.1311	-14.734	109.45
RT (msec)	Give	Female	Happy	Same Gender	11.484	37.578	0.7615	-64.463	87.431
RT (msec)	Give	Female	Happy	Different Gender	-58.788	53.442	0.2779	-166.8	49.221
RT (msec)	Give	Female	Angry	Same Gender	-7.2047	40.817	0.8608	-89.7	75.29
RT (msec)	Give	Female	Angry	Different Gender	-38.929	31.501	0.2238	-102.6	24.738
RT (msec)	Take	Male	Happy	Same Gender	-40.121	16.089	0.0169	-72.639	-7.6028
RT (msec)	Take	Male	Happy	Different Gender	1.5564	20.984	0.9412	-40.854	43.967
RT (msec)	Take	Male	Angry	Same Gender	-24.641	15.352	0.1164	-55.669	6.3879
RT (msec)	Take	Male	Angry	Different Gender	-41.062	16.283	0.0158	-73.972	-8.1526
RT (msec)	Take	Female	Happy	Same Gender	21.909	21.181	0.3072	-20.899	64.717
RT (msec)	Take	Female	Happy	Different Gender	53.461	13.456	0.0003	26.265	80.657
RT (msec)	Take	Female	Angry	Same Gender	40.176	16.905	0.0224	6.0096	74.343
RT (msec)	Take	Female	Angry	Different Gender	-1.554	13.249	0.9072	-28.331	25.223
AUC (Unit)	Give	Male	Happy	Same Gender	0.022879	0.017329	0.1943	-0.01215	0.057903
AUC (Unit)	Give	Male	Happy	Different Gender	-0.01022	0.016018	0.5272	-0.04259	0.022157
AUC (Unit)	Give	Male	Angry	Same Gender	0.016466	0.016968	0.3377	-0.01783	0.050759
AUC (Unit)	Give	Male	Angry	Different Gender	0.002669	0.017397	0.8789	-0.03249	0.037829
AUC (Unit)	Give	Female	Happy	Same Gender	0.014806	0.018664	0.4323	-0.02292	0.052528
AUC (Unit)	Give	Female	Happy	Different Gender	0.010692	0.018527	0.5671	-0.02675	0.048136
AUC (Unit)	Give	Female	Angry	Same Gender	-0.00078	0.018239	0.9660	-0.03765	0.03608
AUC (Unit)	Give	Female	Angry	Different Gender	-0.00991	0.017072	0.5647	-0.04442	0.02459
AUC (Unit)	Take	Male	Happy	Same Gender	-0.00237	0.007182	0.7429	-0.01689	0.012144
AUC (Unit)	Take	Male	Happy	Different Gender	0.000745	0.010528	0.9439	-0.02053	0.022024
AUC (Unit)	Take	Male	Angry	Same Gender	-0.01174	0.011092	0.2961	-0.03416	0.010674
AUC (Unit)	Take	Male	Angry	Different Gender	-0.00741	0.007668	0.3395	-0.02291	0.008085
AUC (Unit)	Take	Female	Happy	Same Gender	0.002107	0.008578	0.8073	-0.01523	0.019444
AUC (Unit)	Take	Female	Happy	Different Gender	-0.00194	0.011695	0.8692	-0.02557	0.021698
AUC (Unit)	Take	Female	Angry	Same Gender	-0.00629	0.007764	0.4228	-0.02198	0.009402
AUC (Unit)	Take	Female	Angry	Different Gender	0.019661	0.009613	0.0474	0.000233	0.039089

APPENDIX D

CONGRUENCY EFFECTS IN EXPERIMENT 4

CONGRUENCY EFFECTS IN EXPERIMENT 4.

Data Type	Behavior	Target Feature 1	Target Feature 2	Distractor Feature	Congruency Effect	Standard Error	*p*	Lower limits 95% C.I.	Upper limits 95% C.I.
RT (msec)	Give	Happy	Female	Same Gender	-0.9058	47.82	0.9850	-97.63	95.82
				Different Gender	89.013	67.764	0.1967	-48.05	226.08
			Male	Same Gender	26.333	39.064	0.5042	-52.68	105.35
				Different Gender	-13.515	43.783	0.7592	-102.07	75.045
		Angry	Female	Same Gender	106.82	56.943	0.0682	-8.363	221.99
				Different Gender	88.268	42.675	0.0453	1.948	174.59
			Male	Same Gender	62.526	55.223	0.2644	-49.17	174.22
				Different Gender	35.217	58.163	0.5484	-82.42	152.86
	Take	Happy	Female	Same Gender	23.816	15.617	0.1353	-7.772	55.404
				Different Gender	19.374	15.965	0.2322	-12.91	51.666
			Male	Same Gender	48.361	17.814	0.0098	12.32	84.393
				Different Gender	14.248	15.664	0.3686	-17.43	45.933
		Angry	Female	Same Gender	-1.2596	12.686	0.9214	-26.91	24.4
				Different Gender	-47.679	16.875	0.0074	-81.81	-13.54
			Male	Same Gender	-40.221	15.636	0.0140	-71.84	-8.594
				Different Gender	-37.758	17.128	0.0335	-72.40	-3.113
AUC (Unit)	Give	Happy	Female	Same Gender	0.005017	0.019071	0.7939	-0.033	0.0435
				Different Gender	-0.00875	0.014784	0.5575	-0.038	0.0211
			Male	Same Gender	0.002809	0.017601	0.8740	-0.032	0.0384
				Different Gender	0.015558	0.018912	0.4157	-0.022	0.0538
		Angry	Female	Same Gender	-0.00877	0.016459	0.5972	-0.042	0.0245
				Different Gender	-0.04263	0.019637	0.0361	-0.082	-0.002
			Male	Same Gender	-0.01	0.016518	0.5484	-0.043	0.0234
				Different Gender	-0.04631	0.014881	0.0035	-0.076	-0.016
	Take	Happy	Female	Same Gender	-0.00631	0.011838	0.5968	-0.030	0.0176
				Different Gender	0.010065	0.010525	0.3448	-0.011	0.0313
			Male	Same Gender	0.014506	0.016932	0.3969	-0.019	0.0487
				Different Gender	-0.00447	0.014438	0.7587	-0.033	0.0247
		Angry	Female	Same Gender	0.01214	0.014849	0.4186	-0.017	0.0421
				Different Gender	-0.01251	0.012008	0.3041	-0.036	0.0117
			Male	Same Gender	0.000544	0.0106	0.9594	-0.0209	0.0219
				Different Gender	0.001187	0.010523	0.9108	-0.0201	0.0224

APPENDIX E

MEAN RESPONSE TIMES AND AUCS FOR EXPERIMENT 5

MEAN RESPONSE TIMES AND AUCS FOR EXPERIMENT 5.

DV	Movement Orientation	Target Size	WM Load	Mean	SD	SE	95% Confidence Interval	
							Lower	Upper
RT (msec)	Vertical Task	Large Target	Control	622.2	292.7	46.3	528.6	715.8
			Low Load	743.9	231.3	36.6	669.9	817.8
			High Load	1787.8	714.7	113	1559.2	2016.4
		Medium Target	Control	686.4	287.4	45.4	594.5	778.4
			Low Load	779	241.5	38.2	701.8	856.3
			High Load	1708.5	512.9	81.1	1544.5	1872.5
		Small Target	Control	876.9	302.6	47.8	780.2	973.7
			Low Load	955.7	231.5	36.6	881.6	1029.7
			High Load	2020.9	732.1	115.8	1786.7	2255
	Horizontal Task	Large Target	Control	638.1	320.3	50.6	535.6	740.5
			Low Load	727.7	251.9	39.8	647.2	808.3
			High Load	1764.2	526.3	83.2	1595.8	1932.5
		Medium Target	Control	694.9	307.1	48.6	596.7	793.1
			Low Load	799.9	262.3	41.5	716	883.8
			High Load	1771.3	473.8	74.9	1619.8	1922.9
		Small Target	Control	904	335.5	53.1	796.7	1011.3
			Low Load	945.9	264.4	41.8	861.4	1030.5
			High Load	1856.2	539.1	85.2	1683.7	2028.6
AUC (px)	Vertical Task	Large Target	Control	10504.2	10073.2	1592.7	7282.6	13725.8
			Low Load	11598.3	8188	1294.6	8979.6	14216.9
			High Load	8356.2	11965.1	1891.8	4529.5	12182.8
		Medium Target	Control	11458.2	8945.2	1414.4	8597.4	14319.1
			Low Load	12255.5	8749.2	1383.4	9457.4	15053.7
			High Load	8996.2	11576.7	1830.4	5293.7	12698.6
		Small Target	Control	13982.9	11865.6	1876.1	10188.1	17777.7
			Low Load	10345.1	9063.9	1433.1	7446.3	13243.9
			High Load	7951.5	12784.8	2021.5	3862.7	12040.3
	Horizontal Task	Large Target	Control	11851.5	8196.2	1295.9	9230.2	14472.8
			Low Load	11496.8	7205.7	1139.3	9192.3	13801.3
			High Load	7144.7	8576.3	1356	4401.9	9887.5
		Medium Target	Control	16007.8	12219.3	1932	12099.9	19915.8
			Low Load	13912.5	12418.6	1963.5	9940.8	17884.2
			High Load	8522.7	8103	1281.2	5931.3	11114.2
		Small Target	Control	17997.4	12674.9	2004.1	13943.8	22051.1
			Low Load	14243.3	10068	1591.9	11023.4	17463.2
			High Load	10774.8	8807.1	1392.5	7958.1	13591.5

APPENDIX F

MEAN RESPONSE TIMES AND AUCS FOR EXPERIMENT 6

MEAN RESPONSE TIMES AND AUCS FOR EXPERIMENT 6.

DV	Movement Orientation	Target Size	WM Load	Mean	SD	SE	95% Confidence Interval Lower	95% Confidence Interval Upper
RT (msec)	Vertical Task	Large Target	Control	1796.9	425.5	66.5	1662.6	1931.2
			Low Load	1715.9	375.1	58.6	1597.5	1834.2
			High Load	2667.2	905.4	141.4	2381.4	2953.0
		Medium Target	Control	1771.6	460.8	72.0	1626.2	1917.1
			Low Load	1626.6	326.1	50.9	1523.6	1729.5
			High Load	2686.8	859.2	134.2	2415.6	2958.0
		Small Target	Control	1910.8	494.0	77.1	1754.9	2066.7
			Low Load	1693.6	308.7	48.2	1596.2	1791.1
			High Load	2824.4	682.5	106.6	2609.0	3039.8
	Horizontal Task	Large Target	Control	2021.8	567.4	88.6	1842.7	2200.9
			Low Load	1826.0	389.5	60.8	1703.0	1948.9
			High Load	2786.0	746.3	116.6	2550.5	3021.6
		Medium Target	Control	1982.5	419.0	65.4	1850.3	2114.8
			Low Load	1833.4	371.2	58.0	1716.2	1950.6
			High Load	2851.0	695.2	108.6	2631.5	3070.4
		Small Target	Control	2230.2	543.6	84.9	2058.6	2401.8
			Low Load	1955.7	383.8	59.9	1834.5	2076.8
			High Load	2899.9	462.8	72.3	2753.9	3046.0
AUC (px)	Vertical Task	Large Target	Control	3108.0	4524.1	706.5	1680.0	4536.0
			Low Load	7134.1	4300.8	671.7	5776.6	8491.6
			High Load	8410.9	6817.5	1064.7	6259.0	10562.8
		Medium Target	Control	3558.4	4190.2	654.4	2235.8	4881.0
			Low Load	6561.0	4462.1	696.9	5152.6	7969.4
			High Load	7321.2	7385.0	1153.3	4990.2	9652.3
		Small Target	Control	2866.8	4263.8	665.9	1520.9	4212.6
			Low Load	6007.8	5131.7	801.4	4388.0	7627.6
			High Load	7693.0	6848.1	1069.5	5531.4	9854.6
	Horizontal Task	Large Target	Control	4114.3	6083.0	950.0	2194.2	6034.4
			Low Load	8006.9	5507.6	860.1	6268.5	9745.3
			High Load	11104.6	7405.4	1156.5	8767.2	13442.1
		Medium Target	Control	3439.4	6276.3	980.2	1458.3	5420.5
			Low Load	11002.4	7163.2	1118.7	8741.4	13263.4
			High Load	11445.9	6673.9	1042.3	9339.3	13552.5
		Small Target	Control	3820.7	5954.7	930.0	1941.1	5700.3
			Low Load	10932.5	7778.4	1214.8	8477.3	13387.7
			High Load	11823.7	9407.4	1469.2	8854.4	14793.1

APPENDIX G

LIST OF HAND-MOTION FEATURES FED TO SVM CLASSIFIERS

LIST OF HAND-MOTION FEATURES FED TO SVM CLASSIFIERS.

	Hand Motion Feature	Feature Selection Result		
		Computer Mouse (Load)	Touch Screen (Load)	Touch Pad (Scale)
1	Response time	*		
2	Time to initiate the first movement (initiation time)	*	*	*
3	Absolute area under curve			
4	Area under curve	*		
5	Absolute maximum deviation		*	*
6	Maximum deviation			
7	Length of the trajectory		*	*
8	Mean velocity	*	*	*
9	Maximum velocity	*	*	*
10	Minimum velocity			
11	Mean acceleration			
12	Maximum acceleration			
13	Minimum acceleration			
14	Velocity peak onset	*	*	*
15	Onset of the lowest velocity			
16	Acceleration peak onset			*
17	Onset of the lowest acceleration	*		
18	Number of movement flips along the x axis (x flip)			*
19	Movement flips along the y axis (y flip)			
20	x entropy			
21	y entropy		*	
22	Length traveled along x axis beyond the target (x overshoot)			
23	Length traveled along y axis beyond the target (y overshoot)	*		
24	Euclidean-distance-based flip (2D flip)			
25	Euclidean-distance-based sample entropy (2D entropy)			*
26	Euclidean-distance-based overshoot (2D overshoot)			
27	Movement time (RT- Initiation time)			
28	Mean velocity at quartile 1 (Q1 Velocity)	*	*	*
29	Q2 Velocity		*	*
30	Q3 Velocity		*	
31	Q4 Velocity			
32	Q1 Acceleration	*		
33	Q2 Acceleration	*		
34	Q3 Acceleration			
35	Q4 Acceleration			
36	Q1 Radian Angle	*		*
37	Q2 Angle			
38	Q3 Angle			
39	Q4 Angle	*		
	Prediction Accuracy (Chance level)	53.2% (33.3%)	52.2% (33.3%)	34.4% (11.1%)

APPENDIX H

CONGRUENCY EFFECTS IN EXPERIMENT 7

CONGRUENCY EFFECTS IN EXPERIMENT 7.

	Working Memory Load	N-1th Trial Congruency	Congruency Effect	Standard Error	*p*	Lower limits 95% C.I.	Upper limits 95% C.I.
Gender Task RT (msec)	High	Congruent	22.1	13.936	0.1195	-5.9353	50.136
		Incongruent	6.8	14.401	0.6388	-22.167	35.774
	Low	Congruent	67.4	14.98	0.0000	37.223	97.496
		Incongruent	-36.4	13.676	0.0106	-63.918	-8.893
Emotion Task RT (msec)	High	Congruent	74.4	19.99	0.0005	34.23	114.66
		Incongruent	24.2	16.968	0.1600	-9.9132	58.357
	Low	Congruent	48.2	14.119	0.0013	19.841	76.65
		Incongruent	65.7	16.009	0.0002	33.452	97.865
Gender Task AUC (px)	High	Congruent	19764	11659	0.0967	-3691.5	43219
		Incongruent	-18051	13361	0.1832	-44930	8827.6
	Low	Congruent	23832	11612	0.0457	471.35	47192
		Incongruent	-53109	11139	0.0000	-75517	-30701
Emotion Task AUC (px)	High	Congruent	29472	12408	0.0217	4510.3	54435
		Incongruent	13885	8136.9	0.0945	-2484.7	30254
	Low	Congruent	50877	11968	0.0001	26801	74954
		Incongruent	56129	10497	0.0000	35012	77245

APPENDIX I

CONGRUENCY EFFECTS IN EXPERIMENT 8

CONGRUENCY EFFECTS IN EXPERIMENT 8.

	Working Memory Load	N-1th Trial Congruency	Congruency Effect	Standard Error	*p*	Lower limits 95% C.I.	Upper limits 95% C.I.
Gender Task RT (msec)	High	Congruent	36.208	12.153	0.004483	11.787	60.63
		Incongruent	1.4868	10.585	0.88887	-19.785	22.759
	Low	Congruent	24.586	12.142	0.048344	0.18614	48.986
		Incongruent	4.0389	11.884	0.7354	-19.842	27.92
Race Task RT (msec)	High	Congruent	14.688	10.671	0.17493	-6.7556	36.131
		Incongruent	-2.8382	9.634	0.76954	-22.199	16.522
	Low	Congruent	18.697	11.548	0.11186	-4.5102	41.904
		Incongruent	-20.817	13.844	0.13908	-48.637	7.0033
Gender Task AUC (px)	High	Congruent	35092	12500	0.007148	9973.3	60211
		Incongruent	1989.6	11956	0.86852	-22036	26016
	Low	Congruent	6140.8	8114.3	0.4528	-10166	22447
		Incongruent	5452.4	9925.2	0.58526	-14493	25398
Race Task AUC (px)	High	Congruent	526.92	9980.3	0.95811	-19529	20583
		Incongruent	1420.8	8496.6	0.86789	-15654	18495
	Low	Congruent	11029	11482	0.3415	-12045	34103
		Incongruent	-13996	11919	0.24596	-37949	9956.3

APPENDIX J

IRB PERMISSION FOR HUMAN SUBJECT TESTING

IRB PERMISSION FOR HUMAN SUBJECT TESTING

ASU Knowledge Enterprise Development

APPROVAL:CONTINUATION

David Becker
Human Systems Engineering (HSE)
480/727-1151
Vaughn.Becker@asu.edu

Dear David Becker:

On 8/16/2018 the ASU IRB reviewed the following protocol:

Type of Review:	Continuing Review
Title:	Investigating implicit bias on gender and race through unconscious responses.
Investigator:	David Becker
IRB ID:	STUDY00006766
Category of review:	(4) Noninvasive procedures, (7)(b) Social science methods, (7)(a) Behavioral research
Funding:	Name: Sigma Xi
Grant Title:	None
Grant ID:	None
Documents Reviewed:	• Participant Demographic Form, Category: Screening forms; • Recruitment material, Category: Recruitment Materials; • Consent_Form_Word_Categorization, Category: Consent Form; • Consent_Form_Mousetracking, Category: Consent Form;

The IRB approved the protocol from 8/16/2018 to 8/15/2023 inclusive. Three weeks before 8/15/2023 you are to submit a completed Continuing Review application and required attachments to request continuing approval or closure.

If continuing review approval is not granted before the expiration date of 8/15/2023 approval of this protocol expires on that date. When consent is appropriate, you must use final, watermarked versions available under the "Documents" tab in ERA-IRB.

In conducting this protocol you are required to follow the requirements listed in the INVESTIGATOR MANUAL (HRP-103).

Sincerely,

IRB Administrator

cc: Hansol Rheem
Hansol Rheem
David Becker
Vipin Verma