THE USE OF SOCIO-COGNITIVE MINDFULNESS IN MITIGATING IMPLICIT BIAS
AND STEREOTYPE-ACTIVATED BEHAVIORS

By
Christelle Tchangha Ngnoumen

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
in
Psychology

at
Harvard University Graduate School of Arts and Sciences

May 2019

Under the supervision of Professors Ellen Langer, Ph.D., Mina Cikara, Ph.D.,
Jack Demick, Ph.D. and Leah Somerville, Ph.D. at Harvard University
The following seven experiments were founded on the observation of a significant relationship between trait socio-cognitive mindfulness, as measured using the Langer Mindfulness Scale (LMS) and reduced implicit racial bias on the Black-White Implicit Attitudes Test (IAT), which suggested that more mindful individuals have a weaker implicit preference for white Americans over black Americans. This initial observation led to a subsequent interest in experimentally-introducing state socio-cognitive mindfulness as a prejudice reduction intervention, with implications for those contexts in which implicit bias has been found to predict social injustices and discrimination including health care, education, law enforcement, and the labor force.

Five experiments in Chapter 3 evaluated the effectiveness of a state socio-cognitive mindfulness-based paradigm, rooted in novelty-production and engagement, in reducing implicit racial bias, in reducing implicit associations between race and criminality, in reducing implicit face-based bias related to facial maturity, and in increasing intergroup empathy. Across four experiments, having participants re-sort the same social stimuli (e.g., faces of people) according to novel, psychological categories was most effective at reorganizing their preconceptions and at reducing implicit social bias compared to re-sorting according to familiar, physical categories that likely activated stereotypes related to the respective social stimuli. A fifth, follow-up
experiment examining the driving forces behind the effectiveness of the state mindfulness intervention suggested that the element of novelty (which comes from sorting according to novel categories) is both necessary and sufficient to the effectiveness of the intervention, whereas the element of engagement (which involves participants creating their own sorting categories) is only sufficient. Additionally, the success of the intervention depends on its content and on the content’s degree of relation to the target of intervention. The same follow-up experiment revealed that this mindfulness intervention for reducing implicit racial bias toward African Americans was only effective when the sorting task involved black and white faces; the intervention was not successful toward reducing anti-black bias when the sorting task involved Latino and white faces.

Chapter 3 has implications for closing the observed empathy deficit that contributes to observed social disparities in U.S. health care (in so far as negligence in health care treatment is predicted by differences in intergroup empathy), for social disparities in U.S. education and law enforcement (in so far as they are predicted by racial stereotypes that link non-whites to related notions of “troublemaker” and “criminal”), and for social disparities in U.S. labor markets (in so far as they are predicted by face-based stereotypes that link babyfacedness to incompetence and to lower perceived leadership capability).

Chapter 4 shifts away from examining the implications of state socio-cognitive mindfulness on attitudes and emotions at the group level toward focusing on the capacity for socio-cognitive mindfulness induction to mitigate frustration, particularly third-party punishment at the individual level, with implications for the affective obstacles faced by police officers that likely lend to some of the observed racial disparities in law enforcement involving use of force. This chapter examines the effectiveness of a related socio-cognitive mindfulness intervention,
rooted in novelty-production, at reducing frustration. Across two experiments, participants receiving mindful instructions to understand negative events in novel ways following a frustrating experience demonstrated lower levels of frustration and decreased third-party punishment as indicated by the selection of more positively-valenced photos, on average, for other people (e.g., strangers) to see.
DEDICATION

To the young man from Bafoussam whose courage and optimism were far greater than others' perceived limitations and far bigger than what he could fit into a bag crossing foreign spaces to begin what would ultimately become my story. This document is dedicated to my father. Dad, thank you for being the first person to teach me that mindsets and attitudes are everything and for consistently embodying this philosophy. Against stereotypes and false myths about black fathers, you raised five daughters and a son into six awesome humans. Thank you for instilling in us the notions that courage, success and love begin with choice.
ACKNOWLEDGEMENTS

This work would not have been possible without the support of many individuals including my committee members. I would like to thank my advisor, the Mother of Mindfulness, Ellen Langer, for her incredible support and guidance throughout my years as a graduate student. Huge thanks as well must be extended to the rest of my committee members for their exceptional critical guidance throughout my dissertation process, particularly in terms of helping me shape my research questions and approach; many thanks to Mina Cikara, Jack Demick, and Leah Somerville.

I am fortunate to have been able to work with a great team of research assistants. I would like to thank Delaney Snowden, Blake Evertsen, Hannah Emerson, George Vasili and Julia Sager for their dedication and many hours of work on the projects. Additional thanks must be extended to my collaborators, colleagues and closest cohort members Katherine Bercovitz, Chanmo Park, Andrew Reece and Maya Tateyama. Huge thanks to Nira Liberman, Karina Schumann, Jennifer Eberhardt, Allison Seitchik, Katharina Reinecke, Deborah Phillips, the Todorov Lab, and Robert Most.

Finally, one cannot accomplish such feats without the inspiration of friends and spirit animals. I would therefore like to thank those who have supported me over the past few years while completing this work; in particular, many thanks to Kelsey, Françoise, Jonathan, Victoria, Kara, Deja, Neda, Aly, Perrey, Joe, Amanda, Mary, Meg, Cheryl, Jarred, Tulani, Chiara, and Tristan.

Most importantly, I wish to thank all of the individuals who served as participants in this research. Without their interests and efforts none of this work would have been possible.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>iii</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>vii</td>
</tr>
<tr>
<td>Table of Contents</td>
<td>viii</td>
</tr>
<tr>
<td>List of Tables</td>
<td>xi</td>
</tr>
<tr>
<td>Chapters</td>
<td></td>
</tr>
<tr>
<td><strong>I. Introduction</strong></td>
<td>1</td>
</tr>
<tr>
<td>Contributions of Dissertation</td>
<td>2</td>
</tr>
<tr>
<td>Organization of Dissertation</td>
<td>4</td>
</tr>
<tr>
<td><strong>II. Literature Review</strong></td>
<td></td>
</tr>
<tr>
<td>Prejudice and Implicit Stereotyping</td>
<td>7</td>
</tr>
<tr>
<td>Role of Implicit Stereotyping in Race-Based Social Disparities</td>
<td></td>
</tr>
<tr>
<td>Across Various Domains</td>
<td>8</td>
</tr>
<tr>
<td>Racial Disparities in U.S. Health Care</td>
<td>8</td>
</tr>
<tr>
<td>Racial Disparities in U.S. Education</td>
<td>10</td>
</tr>
<tr>
<td>Racial Disparities in U.S. Law Enforcement</td>
<td>12</td>
</tr>
<tr>
<td>Frustration in Police Officers and Implications for Research on Implicit Stereotyping and Interventions to Address Racial Disparities in U.S. Law Enforcement</td>
<td>14</td>
</tr>
<tr>
<td>Race- and Face-Based Discrimination in U.S. Labor Market</td>
<td>15</td>
</tr>
<tr>
<td>Face-Based Social Disparities in Employment and Hiring Selection</td>
<td>17</td>
</tr>
<tr>
<td>Intergroup Empathy: An Empathy Gap Explains Observed Social Disparities Across Different Domains</td>
<td>20</td>
</tr>
</tbody>
</table>
Empathy and Socio-Cognitive Mindfulness 22

The Malleability of Implicit Stereotypic Beliefs and its Implication for Socio-Cognitive Mindfulness 24

Socio-Cognitive Mindfulness 26

A. Pilot Study Examining Relationship between Trait Socio-Cognitive Mindfulness and Implicit Racial Bias 28

III. Using State Socio-Cognitive Mindfulness Induction to Shift Attitudes and Emotions at The Group Level

1. Using State Mindfulness Induction to Reduce Implicit Racial Bias 34

2. Using State Mindfulness Induction to Reduce Implicit Associations between Race and Criminality 41

3. Using State Mindfulness Induction to Reduce Implicit Face-Based Bias 46

4. Comparing the Effects of Process and Content in Mindfulness-Based Paradigm in Reducing Implicit Racial Bias 52

5. Using State Mindfulness Induction to Increase Intergroup Empathy 58

IV. Using State Socio-Cognitive Mindfulness Induction to Shift Emotions at The Individual Level

6. Using State Mindfulness Induction to Decrease Frustration I 66

B. Validation Pilot Study for Frustration Induction Paradigm 68

C. Validation of Photo Selection Task’s 30 Photos’ Explicit Perceived Harmfulness, as Rated by Naïve American Participants in 2019 72
7. Using State Mindfulness Induction to Decrease Frustration II

V. Discussion

General Summary

Limitations

Recommendations and Future Directions

Appendix

A. State Mindfulness Induction Paradigm Levels and Stimuli
   (Black & White)
   102
B. Black-White Race Implicit Association Test (IAT)
   103
C. Race-Weapons Implicit Association Test (IAT)
   105
D. Face Evaluation Task Instructions
   107
E. Big Five Inventory (BFI-44)
   108
F. Frustration Induction Paradigm
   109
G. State Mindfulness Induction Writing Exercises (Experiments 6 & 7)
   113
H. Anger Rating Scale (ARS)
   114
I. Photo Rating Task (Perceived Harmfulness)
   115
J. Empathic Listening Task (Behavioral Measure of Empathic Effort)
   116
K. Other and Self Emotion Evaluations
   117
L. Interpersonal Reactivity Index (Empathic Concern Subscale)
   118
M. Langer Mindfulness Scale (LMS)
   120

References

121
LIST OF TABLES

1. Correlation Matrix of LMS Total Score, LMS Subscale Scores, and IAT Score (Pilot Study) 32
2. Means and standard deviations of black-white IAT scores across four state mindfulness induction levels (Experiment 1) 40
3. Means and standard deviations of race-weapons IAT scores and of mindfulness task completion time between two state mindfulness induction levels (Experiment 2) 45
4. Means and standard deviations of participants’ selections of babyfaced individuals for high-dominant and high-competent roles between two state mindfulness induction levels (Experiment 3) 51
5. Means and standard deviations of black-white IAT scores across four state mindfulness induction levels (Experiment 4) 57
6. Means and standard deviations of black-white IAT scores between state mindfulness inductions involving black-white and Latino-white content (Experiment 4) 58
7. Means and standard deviations of time spent listening to an emotional story by state mindfulness induction level and target race (Experiment 5) 64
8. Means and standard deviations of differences between participants’ evaluations of Natasha’s emotions and their own between state mindfulness induction levels (Experiment 5) 65
9. Means and standard deviations of participants’ selected photos’ valence,
perceived harmfulness, and color darkness proportion between state mindfulness induction levels (Experiment 6)

10. Means and standard deviations of participants’ selected photos’ valence, perceived harmfulness, and color darkness proportion across three state mindfulness induction levels (Experiment 7)
CHAPTER 1
INTRODUCTION

*A great many people think they are thinking when they are merely rearranging their prejudices.*

— William James, 1890

This project approached the question of social attitude change from the perspective that an individual’s attitude toward a group can be conceptualized as having cognitive (e.g., stereotypic or associational), affective (e.g., valenced), and behavioral components (Er-rafiy & Brauer, 2012; Breckler, 1984; Leyens, Yzerbyt, & Schadron, 1994), and therefore sought to develop and test the effectiveness of state mindfulness inductions for shifting social attitudes at each of these levels. The work builds off the fact that our everyday thinking is a form of prejudice—akin to mindlessness or routinized thinking—that helps us navigate our world, on the one hand, yet which can be limiting and create social distance on the other hand. *How often do we find ourselves exploring new ways of thinking, or trying on views other than those we have held onto for years?* More importantly, *What value could we gain from re-sorting our prejudices according to these new ways of thinking?*

William James (1890) accurately remarked on one of the classic reasons behind why we do not change: most of the time, we are mindlessly recycling ideas and associations in a manner that merely strengthens premature cognitive commitments, or PCCs (Langer, 1989; Chanowitz & Langer, 1981). According to Langer (1989), PCCs are early beliefs about reality that shape our perception and subsequent experiences. These beliefs are adopted out of context—either through repeated exposure or single exposure—and are applied unconditionally, with very little room for
perceived harmfulness, and color darkness proportion between state mindfulness induction levels (Experiment 6) 79

10. Means and standard deviations of participants’ selected photos’ valence, perceived harmfulness, and color darkness proportion across three state mindfulness induction levels (Experiment 7) 87
A great many people think they are thinking when they are merely rearranging their prejudices.

— William James, 1890

This project approached the question of social attitude change from the perspective that an individual’s attitude toward a group can be conceptualized as having cognitive (e.g., stereotypic or associational), affective (e.g., valenced), and behavioral components (Er-rafiy & Brauer, 2012; Breckler, 1984; Leyens, Yzerbyt, & Schadron, 1994), and therefore sought to develop and test the effectiveness of state mindfulness inductions for shifting social attitudes at each of these levels. The work builds on the fact that our everyday thinking is a form of prejudice—akin to mindlessness or routinized thinking—that helps us navigate our world, on the one hand, yet which can be limiting and create social distance on the other hand. How often do we find ourselves exploring new ways of thinking, or trying on views other than those we have held onto for years? More importantly, What value could we gain from re-sorting our prejudices according to these new ways of thinking?

William James (1890) accurately remarked on one of the classic reasons behind why we do not change: most of the time, we are mindlessly recycling ideas and associations in a manner that merely strengthens premature cognitive commitments, or PCCs (Langer, 1989; Chanowitz & Langer, 1981). According to Langer (1989), PCCs are early beliefs about reality that shape our perception and subsequent experiences. These beliefs are adopted out of context—either through repeated exposure or single exposure—and are applied unconditionally, with very little room for
variation. To the degree that our thinking involves a mindless “rearranging of old prejudices,” as James originally noted, this dissertation explores the effectiveness of a socio-cognitive mindfulness intervention designed to re-sort old prejudices according to novel categories. The elements of novelty-production and engagement were borrowed from socio-cognitive mindfulness theory and were applied to loosen the strengths of previously-established associations by prompting people to look at outgroup members in new ways. Changes in the strengths of old ideas as a function of the intervention were assessed using implicit, reaction-time measures of association. How this re-sorting of prejudices according to new ways of thinking affected attitude change at the behavioral level was also examined.

**Contributions of Dissertation**

For well over 60 years, social psychologists have been concerned with finding techniques to reduce prejudice (Allport, 1954; Bar-Tal, 1989; Festinger & Carlsmith, 1959; McKay & Pitman, 1993; Meshel & McGlynn, 2004; Pettigrew, 1998; Stangor, 2000). More “traditional”—traditional in the sense of operating at a more conscious and deliberate level of awareness—prejudiced attitudes have declined significantly in recent decades (Sritharan & Gawronski, 2010). While there has also been evidence of a shift in implicit (e.g., unconscious) biases toward neutrality (Charlesworth & Banaji, 2019), studies examining aggregate levels of implicit bias (e.g., looking by countries; states; counties) still find a strong association between aggregate levels of implicit bias and disparities and discrimination suggesting that individual differences in implicit bias predict individual differences in behavioral outcomes (Payne, Vuletich, & Lundberg, 2017).

Implicit stereotypes and biases can influence people’s beliefs, actions and decisions.
outside of explicit awareness. This means that individuals demonstrating implicit stereotypes may not always be consciously aware of their existence (e.g., are not aware of having them) and may not necessarily explicitly express those beliefs (e.g., are unable to verbally express them to themselves or to others).

What has made the task of reducing prejudice necessarily challenging is this observed split between explicit and implicit prejudice as well as the observed automaticity of implicit stereotype activation (Brewer 1988; Devine 1989; Pratto and Bargh 1991; Bargh et al. 1996; Devine 1989). Since Allport’s initial observations, psychologists have acknowledged the powerful influence of those more indirect, subtle, and seemingly innocuous expressions of implicit stereotypes and prejudice (Banaji & Greenwald, 1994; Bem & Bem, 1970; Brewer, 1988; Crosby, Broomley, & Saxe, 1980).

A prevailing view of implicit stereotypic attitudes and is that they are automatic, slow-forming, and develop gradually through experience and learning. Under this view, implicit attitudes are believed to be more stable and resistant to change compared to more explicit attitudes (Banaji, 2001; Gregg, 2000; Smith & DeCoster, 1999; Bargh & Chartrand, 1999; Devine, 1989). Growing research on implicit attitudes demonstrates some surprising malleability, however (Kurdi & Banaji, 2019; Lai et al., 2016; Lai et al., 2014; Lai, Haidt, & Nosek, 2014; Lai, Hoffman, & Nosek, 2013).

The following work adds to the current body of knowledge by presenting a novel, accessible, efficient, reliable, and adaptable method for shifting implicit attitudes and their potential effects on behaviors. This project develops and tests the effectiveness of state mindfulness-based paradigms, rooted in novelty-production, in reducing implicit social bias, in increasing intergroup empathy, and in reducing frustration. Socio-cognitive mindfulness is
stressed as a unique prejudice-reduction technique that, compared to many other forms of prejudice-reduction techniques—including meditative mindfulness—is more readily accessible (e.g., easy to administer online), economical (e.g., takes under five minutes and requires eight stimulus photographs at minimum), and is adaptable, demonstrating effectiveness at reducing implicit and explicit biases across different domains and sample groups.

**Organization of Dissertation**

This research is motivated by an observation of pervasive race- and face-based social disparities than span multiple domains of American society. The experiments comprising Chapters 3 and 4 of this dissertation either directly or indirectly address some of these documented race- and face-based social disparities by using a socio-cognitive mindfulness-based intervention to target the “criminal” stereotype often involved in greater discrimination toward non-white students compared to their white peers in academic contexts and toward non-white citizens and defendants compared to whites in the context of criminal justice. The studies also validate the intervention’s effectiveness at reducing implicit face-based bias with implications for an observed babyface bias that negatively affects leadership selection and which often overlaps with gender and racial stereotypes. Furthermore, the “empathy gap” believed to account for many of these observed social disparities is also addressed in this work via the use of socio-cognitive mindfulness to increase affective empathy and intergroup empathy toward outgroup members.

The experiments in Chapter 4 are founded on research on police officers that suggests greater police brutality toward non-whites compared to whites might be exacerbated by increased frustration and anger among police officers as a function of their job. The research is also
prompted by prior work showing that anger increases reliance on mental shortcuts and on System I [automatic or mindless] processing, including stereotypes. The studies in Chapter 4 therefore test a socio-cognitive mindfulness-based intervention’s effectiveness at reducing frustration and third-party punishment as a way to indirectly address this larger social problem at the individual level.

This dissertation is organized in five chapters. Chapter 2 delves into an overview of the body of literature that has been written on the major topics of the research and research questions, primarily within the areas of prejudice and stereotyping, the implications of stereotyping for social disparities (including face-ism) and frustration, empathy and intergroup empathy, and socio-cognitive mindfulness.

Chapter 3 will introduce the five experiments focused on the use of a state mindfulness-based paradigm, rooted in novelty-production, in reducing implicit racial bias, reducing implicit face-based bias related to facial maturity, reducing implicit associations between race and criminality, and increasing affective and intergroup empathy. These studies are related in so far as they use a similar socio-cognitive mindfulness-based paradigm to shift attitudes and emotions at the group level. This chapter includes an experiment (Experiment 4) aimed at further delineating the driving force of this intervention’s effectiveness.

Chapter 4 (2 experiments) focuses on the application of a similar state socio-cognitive mindfulness induction paradigm, rooted in novelty-production, in reducing frustration at the individual level. The experiments in Chapter 4 look at frustration using a behavioral dependent measure as well as assess the degree to which socio-cognitive mindfulness predicts prosocial behavior following a frustrating experience with implications for the effects of frustration on implicit stereotyping and on some of the social disparities discussed in Chapter 2.
Both Chapter 3 and Chapter 4 provide detailed information about the methodologies used in all of these experiments and validation/pilot studies, including participants, instruments, materials, procedures, analyses, and results. Where applicable, experiments are identified through the use of numbers (e.g., 1 - 7) and pilot/validation studies are identified using letters (e.g., A; B; C). Where appropriate, results pertaining to any additional analyses deemed either complementary to or supplementary to major analyses will be included. Tables and figures will be referenced in-text.

Chapter 5 interprets the results from Chapters 3 and 4 and provides overall conclusions in light of the research questions and in conjunction with relevant literature. Limitations of interpretation and suggestions for future directions are also presented.
CHAPTER 2
LITERATURE REVIEW

1. Prejudice and Implicit Stereotyping

The relationship between categorical thinking and prejudice was first systematically explored by Gordon Allport (1954) in his seminal book, *The Nature of Prejudice*. Allport claimed prejudice is a natural outgrowth of categorical thinking. He wrote, “The human mind must think with the aid of categories. Once formed, categories are the basis for normal prejudgment. We cannot possibly avoid this process. Orderly living depends upon it (p. 20)”

According to Allport, we categorize people into different kinds of social group memberships similarly to how we can categorize objects into different kinds. Social categorization represents a natural cognitive process that consequently shapes how we think, feel, and behave toward people. We engage in social categorization when we think of someone as an African American person (versus an Asian or Caucasian person), as a man (versus a woman), as an older adult (versus a child), etc. (Allport, 1954).

There is nothing inherently wrong with categorical thinking. In fact, as humans, we are inclined to rely on this kind of heuristic during decision-making in order to efficiently process a wealth of information. Stereotypes help us navigate the social world (Johnson, Kim, & Keil, 2016). Social categorization becomes problematic, however, in cases where demarcations of social categories (e.g., race; sex; age) inform our behavior to the extent of predicting injustices or disparities in certain social groups’ exercise of human rights and access to public systems. This is the point where negative beliefs and qualities people associate with particular social group
categories, more generally, begin to inform their judgments of individual members of those groups, an attributional process commonly known as stereotyping. Social categorization therefore becomes problematic when the mere presence of a dark-skinned male automatically triggers thoughts that he is African American and thus more inclined to become hostile, leading strangers to sit further away from him, to evaluate his behavior as more aggressive, to misremember him as the one holding the razor following the description of a crime, to decide to stop and search him for weapons because of an assumption of his greater likelihood of engaging in criminal activity, and to punish him by suspending him from school at a young age or by imposing a harsher criminal sentence on him compared to his white counterparts (Okonofua & Eberhardt, 2016; Eberhardt et al., 2008, Judd et al., 2004; Eberhardt et al., 2004; Hugenberg & Bodehnausen, 2004, 2003; Devine, 1989; Duncan, 1976).

2. Role of Implicit Stereotyping in Race-Based Social Disparities Across Various Domains

This first section reviews the current body of knowledge on racial disparities in three related domains: U.S. health care, U.S. education, U.S. law enforcement, and the U.S. labor market. The literature demonstrates racial disparities within these systems and suggests implicit racial bias could be an influencing factor in the decision-making processes accounting for the observed differences in outcomes between non-whites and whites.

2.1. Racial Disparities in U.S. Health Care

Evidence demonstrates stark differences in the treatment of whites and non-whites by medical professionals that could be explained by implicit racial bias. Skinner et al. (2003) found
disparities in rates of knee arthroplasties; after controlling for their geographic locations and insurance, African American men received knee arthroscopies less frequently compared to their white counterparts. In another study, African American men received fewer cardiac catheterizations following acute myocardia infarctions compared to white men (Chen et al., 2001).

Racial disparities in pain management have also been observed in the treatment of migraines and back pain, in cancer care among the elderly, and in children with orthopedic fractures (Bernabei et al., 1998). A 2008 review of 13 years-worth of emergency room visits found that, for pain-related visits, opioid prescriptions were less likely for African American patients compared to white patients (Fletcher et al., 2008). Non-white patients were also less likely to receive adequate analgesia (Cleeland et al., 2004).

Different explanations have been proposed for these observed racial disparities in health outcomes between whites and non-whites. While those espousing a “racial bias” outlook claim implicit bias and stereotyping in doctors are responsible for the inferior medical care observed among certain social groups, those espousing a more “socioeconomic” outlook claim that the observed disparities vary by race but not because of it, and assert that patients’ economic and geographic backgrounds are more predictive of their differential access to medical care. It is difficult to disentangle the “socioeconomic” argument from the “racial bias” one, however, when one considers how inextricably linked race, ethnicity, and socioeconomic status are in the United States such that communities are often segregated by the joint combination of these variables. Rather than looking at where medication and medical treatment are and are not available, perhaps the more important question is who receives it and who does not in cases where it is readily available. Empathy research suggests that even in cases where availability is controlled
for and medical treatment is readily available to all, some providers still seem to fail to see the
need to administer the treatments; that is, in some cases there is still a blindness to the pain
experiences of certain patients compared to others.

2.2. Racial Disparities in U.S. Education

Significant racial disparities in education in the United States appear to follow a pattern
in which African American, Native American, Latino, and Southeast Asian groups underperform
academically compared to their Caucasian and Asian-American counterparts. Valuable
experimental work points to differential or biased treatment of racial minority students within the
educational system as the source of some of these observed differences. Racial differences in
achievement resulting from discrimination by educators represent the more obvious and
egregious forms of racial disparities. Racial discrimination takes many different forms and
expressions. Discrimination may reflect hostility or patronizing attitudes, expressed in explicit or
implicit forms (Devine, Plant, Amodio, Harmon-Jones, 2002; Dovidio, Kawakami, & Gaertner,
2002), and can be experienced as microaggressions (Sue, 2010) or as more overt forms of
aggression, all of which stigmatize racial minority groups and contribute to educational
disparities.

Research shows that African American students, particularly African American males,
are disciplined 3.5 times more often (e.g., receive more suspensions and expulsions) compared to
their white peers (Lewin, 2012). During the 2013 – 2014 school period, African American
students comprised 26% - 75% of the students involved in school-related arrests or referred to
law enforcement depending on the state, with Virginia being the highest. While African
American students make up only a small percentage of students enrolled in these sampled schools, they account for a greater proportion of the students suspended and/or expelled (Blad & Harwin, 2017). Skiba (2000) found that in many of these cases, African American students are disciplined for less serious and more subjective reasons such as talking back to teachers and inappropriate dress (Skiba et al., 2009). Contrary to expectations, and particularly to the “troublemaker” stereotype associated with black boys (which precurses the “criminal” stereotype when they become adult black men), there is no support for the hypothesis that African American students act out more than other students.

McKown and Weinstein (2008) demonstrated that teacher implicit racial bias predicted observed racial differences in students’ educational achievement over the course of one academic year. In their study, teacher implicit racial bias was evidenced by teachers’ treatment of children of equivalent academic abilities differently based on their racial status.

More recently, Okonofua and Eberhardt (2015) observed racial disparities in teachers’ disciplining practices marked by a greater likelihood of attributing black students’ second infractions to larger patterns of misconduct compared to their white counterparts. These perceptions further informed teachers’ decisions to suspend students. Implicit racial bias in education and in teachers’ treatment of students is hugely problematic for members of certain social groups, not only because it systematically contributes to academic failure, but because it can also, alongside lower socioeconomic status, carve a direct path toward incarceration. The growing implementation of zero-tolerance policies by U.S. schools funnels more students of color out of the education system through suspensions and expulsions for minor infractions. While away from school, these students fall behind academically, and those entering the juvenile justice system often find it difficult to return to school. Implicit racial bias also predicts the
misunderstanding (e.g., misreferrals) and misdiagnosis of mental health symptomology among students of certain ethnic and socioeconomic backgrounds compared to others, further contributing to a bias in the school-to-prison pipeline.

2.3. Racial Disparities in U. S. Law Enforcement

There is evidence of racial disparities at many levels of law enforcement, from traffic stops to drug-related arrests to use of force. For most law enforcement officers, enforcing the law to protect citizens involves uncontrolled and highly stressful conditions. When we are stressed, we become mindless (e.g., narrow-focused), our cognitive processing becomes limited, making us more susceptible to mindless processing of information and more likely to engage in stereotypical thinking and in stereotype-activated behaviors (Langer, 2009; Kuhlmann, Piel, & Wolf, 2005). In the context of law enforcement and community safety, implicit bias has been shown to have a significant influence in the outcomes of interactions between police and citizens. Implicit bias can distort one’s perception and subsequent treatment either in favor of or against a given individual or social group. In policing, for example, implicit bias has resulted in widespread practices that focus undeserved suspicion or unnecessary use of force and aggression on members of some groups compared to others.

In a study reviewing thousands of use of force episodes from 12 police departments across the nation, Goff et al. (2016) found that African Americans are far more likely, compared to whites and other racial groups, to be the victims of use of force by police officers, even when racial disparities in crime are taken into account and controlled for. Other data show that black people are also more likely to be stopped by police. Eberhardt et al. (2016) analyzed data from
the Oakland Police Department and found that while black residents make up 28 percent of the Oakland population, they accounted for 60 percent of police stops. Moreover, black men were four times more likely than their white counterparts to be searched during a traffic stop, even though officers were no more likely to recover contraband when searching black suspects compared to white suspects.

The stereotype of African American males as hostile and criminal has been widely documented by social psychologists (Eberhardt et al., 2004). This implicit stereotype is robust and strong, as are the behaviors that it elicits and predicts from others. In a study on implicit bias in policing decisions, Greenwald, Oakes, and Hoffman (2003) had undergraduate students in the role of plainclothes police officers make split-second decisions about whether or not to use deadly force (e.g., shoot) after being presented with flash images of unarmed citizens, fellow police officers, or armed criminals. The allotted decision time was less than one second. The subjects in the images were each dressed alike, and were either white males or black males. The study found that blacks were incorrectly shot more often than whites. Researchers at Florida State University replicated the study using only police officers and found that officers were more likely to mistakenly shoot unarmed black suspects than unarmed white suspects. Additionally, when the suspects were armed, the officers were slightly more likely to mistakenly not shoot white suspects compared to black ones.

The results from these studies suggest that implicit racial attitudes can influence a police officer’s decisions in the face of a possible threat. More specifically, the results show that implicit racial attitudes can contribute to a “weapon bias” (the finding that participants are faster at detecting guns faster in the presence of a black face compared to a white face) or to a “shooter bias” (the finding that people are quicker to shoot a black target than they are a white one) and
can affect officers’ abilities to discriminate between weapons and harmless objects; the results also show that the race of a possibly armed person can make officers assume that the person is armed (Mekawi & Bresin, 2015; Correll, Urland, & Ito, 2006; Payne, 2006; Correll et al., 2002). Other manifestations of implicit racial bias in the context of criminal justice include public defendants’ prioritization of white defendants’ cases over those of non-whites.

2.3.1. Frustration in Police Officers and Implications for Research on Implicit Stereotyping and Research on Designing Interventions to Address Racial Disparities in U. S. Law Enforcement

Research examining the impact of affect on social perception has found that negative mood states, such as anxiety and anger, exacerbate implicit stereotypic thinking (Krauth-Gruber & Ric, 2000; Bodenhausen, 1993). One study demonstrated that anxiety increases cognitive load, thereby prompting stereotyping as a means for preserving cognitive-processing resources (Macrae, Milne, & Bodenhausen, 1994; Wilder, 1993; Darke, 1988). In a study by Bodenhausen, Sheppard, and Kramer (1994), angry participants demonstrated greater reliance on heuristic cues, particularly an increase in implicit stereotypic thinking related to race.

Bodenhausen, Sheppard, and Kramer’s findings have implications for some of the previously-discussed social disparities in law enforcement, such as use of force, in light of the overwhelming amount of evidence suggesting that frustration and anger levels are higher among police compared to the population; furthermore, because of being frequently angry and/or frustrated, officers are also more likely to support more physical and/or aggressive policing methods (Pew Research Center Survey, 2016). Fifty-one percent of polled U.S. police officers
reported feeling frustrated as a result of their job compared to only twenty-nine percent of other polled employed Americans (Pew Research Center Survey, 2016). If angry and frustrated officers are more likely to use force, as well as engage in stereotypical thinking, then prejudice reduction techniques must consider ways to mitigate the affective components of law enforcement, as a vocation, that may be contributing to some of the observed racial biases in the context of law enforcement.

2.4. Race- and Face-Based Discrimination in U.S. Labor Market

This next section reviews the current body of knowledge on race- and face-based discrimination in the U.S. labor force and employment opportunities. This literature suggests that implicit stereotyping related to racial and facial characteristics informs many of the decision-making processes that account for observed differences in hiring and employment outcomes between whites and non-whites as well as between babyfaced and maturefaced individuals in organizational settings, respectively. How perceived facial maturity maps onto face-based ageism and overlaps with both gender and racial stereotypes to affect labor outcomes is also discussed in this section.

Modern American labor markets have become increasingly diverse. A glance at the U.S. employment-population ratio (e.g., the proportion of the population that is employed) in 2017 reveals that employment-population ratios across the race and ethnicity groups ranged from 55.6% for American Indians and Alaska Natives to 62.9% for Native Hawaiians and Other Pacific Islanders. Moreover, in 2017, the employment–population ratio was 57.6% for Blacks, 60.4% for Whites, 61.5% for Asians, 62.4% for individuals of Two or More Races, and 62.7%
for Hispanics (Bureau of Labor Statistics, 2017). Despite this diversity in labor force participation, however, a closer look at the composition of executive-level leadership positions reveals that it continues to be dominated by white males (Calvert, 2013; Catalyst, 2016; Federal Glass Ceiling Commission, 1995). Today still, even though Asian Americans and women are relatively well-represented in professional roles, they are nevertheless underrepresented in top-level positions (Gee et al., 2015). An examination of over 130,000 employees at five major technology organizations (e.g., Google, Hewlett-Packard, Intel, LinkedIn, and Yahoo) showed that Asian American men are significantly less likely to hold executive-level leadership positions compared to both white men and white women. Furthermore, fewer than three percent of board directors at Fortune 500 companies are Asian, black, and/or Hispanic women (Lim, 2017; Hill, 2016; Gee et al., 2015). Proposed reasons for the observed underrepresentation of certain social groups in leadership positions are varied. Abundant research has focused on examining the direct contributing factors of race-based discrimination and prejudice and finds that implicit stereotyping plays a role in many of the observed social disparities in the U.S. labor force (Quillian et al., 2017; Avery, 2011; Holzer, Offner, & Sorensen, 2005; Powell & Butterfield, 1997; Cancio, Evans, & Maume, 1996).

Maume (1999) discovered that while candidate race had no impact on the likelihood of becoming a manager for women, it had a negative effect for men. After controlling for position-relevant factors, black men were 52% less likely to attain a managerial position compared to their white counterparts. These findings are only one of many examples of how racial and gender prejudice produce different employment trajectories.
3. Face-Based Social Disparities in Employment and Hiring Selection

Research suggests that when race and ethnicity are placed aside, other factors such as physical characteristics—specifically facial structure—also play a large role in hiring decisions and in social disparities in employment (Antonakis & Eubanks, 2017; Todorov et al., 2005). As humans, we are strongly inclined to make heuristic decisions using facial cues. It is often the case that inferences about candidates’ potential leadership capabilities and competencies are based off their looks as opposed to objective information. This body of work shows that “face-ism”—the use of facial appearance for deriving inferences—matters considerably for leader selection. While on the one hand, face-ism can be seen as adaptive to the extent that using facial cues when other information is unavailable offers an efficient decision-making strategy, on the other hand, face-ism is not rational and can become consequential when it contributes to problems like face-based ageism and when it lends to greater attributions of job competency toward mature-faced individuals compared to their babyfaced peers.

Ageism pertains to stereotyping and discrimination against individuals or groups on the basis of their objective age as well as on physical characteristics (e.g., facial features) that are indicative of age. While age is not a reliable indicator when judging a worker’s potential productivity, employability, or leadership capability, research demonstrates that people nevertheless rely on appearance and facial cues such as facial maturity when making certain kinds of decisions including leadership ability (both in the political realm as well as the corporate context) in a manner that places certain candidates at a disadvantage (Re et al., 2013; Rule & Ambady, 2010; Rule et al., 2010; Livingston & Pearce, 2009; Rule & Ambady, 2009; Rule & Ambady, 2008; Antonakis & Dalgas, 2009; Banducci et al., 2008; Buckley, Collins, & Reidy,
2007; Little et al., 2007; Todorov et al., 2005; Martin, 1978; ). For example, leadership selection is heavily influence by perceived facial masculinity (e.g., sexual dimorphism in face shape, which correlates with perceived facial maturity) such that masculine and mature facial structures (characterized by smaller eyes, a higher forehead, lower eyebrows, wider pupil distance, a larger chin/wider jawline and a smaller facial width-to-height ratio) are preferred in leaders’ faces in times of intergroup conflict (e.g., war) whereas more feminine and babyfaced facial structures (e.g., bigger eyes, a lower forehead, higher eyebrows, smaller pupil distance, a smaller chin/more narrow jawline and a greater facial width-to-height ratio) are preferred in leaders’ faces during periods where within-group relationship maintenance or cooperation is emphasized (Spisak et al. 2012; Spisak et al., 2011; Little et al., 2007).

According to Zebrowitz and Montepare (2008), babyfaceness creates impressions of people’s traits. Babyfaced adults are perceived to have childlike traits—e.g., they are assumed to be naïve, submissive, weak, warm, and honest. Other traits inferred from babyfacedness include friendliness, health, openness, extroversion, emotional stability, confidence, intelligence, aggressiveness, dominance, and leadership ability (Cuddy et al., 2009; Albright et al., 1997; McArthur and Berry, 1987; Berry and McArthur, 1985).

Babyface stereotypes can bias social life outcomes, including elections, financial rewards, job applications, academic performance, prison sentences, and prosocial behavior (Zheng, 2016; Zebrowitz et al., 1998; Collins & Zebrowitz, 1995; Zebrowitz & McDonald, 1991). Prior research suggests that having a baby face is negatively correlated with success, particularly among white males in high positions of leadership in Western cultures, and may be a liability for those striving for positions of leadership in government and industry (Livingston & Pearce, 2009; Zebrowitz & Montepare, 2008). In most cases, babyfaced people of all ages and
sexes experience social outcomes that are consistent with the perceived traits associated with their appearance in so far as they are passed over for leadership positions and mentally-demanding tasks and, instead, are favored for jobs that require sociability.

The literature has identified a few notable exceptions to this effect, including a “disarming” effect of babyfacedness for black males that actually facilitates their perceived leadership capability by attenuating the salient, stereotypical perception of black males as threatening in a manner that allows them to be viewed as warm and competent and thereby successful in positions of leadership (Livingston & Pearce, 2009). Zheng et al. (2016) also identified potential cultural variables moderating the effect of babyfacedness on perceived leadership capability. Specifically, they found for male Chinese faces, both Chinese and Americans believed that babyfaces show less competence than mature faces; for female Chinese faces, however, whereas Chinese participants did not consider female babyfaces as less competent, American participants viewed them as less competent. The authors conclude that whereas the American participants’ inferences about both male and female Chinese babyface exemplars followed the same babyface generalization, Chinese participants’ inferences were likely informed by prior real-life experience and encounters of competent Chinese women who happen to have babyfaces in a culture where babyfaced females have the same social status and power (if not more) as their peers. Furthermore, they explain that, in Chinese culture, the babyface is a wise strategy to help Chinese people gain more resources: a round face shape, especially a babyface, can help an individual get limited resources in limited time, which may increase, rather than decrease their competence inferences.

In a study exploring how racial and gender stereotypes overlap to affect leadership selection, Galinsky, Hall, & Cuddy (2013) discovered that blacks were more likely and Asians
less likely than whites to be selected for a masculine leadership position due to implicit and explicit gendered racial stereotypes of Asians as more feminine and of blacks as more masculine compared to Whites. Studies have found that race affects the perceived sex categorization of faces such that participants can more easily categorize Asian female and black male faces compared to black female and Asian male faces, likely due to the fact that in the former cases the facial phenotypes associated with the racial categories are congruent with those associated with the sex categories whereas in the latter cases they are incongruent (Johnson, Freeman, & Pauker, 2011; Goff, Thomas, & Jackson, 2008). Galinsky, Hall, & Cuddy’s (2013) study demonstrates how the gender content of racial stereotypes, which correlates with facial maturity and babyfacedness, can have significant real-world implications and contribute to some of the aforementioned social disparities regarding representation in executive-level leadership.

4. Intergroup Empathy: An Empathy Gap Explains Observed Social Disparities Across Different Domains

Empathy refers to the capacity to understand and respond to the affective experience of another person. In addition to automatic tendencies to mimic the emotional expressions of other people, empathy also involves a role-taking ability to imagine what other people are thinking and feeling (Batson et al., 1997).

Across a variety of fields, researchers have found that empathy—once thought to be a purely automatic and reflexive process—is quite context-dependent and breaks down in the presence of outgroup members. It is often difficult for people to relate to social targets who they feel are dissimilar to themselves (Schumann, Zaki, & Dweck, 2014; Mitchell, Macrae, & Banaji,
Such work demonstrates an “empathy gap,” marked by a greater challenge to empathize with members of a racial outgroup, and the evidence for this exists at the behavioral, functional, and physiological levels (Cikara, Bruneau, & Saxe, 2011; Gutsell & Inzlicht, 2012; Hein et al., 2010).

Forgiarini, Gallucci, and Maravita (2011) found reduced reactivity to the pain of black actors compared to white actors that also correlated with the observers’ degree of implicit racial bias. In a similar vein, Trawalter, Hoffman and Waytz (2012) demonstrated how implicit biases in both white and black participants, medical professionals, and medical students may explain why black patients are undertreated for pain; across all these groups of people they found a racial empathy gap in the perception of pain marked by an assumption that black people feel less pain compared to white people. More specifically, participants interviewed in their study believed black people feel less pain as a result of having “thicker skin.” Assumptions about what it means to be African American—rooted in a stereotype of the African American experience as monolithic and involving low social status and hardship—seemed to instill an implicit belief that blacks have been “hardened” by racism and are therefore less sensitive to pain.

Many researchers have argued that the “empathy gap” explains many of the aforementioned harsher punishment decisions observed in educational contexts among youths of color compared to their white counterparts as well as the harsher criminal sentencing decisions and greater use of force inflicted on adult minorities compared to whites by the legal justice system. For example, white jurors are less likely to empathize with black defendants and victims, resulting in harsher punishments black defendants and more lenient sentencings for white defendants who harmed black victims (Dieter, 1998; Linder, 1996). Bruneau, Cikara, and Saxe (2017) note that empathic failures are common in hostile intergroup contexts. More specifically,
they found that parochial empathy—or the degree to which people felt good about in-group versus out-group fortunes and bad about in-group versus out-group misfortunes—predicted intergroup attitudes and behaviors (e.g., outgroup altruism).

Given how closely-tied to empathic failures many social disparities seem to be, and given that empathic concerns have been shown to predict prosocial behaviors, prejudice reduction techniques therefore should also target empathy induction.

4.1. Empathy and Socio-Cognitive Mindfulness

Research regarding the relationship between socio-cognitive mindfulness and empathy is sparse, particularly given how empathy seems to require hallmark elements of mindfulness such as attention (e.g., directing attention toward another) and flexibility (e.g., being cognitively and emotionally flexible regarding another’s internal state) (Gunaratana 2011; Block et al., 2007; Andersen, 2005). Furthermore, most of the research exploring the presumed connection between mindfulness and empathy has done so via correlational studies (Fan et al., 2011; Chiesa & Serretti, 2010; Greason & Cashwell, 2009; Beitel, Ferrer, & Cecero, 2005); whether mindfulness, in fact, cultivates empathy is less frequently investigated experimentally. Moreover, most of the research exploring these two topics has been confined to clinical contexts (Trent et al., 2016).

Some research has explored the relationship between socio-cognitive mindfulness and compassion, a construct very relevant to empathy and to prosocial behavior in so far as it involves feeling concern and care for others in response to their suffering and with the goal of improving their well-being (Germer & Barnhofer, 2017; Hilderbandt, McCall, & Singer, 2017; Goetz et al., 2010). While not necessarily looking at mindfulness from the socio-cognitive
approach, studies examining the impacts of more meditative mindfulness-based training programs, such as MBSR (Kabat-Zinn, 1982), have found an associated increase in compassion and prosocial behavior (Gu et al., 2015; Neff & Dahm, 2015; Keng et al. 2012, Salzberg, 2011; Birnie et al., 2010; Shapiro, Schwartz,& Bonner, 1998). In one study (Lim, Condon, & DeSteno, 2015), a three-week app-based mindfulness training involving 10-minute sessions was found to increase prosocial behavior (e.g., offering one’s seat to a person in need) compared to cognitive therapy. This same mindfulness training did not affect facial and vocal emotion recognition, however, leaving it up to interpretation whether prosocial behavior, alone, is an adequate proxy for empathy. Although intervention studies such as this one suggest potential empathy-enhancing effects of mindfulness, some studies lack appropriate control conditions (e.g., controlling for positive treatment expectancies) and others primarily rely on self-report questionnaires as empathy indices as opposed to behavioral measures.

Trent et al. (2016) found an association between trait socio-cognitive mindfulness and both affective and cognitive empathy. In this same study, meditative mindfulness only correlated with affective but not cognitive empathy. This supports the socio-cognitive mindfulness construct’s greater reliance on cognitive processes (e.g., perspective-taking; flexibility) compared to meditative mindfulness, and it has implications for the development of state socio-cognitive mindfulness interventions geared toward empathy induction in so far as perspective-taking and flexibility are processes that are relatively amenable to experimental manipulation.

This dissertation contributes to the lacking body of knowledge concerning the relationship between socio-cognitive mindfulness and empathy. Specifically, it experimentally investigates the relationship between socio-cognitive mindfulness and empathy induction.
5. The Malleability of Implicit Stereotypic Beliefs and its Implication for Socio-Cognitive Mindfulness

Social cognition research has identified several methods for inducing malleability and change in implicit social biases. This research reveals that implicit social group preferences can be shifted (albeit at varying degrees) via (1) exposure to counterstereotypical exemplars (Joy-Gaba & Nosek, 2010; Dasgupta & Rivera, 2008; Dasgupta & Asgari, 2004; Dasgupta & Greenwald, 2001; Blair, Ma, & Lenton, 2001), (2) intentional strategies to overcome and/or suppress bias (Cvencek et al., 2010; Kawakami et al., 2000), (3) evaluative conditioning (Kurdi & Banaji, 2019; BarAan, De Houwer, & Nosek, 2010; Olson & Fazio, 2006; De Houwer, Thomas, & Baeyens, 2001; Karpinski & Hilton, 2001; Olson & Fazio, 2001), (4) perspective-taking (Guinote, Willis, & Martellotta, 2010; Turner & Crisp, 2010; Richeson & Ambady, 2001; Finlay & Stephan, 2000), (5) increased social relations with outgroup members (Lowery et al., 2001; Sinclair & Kunda, 1999), (6) shifting focus of attention (Gilbert & Hixon, 1991; Macrae et al., 1999), (7) setting egalitarian goals (Legault et al., 2011; Stewart & Payne, 2008; Lun, Sinclair, Whitchurch, & Glenn, 2007; Gaertner et al., 1993), (8) inducing emotional states (Dasgupta et al., 2009; DeSteno et al., 2004), and (9) meditative mindfulness (Clobert, Saroglou, & Hwang, 2014; Kang, Gray, & Dovidio, 2014; Lueke & Gibson, 2014). Altogether, this line of work suggests that change in implicit prejudice can be attained via different mechanisms including retraining underlying associations, shifting the context of evaluation, and by controlling the activation or application of associations (Lai et al, 2013; Blair, 2002; Lai & Banaji, in press).
Some of the aforementioned prejudice reduction techniques present both strengths and challenges. Regarding those that involve exposure to counter associations (e.g., evaluative conditioning), for example, these often change implicit attitudes without affecting explicit attitudes. Olson and Fazio (2006) presented participants with positive images and words paired with Black faces and negative images and words paired with White faces and found that while exposure to these pairings reduced implicit racial bias quite quickly, explicit racial attitudes didn’t change. Furthermore, similar pairing paradigms and interventions involving exposure to counterstereotypical exemplars seem to “work” according to a compensatory and/or contrastive mechanism whereby outgroup members must be presented favorably (e.g., paired with positive content) at the cost of attitudes toward ingroup members (e.g., ingroup members are paired with negative content, and in some cases this lends to a negative shift in implicit attitudes toward ingroup members). One can see how the latter case is equally problematic; the ideal prejudice reduction technique should improve attitudes and behavior toward outgroup members without negatively impacting attitudes and behavior toward any other groups. Furthermore, interventions that rely on higher-level mental states and beliefs (e.g., perspectives-taking) for shifting bias are effective, however their effects are inconsistent as is the longevity of the resulting attitude shift (Lai & Banaji, in press).

This dissertation developed and assessed the effectiveness of state socio-cognitive mindfulness-based paradigms in reducing implicit racial and face-based bias, in increasing intergroup empathy, and in reducing frustration. Socio-cognitive mindfulness is stressed as a unique prejudice-reduction technique that, compared to the previously-mentioned approaches (including meditative mindfulness), is readily accessible (e.g., easy to administer), is economical (e.g., takes under five minutes and only requires eight stimulus photographs), and is adaptable,
demonstrating effectiveness at reducing implicit and explicit biases along different domains. The advantage of using a socio-cognitive-based mindfulness approach lies in its rapid approach of noticing novelty and sensitivity to context, which does not require the longer time commitment of meditative practices.

6. Socio-Cognitive Mindfulness

The concept of mindfulness originates from ancient Buddhist, Hindu, and Chinese philosophies. These more Eastern approaches to mindfulness are meditative in their nature and emphasize nonreactive awareness and concentration of one’s self and experiences (Thanissaro Bhikkhu, 2007). The role of such mindfulness practices is to keep the mind properly grounded in the present moment and to decrease reactivity to what happens in the moment. It is a way of relating to all experience—positive, negative, and neutral—such that overall levels of suffering are reduced, and sense of well-being increases (Germer, Siegel, & Fulton, 2005). Today, there are a variety of definitions of mindfulness within both Eastern and Western approaches. The Eastern approach to mindfulness has undergone several transformations following its introduction into Western culture and contemporary psychology.

traditions. This notion of mindfulness originated from a social psychological approach, and emphasizes active novel distinction-making.

Mindfulness has been shown to have a de-automizing effect that can impede the automatic effects of stereotype activation on behavior (Bargh & Chartrand, 1999). It has been shown to reduce stress, anxiety and promote resilience (Langer & Rodin, 1976; Langer, Janis, & Wolfer, 1975). Research has also shown that mindfulness can positively affect peoples’ lives by relying less on previously established associations (Lueke & Gibson, 2015; Langer, Bashner, & Chanowitz, 1985).

During the decision-making process, mindfulness techniques can assist in being proactive and identifying when a decision should be made such as by clarifying the objectives, generating options, avoiding irrational escalations, and avoiding commitments to previous bad decisions (e.g., sunk cost bias) (Hafenbrack, Kinias, & Barsade, 2013). Mindful individuals take more time before making decisions to reflect and listen to their inner selves and own values. Mindfulness can also help in identifying whether a decision is necessary. Heightened awareness allows more mindful persons to consider more relevant information to the decisions that are at hands. When choosing their course of action, more mindful individuals have a better ability to separate relevant from irrelevant information and are less likely to rely on stereotypes when deciding (Mallya, Huang, Fiocco, 2017; Djikic, Langer, & Stepleton, 2008).

Within the social scientific literature, there are two major branches of mindfulness research: meditative mindfulness (Ludwig & Kabat-Zinn, 2008; Kabat Zinn, 2003; Kabat-Zinn, 1982) and non-meditative, socio-cognitive mindfulness (Langer, 2009; Langer, 1977). Whereas meditative mindfulness is concerned with the awareness that arises through paying attention non-judgmentally to the present moment, socio-cognitive mindfulness is the process of drawing novel
distinctions which promotes greater attention to the present moment and sensitivity to context (Creswell, 2017; Brown, Ryan, & Creswell, 2007; Ie, Ngnoumen, & Langer, 2014).

Socio-cognitive mindfulness is a state of awareness marked by orientation in the present, openness to novelty, alertness to distinctions, sensitivity to different contexts, and awareness of multiple perspectives. It is brought about by actively noticing new things and attending to variability (Langer, 2009). These characteristics are represented in the major subfacets of the Langer Mindfulness Scale (LMS). This dissertation is concerned with understanding the link between socio-cognitive mindfulness and its potential role in mitigating implicit bias and stereotype-activated behaviors.

Socio-cognitive mindfulness is also importantly both a dispositional tendency and a state, which means that it can develop naturally as well as be induced externally. This quality is especially important and has applied research implications because it suggests that any observed moderating effects of trait socio-cognitive mindfulness in attenuating implicit bias can be exploited, such as via the introduction of a socio-cognitive state mindfulness induction.

A. Pilot Study Examining Relationship between Trait Socio-Cognitive Mindfulness and Implicit Racial Bias

The purpose of this pilot study was to validate the existence of a significant inverse relationship between trait socio-cognitive mindfulness and implicit racial bias prior to designing and testing the effectiveness of a state socio-cognitive mindfulness-based prejudice reduction technique. The value of assessing the relationship between trait socio-cognitive mindfulness and implicit bias in this way was that it would help inform which of the
construct’s four subfacets to emphasize during the design of the state intervention (e.g., based off the subfacets’ unique contributions to implicit racial bias scores).

A. 1. Participants

131 participants in this study were recruited through the Amazon Mechanical Turk (MTURK) participant pool and received $2.00 for partial or full completion of all study tasks, in accordance with Harvard University Institutional Review Board (IRB) policies. Participant ages ranged from 18 - 68 years, and the average was 36.29 years ($SD_{age} = 11.67$). The sample was 54% female. All participants were U.S. citizens and/or permanent residents. 74.4% of the sample were White. The remainder of the study sample was 8.5% African American, 8.5% Asian, .80% Latino, 1.6% Native American, and 6.2% Other/Unspecified.

A. 2. Measures

*Langer Mindfulness Scale (LMS).*

The Langer Mindfulness Scale is a 21-item questionnaire that assesses four domains associated with mindful thinking including novelty-seeking, novelty producing, engagement, and flexibility. An individual who seeks novelty perceives each situation as an opportunity to learn something new. A novelty-producing person generates new information to learn more about the current situation. An individual who scores high in engagement is likely to notice more details about his or her specific relationship with the environment. Flexible people welcome a changing environment rather than resist it (Pirson & Langer, 2015). On the LMS, participants are asked to rate the degree to which they agreed or disagree with statements pertaining to their own ability to be open to new ideas, make new distinctions, and take on different perspectives. Participants respond to items such as ‘I generate few novel
ideas.’ (reverse-scored) on a 7-point Likert scale (1 = strongly disagree; 7 = strongly agree). Higher scores indicate higher trait mindfulness. Previous research has reported that the LMS has robust validity ($\alpha = .81$) through correlations with theoretically relevant individual-difference constructs (e.g., the LMS is correlated with the tendency to entertain multiple perspectives).

**Black-White Implicit Association Test (picture IAT).**

The same Black-White Race IAT was used as in the Pilot Study (A). The Black-White IAT measures strengths of automatic associations. It consists of a computerized task that requires participants to rapidly categorize four target categories (e.g. good | bad attributes and black | white faces) such that faster responses to particular pairings (e.g., good-white; bad-black) are taken to reflect stronger implicit associations, whereas slower responses to pairings reflect weaker, more difficult associations (Greenwald, McGhee, & Schwartz, 1998). The particular categories used in this study include positive and negative attributes (e.g., marvelous-tragic; pleasant-horrible) and photographs of black and white faces. The IAT is traditionally scored by computing a $D$ score, which is the difference between the average response latencies between contrasted conditions divided by the standard deviation of response latencies across the conditions, akin to an effect size (e.g., Cohen’s $d$).
A. 3. Research Design

This correlational pilot study involved the completion of a questionnaire and a behavioral categorization task. The major dependent variables included participants total and subscale scores on the trait socio-cognitive mindfulness questionnaire (e.g., \textit{LMS Total Score}; \textit{LMS Novelty Production Total}; \textit{LMS Novelty Seeking Total}; \textit{LMS Engagement Total}; \textit{LMS Flexibility Total}) as well as their overall standardized score on the reaction time-based measure of implicit racial bias (e.g., \textit{IAT D}).

A. 4. Procedure

Participants from the Amazon Mechanical Turk online pool were informed that they would take part in a psychology study through a listing posted on the MTURK website. After providing electronic consent, all participants were redirected to two different third-party sites (\textit{Qualtrics}; \textit{Inquisit}) on which they completed a series of tasks including a trait measure of socio-cognitive mindfulness, a reaction time-based measure of implicit racial bias, and a short demographic questionnaire (e.g., age; sex; race/ethnicity; highest level of educational attainment; self-identified political affiliation).

A. 5. Data Analysis

To test the association between trait socio-cognitive mindfulness and implicit racial bias, LMS scores (total scores as well as four subfacet scores) and IAT D scores will be entered in a one-tailed bivariate Pearson Product-Moment correlation analysis.
A. 6. Results

The correlation analysis suggested that individuals higher on trait socio-cognitive mindfulness exhibit a weaker implicit preference for white Americans over black Americans, \( r (129) = -.210, p = .008 \). All major subfacets of the Langer Mindfulness Scale (LMS) significantly predicted implicit racial bias including novelty-production,\(^1\) novelty-seeking,\(^2\) engagement,\(^3\) and flexibility.\(^4\)

---

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Correlation Matrix of LMS Total Score, LMS Subscale Scores, and IAT Score (( N = 131 ))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LMS Total</td>
</tr>
<tr>
<td>LMS Total Score</td>
<td>(.917^{**})</td>
</tr>
<tr>
<td>LMS Novelty Production (NP)</td>
<td>(.724^{**})</td>
</tr>
<tr>
<td>LMS Novelty Seeking (NS)</td>
<td>(.442^{**})</td>
</tr>
<tr>
<td>LMS Engagement (E)</td>
<td>(.471^{**})</td>
</tr>
<tr>
<td>LMS Flexibility (F)</td>
<td>(</td>
</tr>
</tbody>
</table>

\(^{**} p < .01\)

\(^{*} p < .05\)

---

\(^1\) Generating new information in order to learn more about the current environment

\(^2\) Viewing every situation as an opportunity to learn something new

\(^3\) Noticing more details about one’s specific relationship with the environment

\(^4\) Welcoming a changing and malleable environment rather than resisting it
A. 7. Discussion

On the basis of these results, a state socio-cognitive mindfulness intervention was designed that emphasized novelty-production and engagement. More specifically, this state socio-cognitive mindfulness intervention was designed to shift attitudes and emotions at the group level (Chapter 3) as well as to shift frustration at the individual level (Chapter 4).
CHAPTER 3

USING STATE SOCIO-COGNITIVE MINDFULNESS INDUCTION TO SHIFT ATTITUDES AND EMOTIONS AT THE GROUP LEVEL

METHODS AND RESULTS

1. Experiment 1: Using State Mindfulness Induction to Reduce Implicit Racial Bias

The purpose of this first experiment was to determine whether a state socio-cognitive mindfulness induction, rooted in novelty-production (e.g., sorting according to novel, psychological categories as opposed to familiar, physical categories) and re-categorization, could reduce implicit racial bias.

1.1. Participants

An a priori power analysis indicated that this experiment would need a total of 180 participants (45 in each group) to have 80% power for detecting a medium-sized effect when employing the traditional .05 criterion of statistical significance. The medium-sized effect was based off prior research stemming from Langerian mindfulness that applied a similar paradigm, rooted in novelty production, toward reducing expectancy effects related to aging in the form of walking speed and obtained a Cohen’s $d$ of 0.325 (Djikic, Langer, & Stapleton, 2008). The decision to power Experiment 1 for a medium effect size was also based off studies finding a medium-sized correlation between trait socio-cognitive mindfulness and ageism. Lastly, the decision was informed by our pilot study’s discovery of an association between trait socio-
cognitive mindfulness and implicit racial bias which fell between a small and medium effect when assessed using a relatively small sample size.

200 participants were therefore recruited for Experiment 1 through the Amazon Mechanical Turk (MTURK) participant pool and received $2.00 for partial or full completion of all study tasks, in accordance with Harvard University Institutional Review Board (IRB) policies. Of the participants, 12 provided incomplete data and 4 presented as statistical outliers, leaving 184 participants in the final dataset. Participant ages ranged from 18 - 76 years, and the average was 39.81 years ($SD_{age} = 13.75$). The sample was 59.8% female. All participants were U.S. citizens and/or permanent residents. 66.3% of the sample were White. The remainder of the study sample was 9.2% African American, 2.7% Asian, .50% Latino, 5.4% Middle Eastern, .50% Native American, and 15.4% Other/Unspecified.

1.2. Measures

*State Mindfulness Induction Paradigm (Black & White).*

The state mindfulness induction paradigm was administered via a computerized task involving the sorting of faces according to categories. All participants viewed the same eight photographs on the screen that, unbeknownst to them, varied according to race. The photos came from the Eberhardt Face Database (Mind, Culture, & Society Laboratory at Stanford University, [http://www.stanford.edu/group/mcslab/cgi-bin/wordpress/examine-the-research/](http://www.stanford.edu/group/mcslab/cgi-bin/wordpress/examine-the-research/)) and were matched on facial attractiveness and affect based off naïve raters’ evaluations of the faces. The set included four photographs of white males and four photographs of black males. The eight photographs were randomly presented on the screen between trials. The four state mindfulness induction levels (and respective experimental conditions) are as follows:
Condition 4: High Mindful (Sorting by Four Psychological Categories)

Participants in the High state mindfulness induction condition were told that they would sort the picture set of eight photos four separate times according to a new category each time. They sorted the eight photos into two equal piles according to how perceptive | not perceptive and social | nonsocial they looked. Additionally, participants in this condition created two psychological dimensions by which to sort the faces. The first time they sorted, participants divided the photos into two equal piles on the screen. Photos were all recombined randomly prior to resorting by the second dimension. For the third and fourth rounds, the same process as before was repeated except using the psychological dimensions participants had created as the basis to sort. The resorting element was done so that there would be a visual aid to the mental concept of placing people in various categories and seeing that the same person can belong in different categories at different times. This condition was considered to be the most mindful as psychological dimensions are more novel in this context (compared to physical ones) and the task allowed participants to be actively engaged in the generation of the categories used for their sorting.

Condition 3: Moderate Mindful (Sorting by Four Physical Categories)

Participants in the Moderate state mindfulness induction condition sorted the same set of eight photos in the same way as those in Condition 4, except according to four categories all provided by the researcher and physical in nature (e.g., angular-shaped | round-shaped; short | tall; small | large; pale | tan). Participants in this condition did not have the elements of engagement/choice in the form of creating their own categories.

Condition 2: Low Mindful (Sorting by One Physical Category – Facial Maturity)
Participants in the Low state mindfulness induction condition sorted the same set of eight photos in the same way as those in Conditions 4 and 3, except according to the same, single physical category each time (e.g., babyfacedness | mature-facedness). This condition was not expected to increase mindfulness since participants were not asked to notice new things or to generate new categories about the individuals in the photos. They were merely sorting according to just one stereotypic dimension that prior research has found to be loosely related to race and to racial bias and therefore this condition was expected to not aid in the loosening of implicit racial associations compared to the previous two conditions.

Condition 1: Lowest Mindful (Sorting by One Physical Category – Physical Tone)
Participants in the Lowest state mindfulness induction condition sorted the same set of eight photos in the same way as those in Condition 2, except according to the same, single physical category of “pale | tan” each time. This condition was expected to promote the least amount of mindfulness due to its high degree of stereotypic activation related to race.

Across all conditions, the total number of sorting trials was constant with all participants sorting the same set of faces a total of four times. For example, participants in the lowest and lower conditions sorted faces on a single dimension--e.g., pale | tan or babyfaced | mature-faced--four times, yielding a total of four sorting rounds, while participants in the moderate and high condition sorted faces once on each of four different dimensions, yielding a total of four sorting rounds.

Black-White Implicit Association Test (picture IAT).

The same Black-White Race IAT was used as in the Pilot Study (A).
1.3. Research Design

The experiment used a 4-group between-subjects design. The independent variable was the four levels of *State Mindfulness Induction* (e.g., Lowest; Low; Moderate; High). The major dependent variable was the Black-White picture *IAT Score* for test trials. Scores from IAT test trials only were used (as opposed to practice trials or the average of practice and test trials) because IAT test trials usually provide the “truest” and cleanest (e.g., least variable) measure of people’s performance. IAT practice trials are often noisy in the sense that they are confounded by many factors including individuals’ degree of familiarity with the task and task demands (e.g., task instructions). The average of the practice and test trials can also be significantly affected by practice trials to the degree that a participant’s practice performance is significantly skewed as s/he gets acquainted with the task itself. For this reason, only the scores from practice trials were used as dependent measures across all experiments.

1.4. Procedures

Participants from the Amazon Mechanical Turk participant pool were informed that they would take part in a psychology study on person perception through a listing posted on the MTURK website. After providing electronic consent, participants were randomly assigned to one of four conditions each corresponding to one of four varying levels of state mindfulness induction (e.g. Lowest; Low; Moderate; High). Following the mindfulness induction, all participants completed the same Black-White picture IAT and a short demographic questionnaire (e.g., age; sex; race/ethnicity; highest level of educational attainment; self-identified political affiliation).
1.5. Data Analysis

The hypothesis tested in this experiment was that an increase in socio-cognitive mindfulness—operationalized as active categorization across increasingly novel dimensions, from physical to psychological—would lead to a decrease in the automatic activation of implicit associations related to stereotypes of Africans Americans, as measured by the Black-White IAT.

To test the directional hypothesis regarding a significant linear relationship between state mindfulness induction level and degree of implicit racial bias as measured by the IAT, a planned contrast analysis with weights of 3 (High state mindfulness induction condition), 1 (Moderate state mindfulness induction condition), -1 (Lower state mindfulness induction condition) and -3 (Lowest state mindfulness induction condition) was applied to the study conditions with IAT score as the major dependent variable. The linear contrast analysis was performed twice, once on the entire dataset and a second time on only those participants who self-identified as white given that they comprised a disproportionate amount of the dataset (e.g., 67.4%).

1.6. Results

As predicted by pilot research that demonstrated a significant linear relationship between higher trait mindfulness on the LMS and reduced implicit racial bias on the Black-White IAT, as well as by prior research on the impact of mindfulness on prejudice related to individuals with disabilities (Djikic, Langer, & Stapelton, 2008), a planned linear contrast analysis with weights of 3 (High state mindfulness induction), 1 (Moderate state mindfulness induction), -1 (Lower state mindfulness induction) and -3 (Lowest state mindfulness induction) revealed a significant linear relationship between the levels of state mindfulness induction and degree of implicit racial bias, $F(3, 180) = 4.22, p = .04, d = .20, 1-\beta = .57$. The High state mindfulness induction
condition, marked by both novelty (e.g., sorting according to psychological categories) and engagement (e.g., choice in the form of user-created categories for sorting) was associated with the greatest reduction in implicit racial bias on the IAT.

| Table 2 | Means and standard deviations for each condition |
|-----------------|-----------------|-----------------|-----------------|-----------------|
| Sorting according to: | 1 | 2 | 3 | 4 |
| 1 Physical Category | 1 Physical Category | 4 Physical Categories | 4 Psychological Categories |
| IAT d score (test trials) | .41 | .30 | .28 | .21 |
| $M$ | (.48) | (.42) | (.42) | (.41) |
| $SD$ | (N = 44) | (N = 53) | (N = 47) | (N = 40) |

Results Collapsed Across Participants of all Races:

Results for Self-Identified White Participants Only:

<table>
<thead>
<tr>
<th>IAT d score (test trials)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Physical Category</td>
<td>2 Physical Category</td>
<td>4 Physical Categories</td>
<td>4 Psychological Categories</td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>.44</td>
<td>.32</td>
<td>.29</td>
<td>.18</td>
</tr>
<tr>
<td>$SD$</td>
<td>(.49)</td>
<td>(.46)</td>
<td>(.35)</td>
<td>(.46)</td>
</tr>
<tr>
<td>(N = 30)</td>
<td>(N = 35)</td>
<td>(N = 28)</td>
<td>(N = 29)</td>
<td></td>
</tr>
</tbody>
</table>

+ Black-White IAT Mean Score for population = .45 - .52 (Nosek, Greenwald, & Banaji, 2005);

.30 - .33 (Charlesworth & Banaji, 2019)

None of the assessed demographic variables varied significantly with the major dependent variable. While a study with a research question such as this one ideally should control for the possible confounding variable of participant race prior to data collection (e.g., by limiting the sample to all white participants or to one constant racial group), policies involving
the recruitment of online participants through MTURK understandably prevent researchers from screening out workers on the basis of self-identified race. Race was therefore a variable that had to be isolated and controlled for only after data collection was fully complete. One obvious disadvantage to this was the significant loss in sample size and therefore power (e.g., total $N$ dropped to 122). When the same linear contrast analysis was performed to test the relationship among the four state mindfulness conditions on IAT scores for participants who self-identified as white only (to examine the effectiveness of the intervention at reducing implicit bias toward outgroup members according to self-identified race), the effect obtained was similar to the previous ($p = .036, d = .20$).

Total time participants spent on the state socio-cognitive mindfulness tasks did not significantly vary among the four groups. Across conditions, total number of sorting was controlled for. These results therefore cannot be explained in terms of task complexity (e.g., potentially greater task complexity within the highest mindfulness condition) for prior research would actually predict an opposite effect of task complexity on implicit bias, namely that greater task demands would yield greater implicit bias due to the introduction of stress, which tends to cause people to rely less on prefrontal executive control and more on habitual ways of processing (see Discussion).

2. Experiment 2: Using State Mindfulness Induction to Reduce Implicit Associations between Criminality and Race

The purpose of this second experiment was to determine whether the same state mindfulness induction that reduced a more general implicit racial bias against black Americans
in Experiment 1 would also prove effective in attenuating a more specific implicit association between criminality and black Americans. This question is pertinent considering prior work that suggests much of the observed racial disparities in law enforcement are rooted in a pervasive stereotype of African American males as hostile and in assumptions that they are more likely to engage in criminal activities.

2.1. Participants

An a priori power analysis indicated that this experiment would need a total of 100 participants (50 in each group) to have 80% power for detecting a medium-sized effect when employing the traditional .05 criterion of statistical significance.

150 participants in this experiment were recruited through the Amazon Mechanical Turk participant pool and received $2.00 for partial or full completion of all study tasks, in accordance with Harvard University IRB policies. Of the participants, 5 provided incomplete data, 4 demonstrated consistent responses on the IAT indicative of task inattention (e.g., responding randomly as quickly as possible to get through the task), and 3 were statistical outliers on the major dependent variable, leaving 138 participants in the final dataset. Participant ages ranged from 25 - 76 years, and the average was 39.95 years ($SD_{age} = 10.49$). The sample was 51.4% male. All participants were U.S. citizens and/or permanent residents. 75.3% of the sample were White. The remainder of the study sample was 8% African American, 8% Asian, 5.9% Latino, 1.4% Other/Unspecified, 0.7% Middle Eastern, and 0.7% Other (e.g., Biracial).

2.2. Measures
State Mindfulness Induction Paradigm (Black & White).

The state mindfulness induction paradigm was similar to the one used in Experiment 1, however only the High and Lowest conditions were used in Experiment 2.

Race-Weapons Implicit Association Test (IAT).

The Race-Weapons IAT requires participants to rapidly categorize four target categories including images of black and white Americans presented alongside weapons and non-weapons. On this particular IAT, most participants show stronger associations between black Americans and weapons and between whites Americans and harmless objects compared to the reverse pairing. In fact, Nosek et al.’s (2007) summary of 85,742 IAT scores on the Race-Weapons IAT suggests 72% of the sample showed the dominant effect, whereas 9% showed the reverse association. This effect was also relatively strong among black participants in the study \( D = 0.59 \).

2.3. Research Design

The experiment used a 2-group between-subjects design. The independent variable was the two levels of State Mindfulness Induction (e.g., Lowest and High). The major dependent variable was the Race-Weapons IAT Score for test trials.

2.4. Procedures

Participants from the Amazon Mechanical Turk participant pool were informed that they would take part in a psychology study on person perception through a listing posted on the
MTURK website. After providing electronic consent, participants were randomly assigned to one of two conditions each corresponding to the same High and Lowest levels of state mindfulness induction as in Experiment 1. Following the mindfulness induction, all participants completed the same Race-Weapons IAT and a short demographic questionnaire (e.g., age; sex; race/ethnicity; highest level of educational attainment; self-identified political affiliation).

2.5. Data Analysis

The hypothesis tested in this experiment was that an increase in socio-cognitive mindfulness—operationalized as active categorization across increasingly novel dimensions, from physical to psychological—would lead to a decrease in the automatic activation of implicit associations related to Africans Americans and criminality, as measured by the Race-Weapons IAT.

To test whether state mindfulness induction influenced implicit associations between race and criminality, a one-tailed independent samples t-test was performed on IAT scores between the High and Lowest state mindfulness conditions.

2.6. Results

Comparisons of participants in the High and Lowest state mindfulness induction conditions revealed lower implicit associations between African Americans and criminality on the Race-Weapons IAT in the former group compared to the latter, $t(136) = 2.06, p = .003$, Cohen’s $d = .255$, $1 − \beta = .58$. 
None of the assessed demographic variables varied significantly with the major dependent variable. Total time participants spent on the mindfulness tasks did, however, significantly vary between groups. This latter finding will be elaborated upon in the discussion section, particularly the open empirical question regarding whether it is reflective of cognitive effort, motivation to engage, or mindless disengagement. Task complexity was equalized between the groups in the form of equal number of sorting experiences, therefore observed differences in completion times between conditions could be a function of cognitive effort or of the phenomenon. Importantly, total time spent on the mindfulness tasks did not predict performance on the major dependent variable, IAT score, which suggests it does not mediate the observed relationship between High state mindfulness induction and reduced implicit bias.

<table>
<thead>
<tr>
<th>Sorting according to:</th>
<th>Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 Low Mindful (N = 70)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>IAT d score (test trials)</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>.40</td>
<td>.21</td>
</tr>
<tr>
<td>SD</td>
<td>(.42)</td>
<td>(.32)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Total Time on M+ Task (s)</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>208.01</td>
<td>450.22</td>
</tr>
<tr>
<td>SD</td>
<td>(145.64)</td>
<td>(293.06)</td>
</tr>
</tbody>
</table>

+ Race-Weapons IAT Mean Score for population = .37 (Nosek et al., 2007b)
3. Experiment 3: Using State Mindfulness Induction to Reduce Implicit Face-Based Bias

The purpose of this experiment was to test the adaptability of the state mindfulness induction paradigm used in Experiments 1 and 2. Experiment 3 sought to determine whether a similar state mindfulness induction could be adapted and applied toward shifting implicit face-based attitudes related to facial maturity. This work has implications for face-based discrimination in the workforce related to implicit associations between babyfacedness and perceived incompetence and/or leadership incapability that informs job candidate selections as well as promotion decisions.

3.1. Participants

An a priori power analysis indicated that this experiment would need a total of 100 participants (50 in each group) to have 80% power for detecting a medium-sized effect when employing the traditional .05 criterion of statistical significance.

100 participants in this experiment were recruited through the Amazon Mechanical Turk participant pool and received $2.00 for partial or full completion of all study tasks, in accordance with Harvard University IRB policies. Ages ranged from 23 - 70 years, and the average was 35.60 years ($D_{age} = 8.49$). The sample was 60% male. All participants were U.S. citizens and/or permanent residents. 70% of the sample were White. The remainder of the study sample was 12% Asian, 10% Latino, 7% African American, and 1% of Middle Eastern origin.

3.2. Measures

*State Mindfulness Induction Paradigm (Facial Maturity).*
The state mindfulness induction paradigm was similar to the one used in Experiments 1 and 2, however stimuli involved eight computerized male faces varying on facial babyfacedness and facial maturity. The photos came from the Todorov Face Database (Todorov et al., 2013; Todorov & Oosterhoof, 2011) and were matched on facial attractiveness and affect. The eight photographs were randomly presented on the screen between trials. Similar to Experiment 2, only the High and Lowest conditions were used.

Face Evaluation Task.
In a computerized task, participants viewed a set of eight novel faces that, unbeknownst to them, varied according to facial babyfacedness and facial maturity. Four photographs were of babyfaced males and four photographs were of mature-faced males. All faces belonged to white males of matching attractiveness and affect, obtained from the Karolinska Directed Emotional Faces (KDEF) Database (Lundqvist, Flykt, & Öhman, 1998). (Perceived age was difficult to control when using real faces in light of its close relation to perceived facial maturity and babyfacedness.) Participants were asked to select, from among the faces, those which seemed likely to belong to professions or positions perceived as high competent and high dominant, including (1) a finance officer (2) a doctor, (3) a police officer, and (4) a physics teacher. Participants were told that there were no requirements for how to select the faces, and they could select anywhere between one and eight (e.g., all) per professional category.

3.3. Research Design
The experiment used a 2-group between-subjects design. The independent variable was the two levels of State Mindfulness Induction (e.g., Lowest and High). The major dependent
variables included four *Face Evaluations* regarding the perceived likelihood of eight novel babyfaced- and mature-faced individuals to serve four different kinds of high-competent and high-dominant roles (e.g., average number of babyfaced individuals included in each of the four professional categories; average number of mature-faced individuals included in each of the four professional categories).

3.4. Procedures

In order to prevent the same Amazon Mechanical Turk (MTurk) workers from Experiment 1 from participating in Experiment 2 (and more generally to protect the contamination of data resulting from the same participants and their prior knowledge from influencing subsequent experiments within this dissertation), all online experiments in this dissertation were posted by the same MTurk requester account and redirected to the same Qualtrics survey in order to (1) save MTurk workers’ IDs and create a new qualification based off these IDs in order to only allow “new” worker IDs to be eligible for participation in future experiments, and (2) to use Qualtrics’ ballot-box stuffing prevention feature which prevents the same IP address from submitting more than one response to the same survey.

Participants from the Amazon Mechanical Turk participant pool were informed that they would take part in a psychology study on person perception through a listing posted on the MTURK website. After providing electronic consent, participants were randomly assigned to one of two conditions each corresponding to the same High and Lowest levels of state mindfulness induction as in Experiment 2. Following the mindfulness induction, all participants completed the same face evaluation task and a short demographic questionnaire (e.g., age; sex; race/ethnicity; highest level of educational attainment; self-identified political affiliation).
3.5. Data Analysis

The hypothesis tested in this experiment was that an increase in socio-cognitive mindfulness—operationalized as active categorization across increasingly novel dimensions, from physical to psychological—would lead to a decrease in the automatic activation of implicit associations related to babyfacedness and incompetence, as measured by participants’ judgments of novel babyfaced and maturefaced individuals’ likelihoods of belonging to various leadership positions post-treatment.

To determine whether the high mindfulness induction reorganized the structure of participants’ representations of high-leadership careers, the frequency of babyfaced individuals included in participants’ responses (e.g., in their conceptualizations of high-competence and high-dominant professions) was compared between groups using independent samples t-tests with Bonferroni correction to control for Type I error inflation due to multiple comparisons. We similarly compared the frequency of mature-faced individuals included in participants’ responses between groups. Both analyses would allow us to see, for example, whether the mindfulness manipulation was successful in shifting people’s conceptualizations by increasing their willingness to see babyfaced individuals as competent and as leaders, by decreasing their willingness to assume mature-faced individuals are more competent leaders, or both. We also compared the total number of responses provided by participants between groups since no guidelines (e.g., minimum or limit) were given.

3.6. Results
The effect of state mindfulness induction on face-based biases related to facial maturity was originally going to be tested through a between-groups comparison of face evaluations using multiple independent samples $t$-tests with Bonferroni correction for Type I error inflation. Prior to running these analyses, however, an examination of the major dependent variables’ distributions revealed both non-normality and positive skew thereby violating the assumptions of parametric two-sample $t$-tests. The non-parametric Mann-Whitney test was therefore used to test this hypothesis because it does not rely on distributional assumptions.

A Mann-Whitney $U$ test with Bonferroni correction showed that there was a significant difference between High and Low state mindfulness conditions in terms of participants’ average selection of babyfaced individuals for high-dominant and high-competent roles. Specifically, participants in the High state mindfulness condition were more likely to view babyfaced individuals as financial officers ($U = 1595, p = .004, d = .54$), police officers ($U = 1575, p = .005, d = .41$), doctors ($U = 1563, p = .020, d = .60$), and physics teachers ($U = 1495, p = .020, d = .35$) compared to those in the Low state mindfulness condition. After the Bonferroni adjusted $p$ value of .013 is applied, only the effect pertaining to evaluations of financial officers and police officers remain significant.

Importantly, there was no difference in the total number of faces selected between the two groups (e.g., per profession, the minimum number of faces selected was 1 and the maximum was 8, with an average of 2.79 faces selected per professional category/evaluation), nor in the number of mature-faced individuals selected (out of a possible maximum of 4 mature faces: $M_{\text{LowM}} = 2.36, SD_{\text{LowM}} = 1.12; M_{\text{HighM}} = 2.48, SD_{\text{HighM}} = .97$). In general, babyfaced individuals were more likely to be selected by participants in the High mindful condition.
Another way of understanding these results is that, overall, individuals from both the Low and High mindful condition were still more inclined to choose mature-faced candidates (e.g., up to about 3 on average) for the various high-leadership professional categories, however those in the High mindful condition were more apt to have one of their three candidates be a mature-faced person. Importantly, it wasn’t the same face that was chosen every time. None of the assessed demographic variables varied significantly with the major dependent variable.

<table>
<thead>
<tr>
<th>Sorting according to:</th>
<th>Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 Low Mindful (N = 52)</td>
</tr>
<tr>
<td>Average # of Baby Faces Included in Representations of Finance Officers</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>SD</td>
</tr>
<tr>
<td>Average # of Baby Faces Included in Representations of Police Officers</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>SD</td>
</tr>
<tr>
<td>Average # of Baby Faces Included in Representations of Doctors</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>SD</td>
</tr>
<tr>
<td>Average # of Baby Faces Included in Representations of Physics Teachers</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>SD</td>
</tr>
</tbody>
</table>
4. Experiment 4: Comparing Effects of Process and Content in Mindfulness-Based Paradigm in Reducing Implicit Racial Bias

The purpose of this experiment was two-fold. As its first aim, Experiment 4 sought to identify the driving force behind the effects observed in the previous three studies. In Experiments 1 – 3, the High Mindful conditions, which predicted the least amount of racial and face-based bias, involved sorting according to both psychological categories (2) as well as self-generated categories (2). Novelty (operationalized through the introduction of psychological as opposed to physical categories) and engagement (operationalized by having participants choose their own sorting categories) were therefore confounded components within the same condition, and it remained unclear which, if either, was the dominant driving force of the effect.

Designing an experiment that isolates and tests the effect of engagement in this context is challenging because it would introduce a lot of variability within and between participants (e.g., it would involve participants sorting according to different categories within the same experimental conditions). It was therefore decided that designing an experiment to test the effect of novelty (e.g., psychological categories) alone, as it compares to familiarity (e.g., physical categories) was an easier start. If a state mindfulness-based paradigm involving four [constant] psychological categories significantly reduced implicit racial bias compared to a state mindfulness-based paradigm involving four [constant] physical categories, then this might suggest that novelty was both necessary and sufficient to the observed effects from the High Mindful conditions in Experiments 1 – 3. Alternatively, if in Experiment 4 the paradigm involving four psychological categories did not reduce implicit bias compared to the paradigm
involving four physical categories, then this might suggest that engagement was necessary to the observed effects from the High Mindful conditions in Experiments 1 – 3.

As its second aim, Experiment 4 examined the degree to which the effectiveness of the state mindfulness-based paradigm in Experiments 1 and 2 requires close similarity between the content of the intervention and the target of the intervention. In other words, it remained an open empirical question whether a similar state mindfulness-based paradigm involving stimuli representing members of different racial group categories (e.g., Latino) would produce a similar shift in implicit racial bias pertaining to African Americans.

4.1. Participants

An a priori power analysis indicated that the study’s first aim would require a sample size of 180 (45 in each of four groups) to have 80% power for detecting a medium-sized effect at \( p < 0.05 \). The study’s second aim would require a sample size of 100 (50 in each of two groups).

160 participants in this experiment were recruited via MTURK’s website and each received $2.00 for partial or full completion of all study tasks, in accordance with Harvard University IRB policies. Ages ranged from 19 - 72 years, and the average was 41.01 years (\( SD_{\text{age}} = 12.52 \)). The sample was 53.7% male. 78.8% of the sample were White. The remainder of the study sample was 8.1% African American, 6.9% Latino, 4.4% Asian, and 1.9% Other.

4.2. Measures

*State Mindfulness Induction Paradigm (Black & White).*

Four similar levels of state mindfulness (e.g. conditions) were used as in Experiment 1 with some important added distinctions: a High mindful group sorted the same eight photos (four
black and four white males) according to four predetermined psychological categories, a Moderate group sorted the eight photos according to four physical categories, a Lowest group sorted according to one psychological category, and a Lowest group sorted according to one physical category. The major difference between this paradigm and the one used in Experiment 1 is the removal of the element of engagement, previously operationalized in the High mindful group by having participants create two of the four psychological categories used in their sorting task.

*State Mindfulness Induction Paradigm (Latino & White).*

This paradigm is very similar to the one just described except the stimuli involves Latino and white photos in the place of black and white photos.

The state mindfulness induction paradigm was administered via a computerized task involving the sorting of faces according to categories. All participants viewed the same eight photographs on the screen that, unbeknownst to them, varied according to race. The photos came from the Chicago Face Database (CFD; Ma, Correll, & Wittenbrink, 2015) and were matched on facial attractiveness and affect based off naïve raters’ evaluations of the faces. The set included four photographs of white males and four photographs of Latino males. The eight photographs were randomly presented on the screen between trials.

*Black-White Race Implicit Association Test (IAT).*

The same Black-White Race IAT was used as in Experiments 1 and 2.

4.3. Research Design
This experiment involved a 2 x 4 between-subjects design with two independent variables of varying levels. The first independent variable was Content (e.g., state mindfulness induction photo content) and had two levels, including black-white photos and Latino-white photos. The second independent variable was Process (e.g., degree of manipulated novelty in state mindfulness induction) and had four levels, including 4 psychological, 4 physical, 1 psychological, and 1 physical. The major dependent variable was the Black-White picture IAT Score.

4.4. Procedures

Participants from the MTURK pool were informed that they would take part in a study on person perception through a posted HIT on the MTURK website. After providing electronic consent, participants were randomly assigned to one of eight conditions involving either black-white or Latino-white content and sorting it in four different ways or according to one of four different processes. Following the state mindfulness induction, all participants completed the same Black-White IAT and a short demographic questionnaire (e.g., age; sex; race/ethnicity; highest level of educational attainment; self-identified political affiliation).

4.5. Data Analysis

Two separate analyses were performed to explore the two separate aims of the study related to process and content. The major dependent variable’s (e.g., IAT score) relationship to the two major independent variables, Process and Content, was explored separately due to the fact that the original research design did not involve a full factorial model (e.g., there was no
interest in assessing interactions between the two independent variables as they pertained to the dependent variable in this context).

To address the question of whether engagement was necessary to the effects observed in Experiments 1 – 3, and whether the novelty of sorting by psychological categories (compared to sorting by physical categories) is, alone, a strong predictor of reduced implicit racial bias, the four levels of the Process independent variable were entered in a planned linear contrast analysis with weights of 3 (High mindful—4 psychological categories), 1 (Moderate mindful—4 physical categories), -1 (Low mindful—1 psychological category) and -3 (Lowest mindful—1 physical category). Black-White Race IAT score was the major dependent variable. A linear planned contrast analysis was used due to the step-wise relationship among the experimental conditions assumed a priori, similar to Experiment 1.

To determine whether the effect of state mindfulness in reducing implicit racial bias toward African Americans depends on the content of the intervention, Black-White Race IAT scores were compared between individuals who completed state mindfulness inductions featuring black and white faces and those who completed the same inductions featuring Latino and white faces using an independent samples t-test.

4.6. Results

Regarding the question of Process, the linear planned contrast analysis with weights of 3, 1, -1, and -3, testing the degree to which four levels of state mindfulness induction predicted implicit racial bias reduction, was significant, $F (3,156) = 5.14, p = .025, d = .20, 1−\beta = .40$. As predicted, the linear relationship among the four state mindfulness induction process levels (e.g., High—4 psychological; Moderate—4 physical; Low—1 psychological; Lowest—1 physical)
4.

Post-hoc one-tailed comparisons of the High and Moderate levels only showed that sorting by four psychological categories significantly reduced implicit racial bias compared to sorting by four physical categories, supporting one of the hypotheses regarding the driving force behind the observed effects in Experiments 1–3, $t(78) = 1.62, p = .05, \text{Cohen's } d = .40, 1-\beta = .55.$

<table>
<thead>
<tr>
<th>Table 5</th>
<th>Means and standard deviations for each condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sorting according to:</td>
<td>Condition</td>
</tr>
<tr>
<td>1</td>
<td>1 Physical Category</td>
</tr>
<tr>
<td>M</td>
<td>.25</td>
</tr>
<tr>
<td>SD</td>
<td>(.29)</td>
</tr>
<tr>
<td>(N = 40)</td>
<td>(N = 40)</td>
</tr>
</tbody>
</table>

+ Black-White IAT Mean Score for population = .45 - .52 (Nosek, Greenwald, & Banaji, 2005);

.30 - .33 (Charlesworth & Banaji, 2019)

The two-tailed independent samples $t$-test comparing IAT $d$ scores on the Black-White race IAT between participants who completed state mindfulness inductions featuring black and white content to those who completed the same paradigms featuring Latino and white content was not significant, $t (158) =-1.34, p = .18, \text{Cohen's } d = .23, 1-\beta = .53.$
Experiment 5: Using State Mindfulness Induction to Increase Intergroup Empathy

5.1. Participants

An a priori power analysis indicated that this experiment would need a total of 180 participants (45 in each group) to have 80% power for detecting a medium-sized effect when employing the traditional .05 criterion of statistical significance.

180 participants in this experiment were recruited via MTURK’s website and each received $3.60 for partial or full completion of all study tasks, in accordance with Harvard University IRB policies. Twelve participants self-identified as male on the demographic survey despite the study recruitment on MTURK specifying that participants was limited to self-identified females. Another nine participants provided incomplete data or did not finish the study. These responses were therefore not included in the final dataset of 159 participants. While MTURK allows participants to know that listed studies or “Hits” have specific criteria (e.g., “females only”) and can make studies visible to only those relevant populations, it still cannot avoid the issue of some participants creating false accounts. Furthermore, in accordance with
MTURKs’ policies, self-identified racial background is not a permissible screening criteria, and therefore even though this study wanted to control for both participant sex and race, both would have to be controlled for following data collection based on participants’ responses to the demographic questionnaire.

Participant ages ranged from 20 - 72 years, and the average was 41.04 years ($SD_{age} = 41.01$). In the final dataset of 159 participants, 100% were female (study intentionally recruited self-identified females to ensure that all participants viewed the target of the empathic listening task, Natasha, as a member of their gender ingroup). 82.4% of the sample were White. The remainder of the study sample was 10.1% African American, 3.1% Asian, 2.5% Latino, and 1.9% Other.

5.2. Measures

*State Mindfulness Induction Paradigm (Black, Latino & White).*

This paradigm is very similar to the ones from Experiments 1 – 4, except the stimuli involve photos of black, Latino, and white males. All three racial groups were used (instead of only black and white) to mask the study’s true focus on perceptions of black Americans as outgroup members and in order to align with the study’s purported aim of assessing personality as it relates to visual and auditory perception more broadly.

The state mindfulness induction paradigm was administered via a computerized task involving the sorting of faces according to categories. All participants viewed the same eight photographs on the screen that, unbeknownst to them, varied according to race. The photos came from the Eberhardt and Chicago Face Databases (Ma, Correll, & Wittenbrink, 2015) and were matched on facial attractiveness and affect based off naïve raters’ evaluations of the faces. The
set included three photographs of white males, three photographs of black males, and two photographs of Latino males. The ratios were designed this way to ensure that black and white photos still comprised a greater proportion of the set in light of Experiment 4’s results, which showed that prejudice reduction was most predicted by the closeness in match between the content and target of the intervention used. We therefore did not want a larger presence of Latino faces to counteract or occlude a potential effect on prejudice reduction targeting anti-black bias and the white-black empathy gap (although it might still be effective toward reducing a self-other empathy gap more generally). The eight photographs were randomly presented on the screen between trials.

*Empathic Listening Task.*

The empathic listening task, borrowed from Schumann, Zaki, & Dweck (2014), is a behavioral measure of empathic effort that assesses the time spent listening to an empathy target. In this task, after seeing the picture of either a black or white female named Natasha, participants listen to a 10-minute recording of Natasha’s experience with her grandmother’s battle with cancer. Participants have the option to fast-forward through the recording, as desired, by dragging the progress line on the audio controller. A script on the webpage records the total number of seconds spent listening to the recording and is used as a major measure of empathic effort. After listening to the audio recording, participants answered four factual questions about the contents of the recording (e.g., “What is something Natasha learned from her grandmother?”)

*Other and Self Emotion Evaluations.*
After listening to the audio recording of Natasha’s experience of her grandmother’s battle with cancer participants rated the degree to which Natasha experienced 11 different emotions (e.g., anger; disgust; excitement; fear) on a 7-pt horizontal slider scale. After providing ratings for Natasha, they also rated the degree to which they themselves experienced these same 11 emotions in listening to her story. Participants’ ratings (7-pt scale) for themselves were presented in a different order and using a visually different scale (e.g., vertical slider scale) to prevent learning and response bias. A difference score between ratings for self and for other (e.g., Natasha) will be calculated to form a measure of affective empathy. If empathy is defined as the ability to share and understand the feelings of another, then by this definition a difference score between emotion evaluations of Natasha (e.g., Other) and self were constructed to determine whether the state socio-cognitive mindfulness intervention would predict how aligned or misaligned emotional evaluations of Natasha’s story would be as a function of Natasha’s race. This way of operationalizing an “empathy gap” and of reflecting racial bias was borrowed from Trawalter, Hoffman, and Waytz’s (2012) work on perceptions of others’ pain in which they compared scores between participants’ anticipated pain evaluations for self and for another person (either a black or white target).

Interpersonal Reactivity Index (Empathic Concern).

The Interpersonal Reactivity Index (Davis, 1983) is a self-report individual differences (e.g., trait) measure of empathy. The original 28-item measure is answered on a 5-pt Likert scale and contains four subscales including perspective taking, fantasy, empathic concern, and personal distress. For this study, we were concerned with only the Empathic Concern subscale (IRI-EC)
given its association to "other-oriented" feelings of sympathy and concern for unfortunate others.

*Langer Mindfulness Scale (LMS).*

The same LMS was used as in the Pilot Study (A).

5.3. Research Design

This experiment involved a 2 x 2 between-subjects design with two independent variables of varying levels. The first independent variable was *State Mindfulness Level* (High | Low). The second independent variable was *Target Race* (White | Black). The major dependent variables included the total amount of time spent listening to Natasha’s story (*Empathic Effort*) and the average difference in emotion evaluations between self and other (*Affective Empathy*). A difference score was calculated between participants’ evaluations of Natasha’s emotional experiences and their evaluations of their own experiences to produce a measure of *Affective Empathy*, and this score was also entered in a similar univariate ANOVA analysis to determine the degree to which *State Mindfulness Level* and *Target Race* predict affective empathy both individually (main effects) and jointly (interactions).

5.4. Procedures

Participants from the MTURK pool were informed that they would take part in a study aimed at understanding how personality is related to visual and auditory perception (cover story). They were told they would answer some questions about themselves, sort some photos, and listen to a story via an audio recording. After providing electronic consent, participants were
randomly assigned to one of two conditions involving Low or High state mindfulness induction. Following the state mindfulness inductions, all participants completed the same Black-White IAT and a short demographic questionnaire (e.g., age; sex; race/ethnicity; highest level of educational attainment; self-identified political affiliation).

5.5. Data Analysis

All major analyses will be limited to participants who self-identify as both female and white only in order to ensure that all participants in the study view the white empathy target (e.g., “White Natasha”) as an ingroup member and the Black empathy target (e.g., “Black Natasha”) as an outgroup member.

The hypothesis tested in this experiment was that an increase in socio-cognitive mindfulness—operationalized as active categorization across increasingly novel dimensions, from physical to psychological—would lead to an increase in empathic effort (e.g., longer listening time) and an increase in affective empathy (e.g., smaller difference between evaluations of Natasha’s and self’s emotional experiences in response to the story) more generally (main effect of state mindfulness). An interaction between State Mindfulness Level and Target Race is also expected that will enable more empathic responding toward Black Natasha in the High mindful group compared to the Low mindful control group.

Demographic variables, trait mindfulness (LMS), and trait empathy (IRI-EC) will be assessed in relation to the major dependent variables, Total Listening Time to Outgroup Member’s Emotional Story (empathic effort) and Affective Empathy, via bivariate correlation analyses and independent samples t-tests (e.g., sex). Any variables that significantly predict the dependent variable will be controlled for in major analyses. A univariate ANOVA will be
performed with *Total Listening Time* as the dependent variable and *State Mindfulness Condition* (High | Low) and *Target Race* (White | Black) as two independent variables to assess main effects as well as interaction effects. A similar univariate ANOVA will be performed for the *Affective Empathy dependent* variable.

5.6. Results

The hypothesis that state mindfulness induction would predict interpersonal empathic effort (e.g. time spent listening to an outgroup member’s emotional story) was tested using a univariate ANOVA in order to look at interaction effects. When looking at only self-identified white females, the 2 (High Mindful | Low Mindful) x 2 (White Natasha | Black Natasha) univariate ANOVA yielded a significant interaction, $F(1, 126) = 4.52, p = .036$, Cohen’s $d = .23, 1−β = .52$. Post-hoc examination of simple effects suggests the interaction effect is primarily driven by a significant difference in average time spent listening to Black Natasha’s emotional story between the High and Low mindful groups, $p = .024$. Post-hoc analyses also revealed that participants in the Low mindful condition spent significantly less time listening to Black Natasha compared to White Natasha, $p = .032$. 
When the difference scores between participants’ evaluations of Natasha’s emotions and their own were averaged across all eleven emotions by participant and entered in a univariate ANOVA as a dependent variable with *State Mindfulness Level* and *Target Race* as the independent variables, state socio-cognitive mindfulness predicted affective empathy, $F(1, 128) = 4.12, p = .045$, Cohen’s $d = .37$, $1−β = .65$. The average difference scores were smaller for individuals in the High mindful condition compared to those in the Low, suggesting greater affective empathy as a function of the intervention.
As expected, trait socio-cognitive mindfulness on the LMS and trait empathy on the IRI-EC were significantly related, $r(128) = .484, p = .000$. There was no difference in state mindfulness task completion time between High and Low conditions. None of the assessed demographic variables varied significantly with the major dependent variable.
CHAPTER 4

USING STATE SOCIO-COGNITIVE MINDFULNESS INDUCTION TO SHIFT EMOTIONS AT THE INDIVIDUAL LEVEL

METHODS AND RESULTS

6. Experiment 6: Using State Mindfulness Induction to Decrease Frustration, I

The purpose of this experiment was to determine whether state mindfulness induction could prove effective in attenuating aggression.

6.1 Participants

An a priori power analysis indicated that this experiment would need a total of 100 participants (50 in each group) to have 80% power for detecting a medium-sized effect when employing the traditional .05 criterion of statistical significance.

100 participants in this experiment were recruited from the Harvard Psychology Department participant pool and received 1.00 course credit for partial or full completion of all study tasks, in accordance with Harvard University IRB policies. When asked at the end of the study whether they suspected anything, four participants mentioned believing that the learning task was rigged and that the Photo Selection Task was related to the first part of the study. Two participants provided incomplete data due to a glitch on the laptop computer used to administer the tasks on one of the days the study was running, another four did not complete the second half.
of the study involving the major dependent variable (e.g., did not choose to engage in a “quick, second” study), and one participant completed the study tasks in the incorrect order. Responses from these participants were therefore removed from the final dataset of 89.

Participant ages ranged from 18 - 24 years, and the average was 21.21 years ($SD_{age} = 4.18$). The sample was 66% female. 32% of the sample were White. The remainder of the study sample was 21.6% Asian, 9.3% African American, 6.2% Latino, 4.1% Other, and 3.1% Native American.

6.2. Measures

*Big Five Inventory (BFI-44).*

The BFI-44 is a 44-item inventory that measures an individual on the Big Five Factors (dimensions) of personality including extraversion, agreeableness, conscientious, neuroticism, and openness (Goldberg, 1993). Each of the factors is then further divided into personality facets.

*Frustration Induction Paradigm.*

A modified version of the Wisconsin Card Sorting Test (WCST; Berg, 1948) was designed and validated as a frustration induction task. In line with the original WCST, participants are presented with stimulus cards of varying color (e.g., red, blue, yellow, green), shapes (e.g., crosses, circles, triangles, stars), and number of shapes (e.g., one two, three, four) and are told to match the cards according to some unknown, underlying sorting rule that keeps changing. Participants receive immediate feedback regarding whether they have correctly (“RIGHT” in green lettering) or incorrectly (“WRONG” in red lettering) matched the cards. In this rigged
version, participants complete a total of 120 trials, which is 53.33% more trials than the original WCST and this extended duration is part of the frustration manipulation. Additionally, unbeknownst to participants, the feedback they receive during this modified version is arbitrary to purposely create an illusion of mastery (e.g., successive correct responses) followed by failure (multiple incorrect responses). Finally, at the end of the task, all participants view a results page informing them that they just completed a task assessing their cognitive flexibility and “ability to engage in set-shifting in the response to changing situation demands” as well as a total score that places them in the 52nd percentile. All this information was purposed designed and validated to trigger frustration within this particular sample Harvard students, given their motivations (e.g., intellectual drives; high needs for achievement). This modified version of the computerized WCST was originally developed by the author using C-style language and HTML on the Inquisit Millisecond platform.

**B. Validation Pilot Study for Frustration Induction Paradigm**

**B.1. Participants**

35 students from the Harvard Psychology Department participant pool participated in the pilot testing of the frustration induction paradigm. All participants received 0.50 course credit for partial or full completion of all study tasks, in accordance with Harvard University IRB policies. Ages ranged from 18 – 23 years, and the average was 19.81 years (SD_{age} = 1.38). The sample was 68.6% female. 35% of the sample were White and 35% were Asian. The remainder of the study sample was 12.5% African American, 5% Latino, 7.5% Multi-ethnic, and 2.5% Arab and/or Arab-American, and 2.5% Other. Of the
participants, 3 provided incomplete data and 1 presented as a statistical outlier, leaving 32 participants in the final pilot dataset.

B.2. Measures

*Frustration Induction Paradigm.*

A modified version of the Wisconsin Card Sorting Test (WCST; Berg, 1948)—as described previously in section 2.2.1.2—was used.

*Anger Rating Scale (ARS).*

The Anger Rating Scale (ARS) is a validated state measure of negative mood and consists of 13 bipolar visual analogue scales (e.g., a 100-pt slider scale with labels on each end) measuring feelings related to irritability, anger, and aggressiveness (Henna et al., 2008; Bond et al., 2004; Bond and Lader 1986).

B.3. Research Design

The pilot study used a within-subjects design. The independent variable was the frustration induction. The major dependent variable was the difference between each participants’ self-reported Anger Rating Scale (ARS) score measured before and after receiving the frustration induction paradigm.

B.4. Procedures

Participants from the Harvard Psychology Department participant pool were informed that they would take part in a psychology study on personality and learning (cover
story). Participants were told they would complete a series of questionnaires assessing their attitudes and views about themselves and the world as well as a learning task (frustration induction paradigm). First, they completed the ARS to establish their levels of aggression pre-test. They then completed the modified WCST followed by another version of the ARS with the items in a different order, as well as with some new items, to minimize learning and demand characteristics.

B.5. Data Analysis

To test whether the modified WCST was effective in increasing participants’ aggression or frustration, the mean difference scores from their self-report ARS, measured before and after the frustration induction, was calculated, averaged across all items, and compared to a “test value” of zero (e.g., against a null hypothesis of no difference) in a one-sample $t$ test with Bonferroni adjustment for multiple comparisons.

B.6. Results

The modified WCST successfully increased negative affective states according to an examination of participants’ pre- and post-test ratings on the ARS, $p = .045$.

*State Mindfulness Induction Writing Exercise (2 Groups).*

This state mindfulness intervention involving statement writing exercise is based off the Langerian socio-cognitive mindfulness theory and was designed to create a situation in which participants could attend to a given situation (e.g., an event) from a different perspective by generating a novel way of looking at it. Within the statement writing exercises is where the state
mindfulness induction takes place. Participants in a High mindful group are shown a list of ten negative events (e.g., “You forget your mother’s birthday”; Langer, Delizonna, Pirson, 2010) and asked to write a statement re-evaluating each one in a novel way so that the experimenter can understand “the cognitive strategies used to interpret daily events as they relate to personality” (a purported purpose for the exercise). By contrast, a control, Low mindful group are shown the same list of events as the High mindful group, but instead are prompted to evaluate each one in a non-novel way (e.g., provide a percentage of Americans who have experienced the event).

Photo Selection Task (Behavioral Measure of Aggression).
This task borrows from Denzler, Forster, and Liberman’s (2009) use of negative picture selection as a measure of aggressive tendencies. During the task, participants were asked to help the experimenter in selecting pictures for an unrelated study, and therefore deceived into believing that the pictures would be viewed by another person. Participants received a folder with 30 photographs, always in random order. The experimenter asked participants to choose 10 of these pictures and put their final selection in an envelope. The pictures were 30 photographs from the International Affective Picture System (IAPS; Center for the Study of Emotion and Attention, 1995), which has been widely used in research on affect (e.g., Ito, Cacioppo, & Lang, 1998; Lang, Greenwald, Bradley, & Hamm, 1993; Mussweiler & Förster, 2000) and extensively pretested with respect to their valence (e.g., Ito et al., 1998; Lang, Bradley, & Cuthbert, 1995). The picture sample in this study is the same as in Mussweiler and Förster (2000), who showed that selecting more negative pictures reflects more aggression: by selecting a negative picture, a person makes the viewer uncomfortable. The paradigm is conceptually equivalent to classic
measures of behavioral aggression that typically involve the delivery of noxious or aversive stimuli such as an electric shock (e.g., Geen, 1994).

Photo Color Analysis Tool (Image Color Summarizer).

Participants’ selected photos during the Photo Selection Task were submitted to Image Color Summarizer—a highly reliable image processor—to determine the proportion (e.g., pixel dominance) of different hues featured in them with a specific focus on the color black. Image Color Summarizer provides descriptive and quantitative color statistics of images, including the range of colors (e.g. hue; brightness; saturation) featured in them and their pixel proportions.

C. Validation of Photo Selection Task’s 30 Photos’ Explicit Perceived Harmfulness, as Rated by Naïve American Participants in 2019

The 30-picture sample used in the mindfulness and aggression experiments I and II were the same as those used in Mussweiler and Förster (2000) due to the fact that they had been validated in their behavioral measure of aggression. However, because these kinds of attitudes can vary across culture and time (e.g., the original measure was validated on a sample of college students in Germany in 2000), the same 30 photos were shown to an American sample to determine their perceptions regarding the likelihood that each photo could be used to make a person feel bad. This was done to ensure that the stimuli would remain relevant to the current studies’ major research questions and American sample in 2019.
C.1. Participants

100 participants from the MTURK participant pool completed a short online questionnaire lasting 3-5 minutes. All participants received $2.00 for partial or full completion of the questionnaire, in accordance with Harvard University IRB policies. Ages ranged from 18 – 68 years, and the average was 36.50 years (SD_{age} = 12.60). The sample was 55.1% male. 69.4% of the sample were White. The remainder of the study sample was 14.3% Asian, 10.2% African American, 4.1% Latino, 1% Native American, and 1% Other. Of the participants, 2 provided incomplete data leaving 98 participants in the final pilot dataset.

C.2. Measures

*International Affective Picture System (IAPS).*

Borrowing from Mussweiler and Förster (2000), 30 pictures from the International Affective Picture System (IAPS; Center for the Study of Emotion and Attention, 1995) were used. Pictures varied in terms of valence (e.g., positive; neutral; negative), controlling for arousal. The IAPS has been extensively used in research on affect and pretested with respect to their valence (e.g., Ito, Cacioppo, & Lang, 1998; Lang, Greenwald, Bradley, & Hamm, 1993). Of the 30 pictures included, ten were negatively-valenced (ratings below 4 on a 9-point rating scale where 9 is positive; e.g., a snake), ten were neutral in valence (ratings between 4 and 6 on a 9-point rating scale; e.g., an umbrella), and ten were positively-valenced (ratings above 6 on a 9-point rating scale; e.g., a romantic couple).

*Online Photo Rating Task (Perceived Harmfulness).*
The Online Photo Rating Task consists of each of the 30 IAPS photos from Mussweiler and Förster’s (2000) behavioral measure of aggression individually presented on a screen, alongside a 9-point slider scale. For each photo, participants are to evaluate it in terms of its perceived harmfulness (e.g., “What is the likelihood that this photo could be used to make a person feel bad.” [9 = very high likelihood; 1 = very low likelihood]).

C.3. Research Design

This study involved creating a questionnaire for the purpose of generating quantitative data on attitudes that would be used in other studies. Data (perceived harmfulness) in this study was collected for the purpose of comparing it to—and validating—data obtained in the mindfulness and aggression studies involving American samples’ selections of valenced photos following a frustration induction and after either receiving state mindfulness induction or not. The questionnaire was designed and administered using the Qualtrics online survey platform.

C.4. Procedures

Using either their computers or mobile phones, participants provided electronic consent and then were automatically redirected to a Qualtrics online survey administration platform where they completed the Online Photo Rating Task followed by a demographic questionnaire (e.g., age; sex; race/ethnicity; highest level of educational attainment; self-identified political affiliation).

C.5. Data Analysis
The hypothesis tested in this experiment was that an increase in socio-cognitive mindfulness—operationalized by reframing negative events in novel ways—would lead to a decrease in frustration, as measured by the average valence of photos selected following a frustrating experience.

To determine whether and how naïve participants’ explicit evaluations regarding the perceived harmfulness of the 30 photos corroborated with the photo selections of participants in the mindfulness and aggression study, the two were compared using bivariate correlation analyses.

*Langer Mindfulness Scale (LMS).*

The 21-item LMS questionnaire will be used and scored similarly to previous experiments.

6.3. Research Design

The experiment used a 2-group between-subjects design. The independent variables were the 2 levels of *State Mindfulness Induction* (e.g., Low and High). The major dependent variable was the average valence of the ten photos selected during the Photo Selection Task (*Average Photo Valence*). An additional major dependent variable included the degree to which participants’ ten selected photos’ were likely to make someone else feel bad, as rated explicitly by naïve observers oblivious to the study and to study objectives (*Average Photo Perceived Harmfulness*).
6.4. Procedures

Participants from the Harvard Psychology Department participant pool were informed that they would take part in a psychology study on “personality and daily events” aimed at understanding how personality is related to people’s evaluations of events and would involve completing a personality questionnaire, a learning task, and a writing exercise. The study was advertised through a listing posted on the SONA website Unbeknownst to participants, the personality questionnaire was a decoy to distract them from real aim of the study, the WCST was introduced as a learning task to prime the idea of their competence being evaluated for the purposes of seeing how it related to their personality, and the writing exercise was introduced as an assessment of their evaluations of daily events and the relation to personality to mask its true function as a state mindfulness induction depending on group assignment. Furthermore, participants were led to believe that those were the only three components of the study, which was estimated to last about 45 minutes. This was an overestimation, however, in order to account for the Photo Selection Task at the end of the study, which was introduced as a separate, unrelated and optional study participants could opt to also do.

After providing written consent, participants were randomly assigned to one of two conditions each corresponding to either the Low or High state mindfulness induction writing exercise. All participants began the study by completing a computerized personality questionnaire followed by the frustration induction paradigm. Both the personality questionnaire and frustration induction paradigm were completed on PC computers in the lab room. Participants were then handed a sheet of paper for the writing exercise with instructions on the sheet corresponding to their respective mindfulness condition. Researchers were blind to
participants’ study conditions; a code on their writing exercise sheet allowed for their condition assignment to be determined later during data entry and analysis.

Following the writing exercise, all participants were informed by the researchers that the study was over and were handed a mock debriefing form and asked if they had any questions. The researcher then asked if they were interested in participating in a second study. If participants agreed, to create the illusion of a new and separate study experience, the researcher returned to the lab room with a new, mock consent form for the second study, a manila envelope containing photos and a new laptop (e.g., MacBook) that was different from the one used to complete the previous computerized tasks.

Participants were told this second study involved the selection of photos for a future study and showed them the envelope containing 30 photos. They were instructed to view all 30 photos once the researcher had left the room, and to select 10 they wanted the researchers to consider for their next study and to place those 10 photos in the envelope while leaving the other 20 on the table. Once done, they were to complete a questionnaire (LMS) and demographic survey already loaded on the MacBook laptop.

6.5. Data Analysis

To determine whether the state mindfulness induction was effective in attenuating frustration, the average valence of participants’ 10 selected photos were compared between Low and High mindfulness conditions in a one-tailed independent samples t-test. The average degree to which participants’ 10 selected photos “could be used to make someone feel bad,” as rated explicitly by naïve observers oblivious to the study and to study objectives, will also be compared between groups via a one-tailed independent samples t-test.
6.6. Results

The one-tailed independent samples t-test assessing the effectiveness of state mindfulness induction in attenuating frustration, as measured using the Photo Selection Task, was significant, \( t(87) = 2.20, p = .026, \) Cohen’s \( d = .47, 1-\beta = .60 \). The average (combined) valence for the 10 photos selected post-frustration by individuals in the High mindful condition was higher (more positive) than the average of those in the Low mindful group. The comparison of High and Low conditions’ selected photos in terms of how they mapped onto naïve raters’ explicit evaluations of their perceived harmfulness was also significant, \( t(87) = 2.234, p = .022, \) Cohen’s \( d = .50, 1-\beta = .65 \). On average, individuals in the Low mindful condition selected photos evaluated by naïve raters as more likely to make other people feel bad if shown to them. State mindfulness induction level predicted 5.4% of the variation in participants’ selected photos according to valence \( (R^2 = .054) \) and slightly more of the variation in their selected photos’ explicit capacities to inflict harm on others \( (R^2 = .059) \).

Prior research finds an association between darker image colors and negative affective states like anger. As an additional analysis supporting the major research question and hypothesis of this experiment, participants’ selected photos during the Photo Selection Task were submitted to a color analysis to determine the proportions of various colors in the images. In particular, a one-tailed independent samples t-test comparing the dominance of the color black in participants’ selected photos between High and Low mindful groups was significant and in-line with expectations, \( t(87) = 2.432, p = .017, \) Cohen’s \( d = .52, 1-\beta = .62 \).
When asked at the end of the experiment, 96% of participants did not suspect the frustration task was rigged despite reporting that they found it agitating, and 96% believed the two studies were unrelated.

Two manipulation checks for the frustration induction were embedded in a demographic and personality questionnaire administered at the very end of the study. The two questionnaire items, borrowed from the Anger Rating Scale, that were used to validate the frustration induction in the original pilot study were, “I feel impatient right now” and “I feel composed right now.” There were no significant differences across groups on the manipulation checks. It remains unclear, however, whether the placement of this manipulation check after the Photo Task was optimal in so far as it would make sense for there to be no differences across groups if the Photo Task did, in fact, serve as a cathartic means for displacing frustration and thereby reducing it
before its evaluation by the time participants completed this manipulation check. See Discussion section for more thoughts on this.

7. Experiment 7: Using State Mindfulness Induction to Decrease Frustration, II

This experiment closely followed the previous one and aimed to address two things: (1) firstly, it sought to address one potential confound that the first study on mindfulness and aggression did not, which was related to the possibility that the observed effects were due to positive priming as opposed to mindfulness, itself. This is because the instructions for the High Mindful group in the first study involved reframing in a positively-valenced direction, which one could argue primed the observed effects for this group (e.g., the selection of more positively-valenced photos) compared to the Low Mindful or control group whose instructions were more neutral in valence. (2) Secondly, this experiment sought to disentangle the effects of mere positive re-evaluation from mindful evaluation by controlling for valence in instructions and by varying degree of mindfulness in the form of novelty-production (while controlling for total number of responses across conditions). Experiment 7 also attempted to ensure and confirm that participants’ levels of frustration were sustained throughout the entire experiment by introducing more frustration experiences throughout its design, and one of two manipulation checks for the frustration inductions administered at the end of the study confirms that these attempts were successful.

7.1 Participants
An a priori power analysis indicated that this experiment would need a total of 160 participants (53 in each group) to attain the traditionally desired power threshold (.80) for detecting a medium-sized effect at an alpha of .05. Due to financial constraints and increased MTURK service fees, the sample size decided upon a priori for this study was 120 (40 in each group) with an understanding that it presents reduced statistical power (.70). A higher compensation rate was provided in this experiment to ensure greater quality of responses and full completion of all tasks from online participants.

120 participants in this experiment were recruited from the MTURK participant pool and received $4 for partial or full completion of all study tasks, in accordance with Harvard University IRB policies. Participant ages ranged from 25 - 49, and the average was 39.33 (SD_{age} = 10.02). The sample was 50.8% male. 85.8% of the sample were White. The remainder of the study sample was 10% African American, 4.2% Asian, and 2.5% Latino.

7.2. Measures

_Frustration Induction Paradigm._

A similar task was used as in Experiment 2 with some modifications specific to the anticipated MTurk sample; due to the online administration of this version of the modified WCST, and given prior experience with working with MTurk samples, the elements of a _goal thwarting_ and of _impatience_ were emphasized in their version of the frustration induction task to both heighten frustration and ensure its duration throughout the study. More specifically, participants were told during the consent process that the “learning task” (what the frustration induction was called) they would be completing as part of the personality study included an opportunity to win bonus money according to performance; they were told, “During this task, you will receive bonus
payment based on attention and accuracy in performance. Participants with overall scores falling in the 75th percentile or higher will be eligible for bonus payment at a rate of $0.025 per correct response (maximum possible bonus amount: $3.5).”

*Goal thwarting* was operationalized in the form of their final score at the end, which placed them in the 52nd percentile and therefore not eligible for bonus payment. *Impatience* was operationalized by automatically redirecting them to a second study (e.g., the major behavioral measure of aggression/photo task) without their permission or consent at the end of the personality study in order to create an impression of being made to do more work for less money.

*State Mindfulness Induction Writing Exercises (3 Groups).*

A similar task was used as in Experiment 2 and was adapted into a computerized version in order to run the study online. Within the statement-writing exercises is where the state mindfulness inductions take place. Participants in a High Mindful group are shown a list of five negative events (e.g., “You forget your mother’s birthday”; Langer, Delizonna, Pirson, 2010) and for each event are asked to provide two examples of how the event could prove useful or beneficial to them as well as two examples of how the same event could prove negative (20 total number of responses across 5 events). Participants in a Moderate Mindful group are shown a list of ten negative events, including the five viewed by the High Mindful group, and for each event are asked to provide one example of how the event could prove useful or beneficial to them as well as one example of how the same event could prove negative (20 total number of responses across 10 events). Finally, a Low Mindful control group are shown a list of the same ten events as those viewed by the Moderate Mindful group, and for each event are asked to provide a percentage of the population they believe has experienced the event as well as explicitly identify the valence.
(e.g., negative or positive) of the event (20 total number of responses across 10 events). This group was instructed to also provide their responses in the form of statements in order to equalize effort across groups as much as possible. This last group’s instructions are considered “Low Mindful” due to their lack of novelty-production compared to the Moderate and High groups.

*Photo Selection Task (Behavioral Measure of Aggression).*

The same task was used as in Experiment 2 and was adapted into a computerized version in order to run the study online.

7.3. Research Design

The experiment used a 3-group between-subjects design. The independent variables were the three levels of *State Mindfulness Induction* (e.g., Low/Control; Moderate; High). The major dependent variables were *Average Photo Valence* and *Average Photo Perceived Harmfulness*.

7.4. Procedures

Participants from MTURK’s recruitment pool were informed that they would take part in a psychology study on “personality and daily events” aimed at understanding how personality is related to people’s evaluations of events. In a short description on the MTURK website, they were informed that the MTURK HIT involved answering a series of questions about themselves, completing a learning task (with a chance to win bonus money based on performance), and doing a writing exercise. Altogether, the HIT would take no more than 10 minutes. Unbeknownst to participants, the learning task was programmed to feature a result at the end informing them that they had earned a score within the 52nd percentile and that based on their performance would not
be eligible for a bonus. Furthermore, at the start of the HIT, participants were also purposely made unaware of the real length of the study, which was 20 minutes instead of 10, in order to allow room for a second, unrelated study that they would be automatically redirected into participating in (e.g., the Photo Selection Task). The inability to receive a bonus during the modified WCST, coupled with the reality of a much longer HIT than expected, was all geared toward magnifying the frustration induction for the online version of the mindfulness and aggression study. As originally planned, participants were compensated for 20 minutes of work at the end of the study.

7.5. Data Analysis

The hypothesis tested in this experiment was that an increase in socio-cognitive mindfulness—operationalized by reframing negative events in novel ways—would lead to a decrease in frustration, as measured by the average valence of photos selected for other people to see following a frustrating experience.

Participants’ Average Photo Valence scores will be correlated against their Average Photo Perceived Harmfulness—derived by recoding each of their ten selected photos according to its average perceived harmfulness rating as ascribed by naïve raters in a pilot study prior to averaging across the ten photos—to gauge the degree to which explicit evaluations regarding which kinds of photos would make someone feel bad end up aligning with participants’ implicit interpersonal decisions post-frustration.

To determine whether any of the measured demographic variables were potential confounding variables that needed to be controlled for in major analyses, their effects on the major dependent variables—e.g., Average Photo Valence and Average Photo Perceived
Harmfulness—were assessed prior to performing major analyses. To assess their degree of association to the major DVs and whether they should be controlled for, the demographic variables were either entered in a bivariate Pearson Product-moment correlation against Average Photo Valence and Average Photo Perceived Harmfulness (e.g., age) or submitted to an independent samples t-test (e.g., sex).

Any demographic variable that correlated significantly with either of the major DVs was entered as a covariate in a univariate analysis of variance with the respective DV (e.g., Average Photo Valence or Average Photo Perceived Harmfulness) as the dependent variable and Study Conditions (e.g., High Mindfulness Induction; Moderate Mindfulness Induction; Low Mindfulness Induction) as the independent variables to determine the degree to which state mindfulness predicts post-frustration levels of aggression, as evidenced through the valence of selected photos and naïve raters’ explicit evaluations of which photos are more likely to induce negative feelings in others.

7.6. Results

An examination of the demographic variables revealed that self-identified sex predicted Average Photo Valence, $r (118) = .310, p = .001$, and Average Photo Perceived Harmfulness, $r (118) = -.309, p = .001$. Self-identified sex was therefore controlled for in all analyses involving both major dependent variables.

A univariate analysis of variance with Average Photo Valance as the dependent variable, State Mindfulness Induction Level (e.g., High; Moderate; Low) as the independent variable, and self-identified sex as a covariate was significant, $F (2, 117) = 3.85, p = .024$, Cohen’s $d = .10$, $I−β = .10$. After controlling for the effects of self-identified sex, state mindfulness induction
levels (the IV) predicted 15.3% of the variation in participants’ selected photos according to valence ($R^2 = .153$). The same analysis without the covariate was also significant, $F (2, 117) = 3.86, p = .024$.

A univariate analysis of variance with Average Photo Perceived Harmfulness as the dependent variable, State Mindfulness Induction Level (e.g., High; Moderate; Low) as the independent variable, and self-identified sex as a covariate was significant, $F (2, 117) = 5.65, p = .005$, Cohen’s $d = .31$, $I^\beta = .60$. After controlling for the effects of self-identified sex, state mindfulness induction levels (the IV) predicted 17.6% of the variation in participants’ selected photos according to their perceived harmfulness ($R^2 = .176$). The same analysis without the covariate was also significant, $F (2, 117) = 5.81, p = .004$.

As was done in the first mindfulness and frustration experiment, a supporting analysis looking at the relationship between state mindfulness induction and Average Photo Color Darkness Proportion was performed. Because self-identified sex significantly predicted Photo Color Darkness Proportion in this experiment, $r (118) = -.250, p = .006$, it was controlled for in a univariate ANOVA, and the test was significant, $F (2, 117) = 4.01, p = .021$, Cohen’s $d = .27$, $I^\beta = .62$. After controlling for the effects of self-identified sex, state mindfulness induction levels (the IV) predicted 12.3% of the variation in participants’ selected photos according to the proportion of the color black contained ($R^2 = .123$). The same analysis without the covariate was also significant, $F (2, 117) = 4.32, p = .015$. 

87
None of the participants suspected that the frustration task was rigged when asked at the end of the experiment. They all also believed the two studies were unrelated. There was no difference in state mindfulness task completion time, nor in response word count average, among the three conditions. None of the other assessed demographic variables varied significantly with the major dependent variable. Two manipulation checks for the frustration induction were embedded in a demographic and personality questionnaire administered at the very end of the study. The two questionnaire items, borrowed from the Anger Rating Scale, that were used to validate the frustration induction in the original pilot study were, “I feel impatient right now” and “I feel composed right now.” There were no significant differences across groups on the manipulation checks. Similar to Experiment 6, it’s unclear whether the placement of this manipulation check after the Photo Task was optimal in so far as it would make sense for there to be no differences across groups if the Photo Task did, in fact, serve as a cathartic means for
displacing frustration and thereby reducing it before its evaluation by the time participants completed this manipulation check. See Discussion section for more thoughts on how future work should consider administering more effective manipulation checks for the frustration induction.

As was anticipated and discussed in this experiment’s methods section, based off these results, a post-hoc power analysis revealed that this experiment was underpowered and needed a larger sample size to have enough data for more confident conclusions, particularly regarding the findings for the variable of Average Photo Valence (see Discussion).
CHAPTER 5
DISCUSSION

General Summary

The seven experiments comprising this dissertation sought to demonstrate the effectiveness of an alternative prejudice-reduction method rooted in socio-cognitive mindfulness principles. Socio-cognitive mindfulness is a state in which individuals continually make novel distinctions about objects of their attention. Langer and Imber (1980) argued that, rather than decreasing the number of distinctions people make about social targets, efforts should instead be placed on increasing the number of distinctions people make about their objects of attention, including people. This speaks to a counterintuitive notion of decreasing prejudice by increasing discrimination. That is, rather than getting people to reduce their tendency of categorizing each other as a way to attain a “we are all the same” egalitarian and homogenizing approach to prejudice reduction, we should instead exploit the natural human tendency of categorizing provided that the categories used are mindful (e.g., novel).

In Experiments 1 – 3, as hypothesized, the experimental induction of socio-cognitive mindfulness (novel distinction-making and recategorization) successfully loosened—or “rearranged”—access to old associations in a manner that lowered implicit racial and face-based bias. Not only did the actual strengths of stereotypic associations linearly decrease across four levels of state mindfulness induction (Experiments 1 and 2) as a function of this “loosening”, but this same socio-cognitive state mindfulness induction also reduced implicit face-based bias in Experiment 3 by reorganizing (rather than expanding) participants’ prior representations of babyfacedness in a manner that enabled them to view new babyfaced individuals as more likely
to be in leadership positions and in positions rated high on perceived competence. The average total number of selected faces between groups did not differ and was quite low. Out of a possible total of 8 photos participants could choose, the average was around 3 faces for individuals in both the High and Low mindful conditions. That the individuals in the High mindful condition chose significantly more babyfaced individuals despite their total number of selections remaining the same as the Low mindful group’s suggests that some substitution, in response to the mindfulness intervention, occurred for this group. This difference between the reorganization and expansion of prior representations is an important one. Were expansion the operating mechanism, it would manifest in the form of a greater selection (e.g., number) of photos that included both mature-faced and babyfaced individuals. Substitution here supports the research hypothesis in terms of the state mindfulness induction operating via a mechanism that rearranges prejudices both literally (e.g., via resorting) and figuratively (e.g., conceptually).

Experiment 4 shed more light on the underlying mechanism involved in the state mindfulness induction’s observed effectiveness. Various mindfulness interventions have been shown to be effective by using a range of different materials, and can be divided into two types: (1) those that are universally applicable but whose content is irrelevant to the target subject; and (2) those whose content is specifically relevant to the target subject. In previous work, domain-irrelevant perspectives of mindfulness interventions mainly stressed a complex transfer of cognitive styles from one domain to irrelevant ones (Gang, Zhang, & Zhang, 2011). For instance, Idusohan-Moizer et al. (2015) used domain-irrelevant content for adults with intellectual disabilities to reduce depression and anxiety symptoms.

Through domain-relevant learning materials, research has shown that individuals are more likely to focus on a deeper acquaintance of subjects in one field (Geng et al., 2019), and the
results from Experiment 4 support this. Experiment 4 showed that domain-relevant content is important to the observed effects; namely, state socio-cognitive mindfulness interventions such as the ones presented in this dissertation are most effective when the contents of the paradigm match the target of the interventions.

That the state mindfulness induction condition in Experiment 4 involving four psychological categories was more effective at reducing implicit racial bias compared to the condition involving four physical categories also helped rule out engagement (e.g., participants choosing and creating their own sorting categories) as a necessary factor in the observed effects from the previous experiments.

The results from the fifth experiment align with what one would expect given trait socio-cognitive mindfulness’ relationship to empathy. Trent et al. (2016) found that trait socio-cognitive mindfulness is related to affective and cognitive empathy. Moreover, compared to meditative mindfulness, socio-cognitive mindfulness involves cognitive processes (e.g., perspective-taking, flexibility) that are potentially amenable to experimental manipulations aimed at increasing empathy. The results from the fifth experiment extend Trent et al.’s work by being the first to demonstrate the use of state socio-cognitive mindfulness induction to increase affective empathy using a behavioral measure.

The sixth experiment demonstrated the effects of mindful reframing on reducing frustration, with implications for interpersonal aggression. One potential explanation for these findings, however, is that they were due to a positivity induction in the High mindful group insofar as individuals in this group were instructed to reframe negative events in terms of how they might be useful to them (“write a statement that captures one positive thing that could result from this event occurring”) which is a positively-valenced re-appraisal process. Furthermore, it
could be argued that the instructions for participants in the Low mindful control group were more neutral in valence by comparison, and therefore the observed effects were a function of positive re-appraisal priming participants in the High mindful group to select more positively-valenced photos compared to control. The seventh experiment was designed to control for valence in instructions across conditions while also controlling for the important theoretical difference between mere positive reframing and mindfulness. Nonetheless, the results from the fifth experiment still extend prior work on cognitive reappraisal and aggression.

Extant research has tested the processes explaining how individuals use re-appraisal (e.g., an emotion regulation strategy) to down-regulate negative emotions such as anger, however no research has tested how re-appraisal is related to aggressive behavior despite several theoretical claims regarding its relations (Bartlett, 2011). In fact, evidence for cognitive reappraisal’s potential to reduce reactive aggression in healthy populations is still severely lacking (Qi Jiang et al., 2018). The results from the sixth experiment support a potential effect of cognitive reappraisal on aggression to the extent that aggressive behavior always presupposes the existence of frustration.

In the seventh experiment, mere positive reappraisal and mindfulness are distinguished by introducing mindfulness instructions that vary degree of novelty-production between groups while controlling for valence by maintaining an equal number of positively- and negatively-valenced instructions within conditions. While the state mindfulness induction in this study is different from those used in Experiments 1 – 4, it is very similar to them in that it shares the element of novelty-production which is fundamental to the socio-cognitive mindfulness construct. The results from this experiment resonate with those from Experiments 1 and 4 in so
far as they demonstrate a linear relationship between degree of novelty-production (e.g., levels of mindfulness) and shifts in attitudes (e.g., social bias) and behavior (e.g., frustration).

**On Rearranging and Re-Sorting Prejudices**

This research has important theoretical implications. By focusing on how we rearrange our prejudices, it identifies a natural human tendency that can be exploited to facilitate attitudinal change at an implicit and behavioral level. The research presented contributes to the social cognition, mindfulness, and empathy literatures by highlighting an important piece of the self-other gap related to attention to variability. At its core, novelty-production is about attending to variability. The mechanism underlying the presented interventions is rooted in attending to variability. In the process of re-sorting the same individual according to different dimensions, one comes to find that the same person can support a variety of profiles and characteristics, many of which do not carry any stereotypic associations. Re-sorting allows for the previously homogeneous view of outgroup members to shift through the visually-guided realization that heterogeneity is a possibility. Mindful resorting brings our attention toward the heterogeneity of humanity. In doing so it teaches us things like empathy by revealing to us the skins of others in ourselves and ourselves in the skins of others.

This work is consistent with prior work showing that an increased perception of outgroup variability and heterogeneity can reduce discrimination (Brauer & Er-rafiy, 2011; Park, Ryan, & Judd, 1992). The state mindfulness intervention introduced in these experiments—and its focus on resorting—presents one effective, efficient, and reliable way of invoking a perception and understanding of variability with regards to members of other social group categories that can reduce implicit bias and increase intergroup empathy.
Limitations

This research does not come without limitations. Compelled to limit the sampling size of some experiments due to economic reasons, this significantly affected power and increases margins of error. This is especially relevant for the seventh experiment, whose small sample size calls into question the validity of the findings. Post-hoc power analyses indicated that many of the experiments were underpowered, and it could be argued that there is insufficient data to confidently conclude that the High mindful interventions are more effective than their comparison groups to the degrees currently demonstrated. Replications of the relevant studies, using larger sample sizes, would need to be performed to confirm the current conclusions.

Another limitation presented by the research involves the lack of “true” control conditions in some experiments. While many of the experiments were built of an assumption that the Low mindful groups were akin to a control group in so far as they were designed to highlight the natural human default state of mindlessness, one could argue that the observed effects were partly driven by the high stereotypic activation in the Low mindful conditions as opposed to an effectiveness of the High mindful condition. Current IAT data obtained from larger samples that are much more representative of the real population, however, suggest that the average IAT scores from the Low mindful groups in Experiments 1 – 4 mirror those of the larger population, which confirms the experiments’ original assumptions regarding the capacity for the Low mindful conditions to serve as kind of control group. In other words, the larger population does not differ from those operating under high stereotypic activation, reaffirming the notion that our everyday thinking is a form of mindless prejudice (Langer, 1989; James, 1890).
It is still possible, however, that even this population data, obtained from research working with data and samples predominantly from Project Implicit\textsuperscript{5}, is also not entirely representative of the real population in so far as Project Implicit’s data might reflect a self-selected group of relatively more liberal individuals whose performance falls on the lower end of the true range of implicit bias scores from the larger population.

The research attempted to control for potential differences in cognitive effort between treatment groups, such as by equalizing the total number of sortings involved during the state mindfulness induction (e.g., Experiments 1 – 5) and by equalizing the total number of responses provided during the writing exercise (Experiments 6 and 7). In most of the experiments there was no significant difference in task completion time between conditions, and for those experiments in which there was a difference, time spent on the state socio-cognitive mindfulness tasks did not predict performance on the major dependent variables. In other words, if those time differences were in, in fact, indicative of differences in cognitive effort, they did not mediate the observed relationships between High state mindfulness induction and prejudice reduction.

There is still a chance, however, that task complexity was a factor left uncontrolled for in the experiments in the sense that sorting by psychological categories can be cognitively more complex and demanding than sorting by physical categories in a manner that cannot be captured by looking at differences in task completion times. In the High mindful conditions, novelty and complexity (e.g., cognitive load) are potentially confounded to the extent that the categories used

\---

\textsuperscript{5} Project Implicit is a non-profit organization and collaboration among researchers who are interested in examining implicit social cognition—e.g., thoughts and feelings outside of individuals’ conscious awareness and control. The goal of the organization is to educate the public about implicit biases and to provide an online platform for collecting data across a variety of implicit measures including various forms of implicit association tests (IATs) like the Black-White and Race-Weapons IATs used in Experiments 1 – 4 of this dissertation.
for sorting along psychological dimensions present a more taxing experience than the categories used for sorting along physical dimensions. To rule this out, future research should consider having naïve participants rate a list of candidate psychological and physical categories on their perceived degree of novelty and complexity in order to select categories that vary along novelty while controlling for complexity.

It still stands to note, however, that greater cognitive load as a function of greater task complexity in the High mindful conditions could not explain the observed patterns of effects in the current experiments. One would expect the opposite pattern of results to occur were cognitive effort at play in light of what prior research has shown. Under cognitive load, decision-making regresses to system I processing (Kahneman, 1973), and therefore were the High mindful conditions marked by greater cognitive load as a function of greater complexity, then one should expect to see increased implicit social bias, decreased intergroup empathy, and increased aggression among these individuals, contrary to the current findings.

Another limitation of the current work is that it cannot speak to the longevity of the observed effects in terms of whether the state socio-cognitive mindfulness interventions produced relatively shorter- or longer-term attitudinal and behavioral change. Similar to many of the prejudice reduction techniques introduced in Chapter 2 of this work, the question of longevity, as it pertains to attitude- and behavior-change interventions, is a necessary and important one to examine. As was previously noted, interventions that can afford quick change and that rely on higher-level mental states and beliefs for shifting bias often come with the challenge of trying to maintain those changes overtime. Oftentimes, rapid onset may also mean rapid decay, therefore with effects such as those obtained in the studies comprising this dissertation, it is important that future work explore their longevity. There currently is no data on
how long the observed declines in implicit bias last nor evidence concerning the true magnitude of the observed behavioral changes in terms of whether they are capable of generalizing toward other aspects of a person’s behavioral repertoire (e.g., whether longer listening time to an outgroup member’s emotional story following state socio-cognitive mindfulness induction translates into other forms of empathic behavior and/or prosocial behavior).

While the results from Experiment 3 demonstrate the intervention’s adaptability, in so far as it is capable of being modified (e.g., changing from faces that vary along race to faces that vary along facial maturity) and applied toward reducing bias in another social domain (e.g., face-based bias), there is yet still no support for the intervention’s true generalizability as evidenced by the results from Experiment 4. True generalizability would require such an intervention to produce a reduction in bias across several domains purely as a function of process only regardless of content, and independent of the kind of stimuli used for sorting. Some researchers have managed to develop generalizable interventions that demonstrated effectiveness at reducing prejudice across a variety of domains, such as by priming counter-stereotypic thinking without the use of any stimuli (Vasiljevis & Crisp, 2013). Future work from this dissertation seeks to explore whether there are any higher-order categories (e.g., non-social stimuli like shapes) whose “mindful” sorting could generalize in terms of enabling prejudice reduction across multiple social domains without the need for changing content/stimuli and possibly even produce positive outcomes in non-social domains such as health.

Finally, both Experiments 6 and 7 could have benefitted from more appropriately-timed manipulation checks for the frustration induction. It remains unclear for how long participants in the study were frustrated and therefore how much of an effect the mindfulness intervention had on the observed effects in reducing frustration and third-party punishment. The experiments, as
well as their conclusions, could have benefitted from the inclusion of more sensitive
manipulation checks such as implicit measures of aggression or frustration like sentence
completion tasks or detection tasks for aggressive stimuli that would help reveal the degree to
which participants from each condition were still relatively frustrated based off the accessibility
of related concepts in their minds. Furthermore, administering the manipulation checks at earlier
points in the study would have been ideal as opposed after the Photo Task which likely had
already suppressed any observable aggressive tendencies.

**Recommendations and Future Directions**

Even though much effort and attention continue to be placed on finding techniques to
reduce prejudice, only a minority of the prejudice reduction methods developed in scientific
laboratories have been systematically tested in the field. Er-Radiy and Brauer (2012) make the
important point that most interventions that demonstrate success in the context of the laboratory
remain untested in the real world due to limited attempts at conducting studies in the field and at
testing the generalizability of obtained effects on prejudices in those contexts. Future research in
this area is therefore in desperate need of submitting itself to testing in relevant contexts and on
non-student populations, for example.

As an example of where the work produced from this dissertation is currently being
applied, less research has been dedicated to assessing teacher’s mental illness identification
accuracy, and to those referral decisions teachers make that might impact the observed biases in
the school-to-prison pipeline. The school-to-prison pipeline begins with teachers’ initial
perceptions of student behavior. While teachers cannot make formal diagnoses, they can choose
to address behavioral problems in the classroom, or to seek administrative help by referring
students for harsher disciplinary sanctions or by recommending special education placement. Their school administrations, in turn, can then choose to either disregard or uphold the teachers’ recommendations. Teachers have a lot of power in this process as administrators heavily weight their final decisions on teachers’ reports under the (often false) assumption that teachers have appropriate training and a better understanding of the situations at hand having spent a greater amount of time with the student/s in question.

How objective—i.e., accurate—teachers’ evaluations of students’ behavior is, particularly in those cases where mental illness is present, and how these evaluations vary as a function of the perceived ingroup-outgroup status of students, remains a relatively unexplored topic. Fewer studies on teachers’ perceptions of mental illness in students have appeared in the literature. It remains unclear how equipped the average U.S. general education teacher is to properly identify mental illness symptomatology and how it differs from mere incivility. Furthermore, the lack of clear policies and the ambiguity in evaluation criteria for differentiating mere disobedience from learning difficulties and/or from mental illness may leave room for teachers and educators to unconsciously rely on their preexisting stereotypes about racial groups in their decision-making.

Questions under a relevant and currently-running study that was informed by this body of work include: (1) How perceptions of mental illness inform teachers’ decision-making regarding which students go down the “disciplinary” and “special education” tracks, and furthermore, how this decision-making process is affected by students’ social group membership; (2) How student race affects the accuracy of teachers’ mental health identification and decisions regarding treatment referral, and the degree to which this varies with the magnitude of their implicit racial bias; and (3) whether the introduction of state mindfulness in the educational context (e.g., in
teachers) could make mental health identification in students more accurate as well as facilitate decision-making around treatment referral.
APPENDIX A

State Mindfulness Induction Paradigm Levels (Black & White)

State Mindfulness Induction Face Sorting Task (for Racial Bias)

**LOWEST M-**
1 physical dimension - pale | tan

**LOWER M+**
1 physical dimension - baby | mature face

**MODERATE M+**
4 physical dimensions

**HIGHEST M+**
2 psychological dimensions; 2 self-created psychological

1 physical dimension: pale | tan

1 physical dimension: baby | mature

4 physical dimensions: angular shaped | round shaped
short | long
small | large
pale | tan

4 psychological dimensions: perceptive | not perceptive
sociable | not sociable
self-created 1 | self-created 1
self-created 2 | self-created 2
## APPENDIX B

Black-White Race Implicit Association Test (IAT)

Sequence of Trial Blocks:

**Order 1:**

<table>
<thead>
<tr>
<th>Block</th>
<th>No. of Trials</th>
<th>Function</th>
<th>Items Assigned to Left Key “A” Response</th>
<th>Items Assigned to Right Key “L” Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20</td>
<td>Practice</td>
<td>Black American Faces</td>
<td>White American Faces</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>Practice</td>
<td>Pleasant Words</td>
<td>Unpleasant Words</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>Practice</td>
<td>Black American Faces + Pleasant Words</td>
<td>White American Faces + Unpleasant Words</td>
</tr>
<tr>
<td>4</td>
<td>40</td>
<td>Test</td>
<td>Black American Faces + Pleasant Words</td>
<td>White American Faces + Unpleasant Words</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>Practice</td>
<td>White American Faces</td>
<td>Black American Faces</td>
</tr>
<tr>
<td>6</td>
<td>20</td>
<td>Practice</td>
<td>White American Faces + Pleasant Words</td>
<td>Black American Faces + Unpleasant Words</td>
</tr>
<tr>
<td>7</td>
<td>40</td>
<td>Test</td>
<td>White American Faces + Pleasant Words</td>
<td>Black American Faces + Unpleasant Words</td>
</tr>
</tbody>
</table>
Order 2:

<table>
<thead>
<tr>
<th>Block</th>
<th>No. of Trials</th>
<th>Function</th>
<th>Items Assigned to Left Key “A” Response</th>
<th>Items Assigned to Right Key “L” Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20</td>
<td>Practice</td>
<td>Black American Faces</td>
<td>White American Faces</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>Practice</td>
<td>Pleasant Words</td>
<td>White American Faces</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>Practice</td>
<td>Black American Faces + Unpleasant Words</td>
<td>White American Faces + Pleasant Words</td>
</tr>
<tr>
<td>4</td>
<td>40</td>
<td>Test</td>
<td>Black American Faces + Unpleasant Words</td>
<td>White American Faces + Pleasant Words</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>Practice</td>
<td>White American Faces</td>
<td>Black American Faces</td>
</tr>
<tr>
<td>6</td>
<td>20</td>
<td>Practice</td>
<td>White American Faces + Unpleasant Words</td>
<td>Black American Faces + Pleasant Words</td>
</tr>
<tr>
<td>7</td>
<td>40</td>
<td>Test</td>
<td>White American Faces + Unpleasant Words</td>
<td>Black American Faces + Pleasant Words</td>
</tr>
</tbody>
</table>
# APPENDIX C

Race-Weapons Implicit Association Test (IAT)

Sequence of Trial Blocks:

Order 1:

<table>
<thead>
<tr>
<th>Block</th>
<th>No. of Trials</th>
<th>Function</th>
<th>Items Assigned to Left Key “A” Response</th>
<th>Items Assigned to Right Key “L” Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20</td>
<td>Practice</td>
<td>Black American Faces</td>
<td>White American Faces</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>Practice</td>
<td>Harmless Objects</td>
<td>Weapons</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>Practice</td>
<td>Black American Faces + Harmless Objects</td>
<td>White American Faces + Weapons</td>
</tr>
<tr>
<td>4</td>
<td>40</td>
<td>Test</td>
<td>Black American Faces + Harmless Objects</td>
<td>White American Faces + Weapons</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>Practice</td>
<td>White American Faces</td>
<td>Black American Faces</td>
</tr>
<tr>
<td>6</td>
<td>20</td>
<td>Practice</td>
<td>White American Faces + Harmless Objects</td>
<td>Black American Faces + Weapons</td>
</tr>
<tr>
<td>7</td>
<td>40</td>
<td>Test</td>
<td>White American Faces + Harmless Objects</td>
<td>Black American Faces + Weapons</td>
</tr>
</tbody>
</table>
Order 2:

<table>
<thead>
<tr>
<th>Block</th>
<th>No. of Trials</th>
<th>Function</th>
<th>Items Assigned to Left Key “A” Response</th>
<th>Items Assigned to Right Key “L” Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20</td>
<td>Practice</td>
<td>Black American Faces</td>
<td>White American Faces</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>Practice</td>
<td>Weapons</td>
<td>Harmless Objects</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>Practice</td>
<td>Black American Faces + Weapons</td>
<td>White American Faces + Harmless Objects</td>
</tr>
<tr>
<td>4</td>
<td>40</td>
<td>Test</td>
<td>Black American Faces + Weapons</td>
<td>White American Faces + Harmless Objects</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>Practice</td>
<td>White American Faces</td>
<td>Black American Faces</td>
</tr>
<tr>
<td>6</td>
<td>20</td>
<td>Practice</td>
<td>White American Faces + Weapons</td>
<td>Black American Faces + Harmless Objects</td>
</tr>
<tr>
<td>7</td>
<td>40</td>
<td>Test</td>
<td>White American Faces + Weapons</td>
<td>Black American Faces + Harmless Objects</td>
</tr>
</tbody>
</table>
APPENDIX D

Face Evaluation Task Instructions

Instructions: “Sort these 8 photos according to those who, in your opinion, could be doctors. There is no required total number of photographs per box, so sort freely. For example, you can place anywhere between 0 - 8 of the photographs into the single box on the page. Any photographs that you don't think qualify can stay on the left-hand side, outside of the box.”
APPENDIX E

Big Five Inventory 44 (BFI-44)

Instructions: Here are a number of characteristics that may or may not apply to you. For example, do you agree that you are someone who likes to spend time with others? Please write a number next to each statement to indicate the extent to which you agree or disagree with that statement.

1 = Disagree strongly; 2 = Disagree a little; 3 = Neither agree nor disagree; 4 = Agree a little; 5 = Agree Strongly

I see Myself as Someone Who...

___1. Is talkative  ___23. Tends to be lazy
___2. Tends to find fault with others  ___24. Is emotionally stable, not easily upset
___3. Does a thorough job  ___25. Is inventive
___4. Is depressed, blue  ___26. Has an assertive personality
___5. Comes up with new ideas  ___27. Can be cold and aloof
___6. Is reserved  ___28. Perseveres until the task is finished
___7. Is unselfish with others  ___29. Can be moody
___8. Can be somewhat careless  ___30. Values artistic, aesthetic experiences
___9. Is relaxed, handles stress well  ___31. Is sometimes shy, inhibited
___10. Is curious about many things  ___32. Is considerate to almost everyone
___11. Is full of energy  ___33. Does things efficiently
___12. Starts quarrels with others  ___34. Remains calm in tense situations
___13. Is a reliable worker  ___35. Prefers work that is routine
___14. Can be tense  ___36. Is outgoing, sociable
___15. Is ingenious, a deep thinker  ___37. Is sometimes rude to others
___16. Generates a lot of enthusiasm  ___38. Makes plans and follows through
___17. Has a forgiving nature  ___39. Gets nervous easily
___18. Tends to be disorganized  ___40. Likes to reflect, play with ideas
___19. Worries a lot  ___41. Has few artistic interests
___20. Has an active imagination  ___42. Likes to cooperate with others
___21. Tends to be quiet  ___43. Is easily distracted
___22. Is generally trusting  ___44. Is sophisticated in art and music

Scoring:

BFI scale scoring ("R" denotes reverse-scored items):

Extraversion: 1, 6R, 11, 16, 21R, 26, 31R, 36
Agreeableness: 2R, 7, 12R, 17, 22, 27R, 32, 37R, 42
Conscientiousness: 3, 8R, 13, 18R, 23R, 28, 33, 38, 43R
Neuroticism: 4, 9R, 14, 19, 24R, 29, 34R, 39
Openness: 5, 10, 15, 20, 25, 30, 35R, 40, 41R, 44
APPENDIX F

Frustration Induction Paradigm

Thank you for agreeing to participate in our study.

The purpose of this research is to understand personality. You will complete a personality questionnaire, a learning task, and a statement-writing exercise about your understanding of daily events.

Here are a number of characteristics that may or may not apply to you. For example, do you agree that you are someone who likes to spend time with others? Please indicate the extent to which you agree or disagree with each statement.

You are someone who...

- Is talkative
- Tends to find fault with others
- Does a thorough job
- Is depressed, blue
- Is original, comes up with new ideas
- Is reserved
- Is helpful and unselfish with others
- Can be somewhat careless
<table>
<thead>
<tr>
<th>Question</th>
<th>Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>Race</td>
<td></td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
</tr>
<tr>
<td>Political Affiliation</td>
<td></td>
</tr>
<tr>
<td>Religious Affiliation</td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td></td>
</tr>
</tbody>
</table>

You are about to begin a learning task in which you will match cards according to an underlying characteristic that changes.

The cards will be presented on the computer screen.

Select cards by clicking on them.
RIGHT
RESULTS: This was a neuropsychological test of "set-shifting", i.e. the ability to display flexibility in the face of changing circumstances or rules.

PERCENTILE SCORE: 52

Scores are based on an analysis of all participants run in the study thus far.
APPENDIX G

State Mindfulness Induction Writing Exercises (Experiments 6 & 7)

Experiment 6

All participants (regardless of condition) are told:
“The following exercise is geared toward understanding how personality is related to our understanding of daily events. You will engage in a series of short statement-writing exercises to help us understand how you view daily events.”

High Mindful Group Instructions:
“You will read examples of ten events that happen to people. For each event, write out a statement that captures one positive thing that could result from this event occurring.” (10 responses)

Low Mindful Group Instructions:
“You will read examples of ten events that happen to people. For each event, write a statement that expresses what proportion of the American population you believe has experienced this event.” (10 responses)

Experiment 7

All participants (regardless of condition) are told:
“The following exercise is geared toward understanding how personality is related to our understanding of daily events. You will engage in a series of short statement-writing exercises to help us understand how you view daily events.”

High Mindful Group Instructions:
“You will read examples of five events that happen to people. For each event, provide two examples of how the event could prove useful or beneficial to you as well as two examples of how the same event could not be useful.” (20 responses)

Moderate Mindful Group Instruction:
“You will read examples of ten events that happen to people. For each event, provide one example of how the event could prove useful or beneficial to you as well as one example of how the same event could not be useful.” (20 responses)

Low Mindful Group Instructions:
“You will read examples of ten events that happen to people. For each event, write a statement that expresses what proportion of the American population you believe has experienced this event as well as the valence of the event (e.g., how negative or positive it is).” (20 responses)
APPENDIX H

Anger Rating Scale (ARS)

Instructions: Use your mouse to move the slider across each line provided to describe where you fall on each dimension at this moment.

1. ANGRY --- PEACEFUL
2. FRIENDLY --- QUARRELSONE (R)
3. FURIOUS --- CALM
4. SOCIABLE --- UNSOCIABLE (R)
5. AGGRESSIVE --- COOL-HEADED
6. BELLGERENT --- RESTRAINED
7. RESENTFUL --- TOLERANT
8. PATIENT --- IMPATIENT (R)
9. FRIENDLY --- HOSTILE (R)
10. SPITEFUL --- BENEVOLENT
11. ANNOYED --- COMPOSED
12. PLEASED --- DISGUSTED (R)
13. REBELLIOUS --- COMPLIANT

(R) indicates reverse-scored items
APPENDIX I

Photo Rating Task (Perceived Harmfulness)
APPENDIX J

Empathic Listening Task

Condition 1

White Natasha Photo

Instructions: ‘On the next page, you will listen to a recording of Natasha sharing a story. The recording is about 10 minutes long and is detailed, so you can fast forward through parts of the recording if you wish. When you are ready to advance to the next part, click on the arrow on the bottom right of the page.’

Condition 2

Black Natasha Photo

Instructions: ‘On the next page, you will listen to a recording of Natasha sharing a story. The recording is about 10 minutes long and is detailed, so you can fast forward through parts of the recording if you wish. When you are ready to advance to the next part, click on the arrow on the bottom right of the page.’

Measure:

listening time & page click frequency
APPENDIX K

Other and Self Emotion Evaluations

“Rate how much Natasha experienced the following emotions:”

1. Anger
2. Disgust

3. Excitement
4. Fear
5. Guilt
6. Happiness
7. Hope
8. Pride
9. Sadness
10. Shame
11. Surprise

“Rate how much you experienced the following emotions in listening to Natasha's story: (1 = Definitely Not | 7 = Definitely Yes)”

1. Excitement
2. Pride
3. Anger
4. Surprise
5. Guilt
6. Shame
7. Sadness
8. Disgust

9. Hope
10. Happiness
11. Fear
APPENDIX L

Interpersonal Reactivity Index (Empathic Concern Subscale)

Instructions: The following statements inquire about your thoughts and feelings in a variety of situations. For each item, indicate how well it describes you by choosing the appropriate letter on the scale at the top of the page: A, B, C, D, or E.

A
DOES NOT
DESCRIBE ME

B
WELL

C

D

E

DESCRIBES ME

VERY

WELL

1. I daydream and fantasize, with some regularity, about things that might happen to me. (FS)
2. I often have tender, concerned feelings for people less fortunate than me. (EC)
3. I sometimes find it difficult to see things from the "other guy's" point of view. (PT) R
4. Sometimes I don't feel very sorry for other people when they are having problems. (EC) R
5. I really get involved with the feelings of the characters in a novel. (FS)
6. In emergency situations, I feel apprehensive and ill-at-ease. (PD)
7. I am usually objective when I watch a movie or play, and I don't often get completely caught up in it. (FS) R
8. I try to look at everybody's side of a disagreement before I make a decision. (PT)
9. When I see someone being taken advantage of, I feel kind of protective towards them. (EC)
10. I sometimes feel helpless when I am in the middle of a very emotional situation. (PD)
11. I sometimes try to understand my friends better by imagining how things look from their perspective. (PT)
12. Becoming extremely involved in a good book or movie is somewhat rare for me. (FS) R
13. When I see someone get hurt, I tend to remain calm. (PD) R
14. Other people's misfortunes do not usually disturb me a great deal. (EC) R
15. If I'm sure I'm right about something, I don't waste much time listening to other people's arguments. (PT) R
16. After seeing a play or movie, I have felt as though I were one of the characters. (FS)
17. Being in a tense emotional situation scares me. (PD)
18. When I see someone being treated unfairly, I sometimes don't feel very much pity for them. (EC) R
19. I am usually pretty effective in dealing with emergencies. (PD) R
20. I am often quite touched by things that I see happen. (EC)
21. I believe that there are two sides to every question and try to look at them both. (PT)
22. I would describe myself as a pretty soft-hearted person. (EC)
23. When I watch a good movie, I can very easily put myself in the place of a leading character. (FS)
24. I tend to lose control during emergencies. (PD)
25. When I'm upset at someone, I usually try to "put myself in his shoes" for a while. (PT)
26. When I am reading an interesting story or novel, I imagine how I would feel if the events in the story were happening to me. (FS)
27. When I see someone who badly needs help in an emergency, I go to pieces. (PD)
28. Before criticizing somebody, I try to imagine how I would feel if I were in their place. (PT)

NOTE:

R denotes item to be scored in reverse fashion

PT = perspective-taking scale
FS = fantasy scale
EC = empathic concern scale
PD = personal distress scale

A = 0
B = 1
C = 2
D = 3
E = 4

Except for reversed-scored items, which are scored:

A = 4
B = 3
C = 2
D = 1
E = 0
APPENDIX M

Mindfulness Scale (LMS)

Instructions: Below are a number of statements that refer to your personal outlook. Please rate the extent to which you agree with each of these statements.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Strongly Disagree</td>
<td>Disagree</td>
<td>Slightly Disagree</td>
<td>Neutral</td>
<td>Slightly Agree</td>
<td>Agree</td>
<td>Strongly Agree</td>
</tr>
</tbody>
</table>

1. I like to investigate things. (NS)
2. I generate few novel ideas. (NP) R
3. I am always open to new ways of doing things. (F)
4. I get “involved” in almost everything I do. (E)
5. I do not actively seek to learn new things. (NS) R
6. I make many novel contributions. (NP)
7. I stay with the old tried and true ways of doing things. (F) R
8. I seldom notice what other people are up to. (E) R
9. I avoid thought provoking conversations. (NS) R
10. I am very creative. (NP)
11. I can behave in many different ways for a given situation. (F)
12. I attend to the “big picture.” (E)
13. I am very curious. (NS)
14. I try to think of new ways of doing things. (NP)
15. I am rarely aware of changes. (E) R
16. I have an open-mind about everything, even things that challenge my core beliefs. (F)
17. I like to be challenged intellectually. (NS)
18. I find it easy to create new and effective ideas. (NP)
19. I am rarely alert to new developments. (E) R
20. I like to figure out how things work. (NS)
21. I am not an original thinker. (NP) R

NP = novelty-producing
NS = novelty-seeking
E = engagement
F = flexibility

R = indicates reverse-scored items
REFERENCES


Er-rafiy, A., & Brauer, M. (2012). Increasing perceived variability reduces prejudice and


https://academicworks.cuny.edu/gc_etds/2293


psychological account of extreme racial disparities in school discipline. Perspectives on

implicit evaluative conditioning. *Perspectives on Psychological Science, 11*, 381–
398.

aggression. *PloS one, 8*(10), e78503.

291.

categorization and the process of intergroup bias. *Journal of Personality and Social
Psychology, 59*(3), 475.

categorization and the process of intergroup bias. *Journal of Personality and Social
Psychology, 59*(3), 475.

of Abnormal Social Psychology, 41*, 437 - 448.

85.


