

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Political Science Department -- Theses,
Dissertations, and Student Scholarship

Political Science, Department of

5-2017

Thinking about Race: How Group Biases Interact with Ideological Principles to Yield Attitudes toward Government Assistance

Frank John Gonzalez

University of Nebraska-Lincoln

Follow this and additional works at: <https://digitalcommons.unl.edu/poliscitheses>



Part of the [American Politics Commons](#), [Other Political Science Commons](#), and the [Social Psychology Commons](#)

Gonzalez, Frank John, "Thinking about Race: How Group Biases Interact with Ideological Principles to Yield Attitudes toward Government Assistance" (2017). *Political Science Department -- Theses, Dissertations, and Student Scholarship*. 41.

<https://digitalcommons.unl.edu/poliscitheses/41>

This Article is brought to you for free and open access by the Political Science, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Political Science Department -- Theses, Dissertations, and Student Scholarship by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

THINKING ABOUT RACE: HOW GROUP BIASES INTERACT WITH
IDEOLOGICAL PRINCIPLES TO YIELD ATTITUDES TOWARD GOVERNMENT
ASSISTANCE

by

Frank John Gonzalez

A DISSERTATION

Presented to the Faculty of
The Graduate College at the University of Nebraska
In Partial Fulfillment of Requirements
For the Degree of Doctor of Philosophy

Major: Political Science

Under the Supervision of Professor Elizabeth Theiss-Morse

Lincoln, Nebraska

May, 2017

THINKING ABOUT RACE: HOW GROUP BIASES INTERACT WITH
IDEOLOGICAL PRINCIPLES TO YIELD ATTITUDES TOWARD GOVERNMENT
ASSISTANCE

Frank John Gonzalez, Ph.D.

University of Nebraska, 2017

Advisor: Elizabeth Theiss-Morse

When are people more likely to evaluate race-targeted government assistance based on ideological principles rather than racial prejudice? In order to answer this question, it is necessary to understand the mechanisms by which prejudice influences political attitudes. In this dissertation, I develop a theoretical model for explaining how deep-seated, automatic group biases interact with higher-order, ideological principles in order to influence attitudes toward race-targeted government assistance. I suggest group-based principles are more important than individualistic values or ingroup favoritism in explaining race-targeted policy attitudes. I argue that when people evaluate race-targeted policies, controlled neural processes translate automatic neural processes into broad group-based principles, which then become the primary tool people use to evaluate race-targeted policies. As such, the degree to which people's race-targeted policy opinions are driven by principles rather than automatic group biases is a function of how much controlled processing has occurred. I find support for this model across an array of empirical investigations. In a survey experiment, I find group-based principles to outperform an array of other constructs in predicting race-targeted policy attitudes. Then, in a laboratory experiment, I replicate the primary findings from the survey experiment and show further that the only influence ingroup favoritism has on race-targeted policy

attitudes is through automatic, implicit processes. Further, group-based principles are comprised of a combination of ingroup bias and individualistic values. Next, I investigate the translation process from automatic to controlled processes by examining the implications of discrepancies between people's implicit and explicit attitudes – known as implicit ambivalence. I find that individuals with the most “resolved” racial attitudes are the most likely to evaluate race-related political objects ideologically. Finally, I directly examine the automatic and controlled neural processes hypothesized to underlay this model by using functional Magnetic Resonance Imaging (fMRI). I find evidence that activation in brain regions implicated in controlled processing is associated with explicit evaluations of race, and this relationship is strongest among individuals with more extreme group-based principles. I conclude by discussing implications of this model for existing literature as well as possible public policy interventions for reducing the role of prejudice in politics.

Dedication

This dissertation is dedicated to my parents, Barbara and Frank Gonzalez, who taught me the value of integrity and how to stand up for truth and change, to my brother, Peter ‘PJ’ Gonzalez, who constantly impresses me and, despite punching me in the face that one time in high school, will always be my best friend, and to Jesse Anderson, who I admire tremendously, who brings joy to every day, and with whom I cannot wait to share many more adventures with. I have been very fortunate to have such remarkable people in my life. I also dedicate this dissertation to Mike Rithjin, from the network.

Acknowledgements

This dissertation is a product of many people’s hard work beyond just my own. I would like to thank the professors at the University of Delaware who allowed me to gain invaluable research experience in their labs, including Steven Most, Samuel Gaertner, and especially, David C. Wilson, who inspired me to pursue political psychology in an introductory political science course, took me under his wing, and guided me through the process of applying to graduate school despite my less-than-impressive GPA (which was mainly due to Chemistry and Calculus courses, I swear; and maybe one Art History course).

At the University of Nebraska – Lincoln, I have been fortunate to work with top-notch scholars who do ground-breaking research while also managing to be incredibly genuine and nice people. Sergio Wals was among the first to welcome me with open arms, and has continued to provide me with invaluable guidance academically and professionally throughout my graduate career. Ingrid Haas has

always made time for me and treated me as nothing less than a colleague. She also taught me everything I know about fMRI, and the fMRI chapter in this dissertation would not exist without her guidance. Sarah Gervais has been a continuing source of knowledge and inspiration through the dissertation process, and her encouragement throughout the process, especially regarding ideas that I feared social psychologists might laugh at, has been invaluable. I have also worked with terrific scholars at the University of Nebraska Public Policy Center – most notably, Lisa Pytlik Zillig, Myiah Hutches, Peter Muhlberger, and Alan Tomkins. These scholars have provided guidance and encouragement at every step of my graduate career.

I owe a special debt of gratitude to John Hibbing and Kevin Smith, who invited me to work on their ground-breaking research in the biology and politics lab since I started my graduate career. I could not have been luckier than to be able to work with such terrific scholars who also happen to be all-around great guys, and much of the work in this dissertation simply would not exist if it weren't for their academic and professional guidance. John and Kevin also make the research process fun, and are exceptionally unpretentious. You would never know how renowned they are from meeting them (and I know they will both take that as a compliment).

I have not met – nor can I think of – a better academic role model than Elizabeth Theiss-Morse. Her academic and professional guidance has been instrumental throughout my graduate career, and she has never treated me as anything less than a colleague. Beth provides advice on research projects as if they are her own yet takes none of the credit. She approaches academia with a curiosity that is contagious, and is truly selfless in what she does for her students and colleagues. I am lucky to have had her as a graduate advisor.

A critical part of being able to complete this dissertation as well as succeed in graduate school more broadly has been the assistance provided by the administrative staff in the department, Helen Sexton and David Csontos. Helen and David have always gone above and beyond to help with any problem I've encountered, and they make working on the 5th floor of Oldfather Hall quite enjoyable.

Finally, I would like to thank the extraordinary group of young scholars I have had the privilege of working with throughout graduate school. These talented individuals include: Johnathan Peterson and Stephen Schneider (my computer lab compatriots); Balazs Feher-Gavra, Kristen Deppe, Nathan Munier, and Jayme Neiman (former computer lab compatriots, but still compatriots in general); Samantha Lauf (newest computer lab compatriot); Tim Collins, Amanda Friesen, Karl Giuseffi, Michael Gruszczynski, Carly Jacobs, Alex Kroeger, and Benjamin Sievert. I should take an extra sentence to emphasize the continued friendship and collaborative support of Johnathan Peterson and Stephen Schneider, or else I probably won't hear the end of it.

Grant Information

A substantial portion of this dissertation was made possible by university or federal funding. The University of Nebraska – Lincoln Department of Political Science has been financially generous in a number of ways such as providing funding for conference registration and travel as well as funding data collection for Chapter 2 through a McPhee Fellowship. Data collection for Chapter 5 was funded entirely through a Political Science Doctoral Dissertation Research Improvement Grant (PS DDRIG) provided by the National Science Foundation Division of Social and Economic Sciences (NSF SES; Award #1560432).

Table of Contents

CHAPTER 1: Lizard Brain Meets “Higher-Order” Thinking.....	1
Two Schools of Thought about Political Thought.....	8
The “Principles or Prejudice” Literature in Political Science.....	14
CCSoT Work on “Principles versus Prejudice”.....	15
Principles or (Implicit?) Prejudice: Bringing in the NMSoT.....	17
Mediation of Implicit Measures by Explicit Measures.....	18
Group-Based Principles.....	20
Evolution, Ingroup Bias, Conscious Control, and the Development of Group-Based Principles.....	24
The “Lizard Brain” in Politics.....	25
Controlled Processing, Rules, and Higher-Order Thought.....	30
Interactions between Controlled and Automatic Processes.....	31
Bringing it back to Race-Targeted Policies.....	34
Caveats.....	35
Chapter Outline.....	38
CHAPTER 2: Disentangling “Principles” and “Prejudice”: Ingroup Favoritism, Individualism, and Group-Based Principles.....	43
Competing Conceptualizations of Individualism.....	46
Social Dominance, Group Identity, and the Multiple Facets of Each.....	48
Is Individualism Just an Excuse for Prejudice?.....	53
Overview and Hypotheses.....	55
Self-Reliance versus Work Ethic.....	55
Group-Based Principles, Social Identity Motives, and the Dimensions of Each...	56
Explaining “Principled Mediation”.....	57
Data and Measures.....	58
Results.....	62
Confirmatory Factor Analyses.....	62
Structural Equation Models.....	65
Putting Merit to the Test: Effects of Work Ethic Manipulations.....	68
Testing the “Principled Mediation” Hypothesis.....	69
Taking Stock of What “Really” Matters.....	73
Is There Room for Non-Group-Based Principles?.....	73
The Dominant Role of Group-Based Principles.....	74
Individualism: Not Just an Excuse for Prejudice.....	75
But Have We Really Ruled Out Ingroup Favoritism?.....	76
CHAPTER 3: Ingroup Bias in the Raw: The Minimal Groups Paradigm.....	77
Ingroup Bias: A Versatile Tool.....	82
Overview and Hypotheses.....	83

Data and Methods	85
Results	90
Replicating Chapter 2 Results.....	90
Preliminary MGP Results	91
Minimal Ingroup Bias and Race-Based Policy Attitudes	93
Ingroup Bias, Merit, and SDO	96
The Subtle Role of Ingroup Bias, and the Not-So-Subtle Role of Group-Based Principles.....	101
Evidence of “Pure” Individualism? Not Really... ..	102
The Group-Based and “Principled” Foundations of Group-Based Principles	103
Implications for How We Think About Race and Politics	104
CHAPTER 4: When Automatic and Controlled Collide.....	105
Do We Really Need Implicit Attitude Measures?	108
Implicit Ambivalence: The Implicit-Explicit Gap	112
Expectations for Implicit Ambivalence	114
Need for Cognition, Education, and Resolving Conflicting Attitudes	114
Implicit Racial Ambivalence and Race-Related Political Outcomes	116
Data and Measures	117
Results	123
What about Ingroup Favoritism?	133
Translating from Automatic to Controlled	135
CHAPTER 5: INHIBITION OR IDEOLOGY? THE NEURAL MECHANISMS OF PRINCIPLES AND PREJUDICE.....	140
The Consequences of “Thinking about Race”	143
A Neural Framework of the Automatic and Controlled Components of Prejudice ...	145
Expectations for the Role of Controlled Processing	148
Data and Methods	150
Experimental Design.....	150
Survey Measures	152
MRI Data Acquisition.....	153
MRI Data Preprocessing and Analysis	153
Results	155
Behavioral Task Results.....	155
fMRI Results	158
Inside the Black Box of “Principles and Prejudice”	162
CHAPTER 6: A Neural Toolkit for Change, Anchored in the Past.....	167
Summary of Results	169
Principles, Prejudice, or Principled Prejudice?	173
Other Group-Based Principles, Symbolic Racism, and System Justification.....	177

Symbolic Racism: Not So “New”	178
SDO or System Justification?	180
Rules for Radicals, and Reactionaries Too	181
Acknowledging Factors Formerly “Beyond Our Scope”	182
Social Desirability “Influence,” Not “Bias”	183
Nudging Away Prejudice?	186
Structuring (Not Simply Encouraging) Racial Thought	187
Leveraging Social Desirability Influence	189
BIBLIOGRAPHY.....	195
APPENDIX.....	203

List of Tables and Figures

Table 1.1: Basic Tenets of Nonconscious Motivations School of Thought (NMSoT) versus Conscious Considerations School of Thought (CCSoT)	12
Table 2.1: Correlation Matrix for All Latent Factors of Individualism, Social Identity Motives, and Group-Based Principles	64
Table 2.2: Manipulation Check – Perceptions of “Hardworking” by Condition.....	68
Table 2.3: Effects of Work Ethic Manipulations on Race-Based Policy Attitudes	69
Table 2.4: Predicting Race-Based Policy Attitudes With and Without SDO-E	71
Table 3.1: Predicting Race-Based Policy Opinions with Group-Based Principles, Social Identity Motives, and Individualism.....	91
Table 3.2: Race-Based Policy Attitudes, Merit, Explicit MG Bias, and Behavioral MG Bias.....	94
Table 3.3: Race-Based Policy Attitudes and Implicit MG Bias	95
Table 3.4: SDO, Merit, Explicit MG Bias, Explicit MG Bias, and Behavioral MG Bias.....	97
Table 3.5: SDO and Implicit MG Bias	101
Table 4.1: Correlations Between Racial Attitude Measures and Covariates of Interest.....	125
Table 4.2: Implicit Racial Ambivalence, Principles, and Race-Related Political Outcomes	137
Table 5.1: Significant Clusters of BOLD Activation in ROIs for Applicant Race*Conscious Awareness	159
Table 5.2: Significant Clusters of BOLD Activation in ROIs for Applicant Race*Conscious Awareness*SDO	161
Figure 2.1: SEM Predicting Race-Based Policy Attitudes with Individualism, Social Identity Motives, and Group-Based Principles.....	66
Figure 2.2: Testing “Principled Mediation” of Work Ethic by SDO-E.....	72
Figure 3.1: Implicit MG Bias across Levels of Support for Race-Based Policies.....	96
Figure 3.2: Effects of Merit Manipulations across Levels of SDO-D	98
Figure 3.3: Implicit MG Bias across Levels of SDO-D.....	101
Figure 4.1: Implicit Racial Ambivalence in the United States (2008).....	124
Figure 4.2: Implicit Racial Ambivalence x Ideology Predicting Support for Affirmative Action.....	130
Figure 4.3: Implicit Racial Ambivalence x Egalitarianism Predicting the Probability of Voting for Obama in 2008	132
Figure 4.4: Implicit MGP Ambivalence and Ingroup Favoritism.....	134
Figure 4.1: Interaction between Applicant Race, Conscious Awareness, and SDO-D	157
Figure 5.2: BOLD Activation in OFC in Response to Explicit > Implicit Blocks for Black > White Trials	160
Figure 5.3: BOLD Activation Associated with SDO-E Extremity in ACC in Response to Explicit > Implicit Blocks for Black > White Trials	162
Figure 5.4: BOLD Activation Associated with SDO-E Extremity in OFC in Response to Explicit > Implicit Blocks for Black > White Trials	162

CHAPTER 1

Lizard Brain Meets Higher-Order Thinking

“There's no reason that we should give up that lever on people's behavior - namely, the inhibition systems of the brain - just because we're coming to understand more about the temptation systems.”

- Stephen Pinker¹

How much control do people have over their racial attitudes and behaviors?

The answer to this question holds considerable implications for not only how individuals of different races interact with one another in face-to-face situations, but also how racial groups interact broadly, whether people think the government should play a role in encouraging group equality, and how societal resources are distributed across groups through government. If it were up to “controlled,” conscious motivations alone, perhaps most people would maintain a consistently egalitarian set of attitudes, beliefs, and behaviors. However, we know from decades of research in social and evolutionary psychology that humans are deeply engrained with motivations to favor ingroups and derogate outgroups. Even beyond ingroups and outgroups, individuals are often motivated to preserve hierarchy within groups such that the dominant members stay on top and the subordinate members stay on the bottom. It can be comforting to think that if we want to, we can override whatever prejudiced instincts we might have and treat individuals the same regardless of race, but subconscious, unintentional group biases can influence a wide swath of attitudes and behaviors even among individuals who for the most part embrace egalitarian values. The central question of this dissertation is: how do “higher-level,” conscious

¹ Pinker, S. (2008). A biological understanding of human nature. In Brockman, J. (Ed.), *Science at the Edge: Conversations with Leading Scientific Thinkers of Today* (pp. 37-56). New York, NY: Sterling Publishing Company.

values, beliefs, and ideological thinking interact with primal, often nonconscious group-based motivations to influence attitudes toward race-targeted policies?

Answering this question is a crucial first step to answering another, more normatively motivated question: can “higher-level,” conscious thoughts and beliefs override group biases? On the one hand, egalitarian principles pervade Western culture, children are taught from a young age that racism is bad, and all in all, overt forms of racial discrimination have decreased substantially over the past several decades. On the other hand, it is abundantly clear that racial prejudice has not vanished. Instead, covert and implicit forms of racial biases continue to cast a shadow on virtually every facet of society, from who gets hired for jobs, to the probability of being shot during an altercation with law enforcement, to whether or not a patient is prescribed painkillers by a doctor, to who people find sexually attractive. Despite people’s ability to, under many circumstances, inhibit racial biases – despite human-made institutions created to instill egalitarian norms and human beings’ ability in many cases to abide by the norms set forth by those institutions – we seem continually compelled by a puppet master from our evolutionary past to prefer a government set up so that our group gets first dibs.

A substantial literature exists in political science and psychology that speaks to the above question in the context of studying attitudes toward race-targeted policies. In addressing this question, which some have referred to as the “principles or prejudice” question, scholars have debated whether opposition to policies like affirmative action is based on individualistic principles – e.g., the belief that everyone should “pull themselves up by their bootstraps” – or racial prejudice. The empirical evidence is largely mixed, which is likely at least in part due to some conceptual

ambiguity. Although much of the literature addresses how individualism and racial prejudice might interact, researchers have essentially skipped answering the central question of this dissertation – that is, how “higher-level” processes interact with automatic group biases – and so we never gained an understanding of the psychological processes that “principles” and “prejudice” actually reflect. Do the effects of individualism simply indicate socially acceptable expressions of deep-seated group biases, or are people genuinely capable of inhibiting their biases and adhering to a consciously held ethos?

There has been an explosion of research in psychology, economics, neuroscience, and more recently political science showing that people’s attitudes and behaviors – including those related to political outcomes – are driven largely by imperceptible, automatic biological processes that hum along under the surface of conscious awareness (e.g., Eagleman 2011; Gazzaniga 2012; Hibbing, Smith, and Alford 2013; Kahneman 2011; Lodge and Taber 2013; Marcus 2013). Many attitudes people have and decisions they make are easily interpreted as the result of conscious deliberation, when in reality they are heavily influenced by forces operating subconsciously. Messy, foul-smelling rooms can influence people’s moral judgments (Schnall, Haidt, Clore, and Jordan 2008). Disease salience can make people more likely to oppose gay marriage (Faulkner, Schaller, Park, and Duncan 2004). Room temperature can influence people’s views on global warming (Risen and Critcher 2011). Drinking Sprite, which increases blood glucose levels, can decrease support for social welfare (Aarøe and Petersen 2013). Exemplifying perhaps the most imperceptible forces on political opinions and behavior is the work showing that

political ideology is between 25% and 60% genetically heritable (e.g., Alford, Funk, and Hibbing 2005; Hatemi et al. 2010; Hatemi et al. 2014; Settle et al. 2010).

This work seems to fly in the face of traditional theories of human decision making that assume individuals weigh costs and benefits, deliberate about important decisions, and in particular relation to this dissertation, form political opinions based on careful consideration of issues in relation to one's own consciously held values. Accordingly, it is reasonable to expect that in the domain of race-targeted government assistance, despite conscious efforts, deep-seated motivations to favor one's ingroup will continue to dominate people's evaluations. In other words, genuine adherence to "principles" may not stand a chance in the face of automatic group biases.

Yet a substantial amount of evidence supports the idea that people are indeed capable of doing exactly these things, under some circumstances. "System 2" thinking, or thinking that is slow, deliberate, and analytical, is readily acknowledged by many scholars as existing alongside "System 1" thinking, which is fast, automatic, and instinctual (Kahneman 2011; Stanovich 2004). A stable personality trait shown to vary significantly across people is Need for Cognition (NFC), which is the degree to which someone is motivated to think effortfully and solve complex problems (Cacioppo and Petty 1982), suggesting there is at least some variation in the degree to which people utilize what might be called "System 2" thinking. Although many of the assumptions of rational choice theories of voting behavior (e.g. Arrow 1963; Downs 1957; Muller and Satterthwaite 1977) have often been lambasted by proponents of behavioral models (e.g. Green and Shapiro 1994), in the aggregate, some rational choice models do fairly well at predicting voting patterns (e.g. Feddersen, Gailmard, and Sandroni 2009; Levine and Palfrey 2007). In many cases, it is hard to argue that

humans do not possess some ability to weigh costs and benefits, even if the weighing of costs and benefits is constrained substantially by affective processes and attentional limits (Simon 1946; 1982). Most proponents of the role of nonconscious processes readily acknowledge some role – even a substantial role – for conscious factors influencing people’s attitudes and behavior (e.g. Jost et al. 2003; Lodge and Taber 2013; Mondak 2010; Smith, Oxley, Hibbing, Alford, and Hibbing 2011). “Strategic” behavior based on expected utility is particularly evident among political elites (e.g. Aldrich 1995; Downs and Rocke 1994; Fearon 1994), who are presumably just as human as anyone else. Even rhesus monkeys have shown evidence of behavior consistent with utility maximization (Pastor-Bernier, Plott, and Schultz 2017).

With regard to race, a great deal of work has shown that individuals vary in the degree to which they are motivated to inhibit racial biases, and that through top-down neural processes individuals can often inhibit prejudiced behavior (Cunningham et al. 2004; Plant and Devine 1998; Stanley, Phelps, and Banaji 2008). Specifically regarding attitudes toward race-based government assistance, although many scholars maintain that the predominant factor driving opposition to such policies is racial prejudice (e.g., Henry and Sears 2002; Kinder and Sears 1981; Reyna, Henry, Korfmacher, & Tucker 2005), the empirical evidence nonetheless suggests ostensibly nonracial ideological principles play at least a substantial role, if not the predominant role (Sniderman, Crosby, and Howell 2000; Sniderman et al. 1996). If people are indeed able to consciously deliberate to exert influence over deep-seated biases, then it is plausible that individuals *can* override racial biases when evaluating race-targeted policies; even further, it is possible that under the right conditions, nonconscious motivations may even become trivial in comparison to conscious considerations.

The main takeaway from the research described above is that although burgeoning literatures in psychology, economics, and political science shed light on the deep-seated, gut-level motivations that silently drive human behavior, other recent work shows that these forces are not the be-all and end-all. Just as millennia-old evolutionary selection processes have imbued humans with tribal motivations to favor ingroups and avoid outgroups, more recent (though still millennia-old), human-specific evolutionary selection processes have endowed humans with existential motivations, “principled” ideological beliefs and thoughts, the ability to deliberately estimate costs and benefits, and the capacity to evaluate issues in relation to consciously held values.

In this dissertation, I develop a theoretical model that updates our understanding of political cognition by extending work on nonconscious, primal factors that drive human behavior in combination with the ability of humans to consciously engage with their environment and carefully evaluate issues according to ideological principles. I focus particularly on political cognition as it relates to race-related political attitudes because the influence of racial prejudice in politics is an oft-studied and central topic in political psychology, and because there are clear implications for how automatic biases might interact with higher-level principles when people evaluate race-related political issues. I utilize a variety of survey and experimental designs, as well as an examination of neural activity via functional Magnetic Resonance Imaging (fMRI), to study the degree to which individuals’ race-targeted policy attitudes are influenced by sheer ingroup bias, ideological principles, and the interaction of these two broad categories of factors.

In the remainder of this chapter, I outline two broad schools of thought underlying how we understand political cognition and how most contemporary work suggests some middle ground between these schools of thought likely holds the most promise for advancing understandings of political attitudes. Next, I describe the state of the political science literature that seeks to understand attitudes toward government assistance, and in particular whether “principles” or “prejudice” are primarily responsible for opposition to policies like affirmative action. I explain how the most recent work on this topic acknowledges to some degree how automatic processes might interact with conscious perceptions and attitudes to influence policy opinions, but this work nonetheless fails to fully explicate the implications of these different processes. I leverage neuroscience and evolutionary psychology literature on how human-specific, evolutionarily young neural systems have evolved on top of basic, evolutionarily old systems in order to explain how nonconscious and conscious processes may interact to influence policy attitudes. Then, I summarize where the political science literature has left us, and how acknowledging the interaction between deep-seated nonconscious motivations and higher-level ideological principles is necessary to understand the psychological mechanisms of how people evaluate race-targeted government assistance. Thus, ultimately, and perhaps somewhat unsatisfyingly, I suggest the answer to the question of whether “principles” or “prejudice” underlay race-targeted policy attitudes is: both do. However, this answer has less to do with the main effects of two distinct psychological variables and more to do with the neural processes that force principles and prejudice to interact with one another. Finally, I describe several caveats regarding the framework I am proposing before outlining the remaining chapters of the dissertation.

Two Schools of Thought about Political Thought

An overarching distinction can be made in political science between work that focuses on the controlled, conscious factors influencing political outcomes and work that focuses on the automatic, often nonconscious factors influencing these outcomes. I refer to these branches of work as emanating from distinct schools of thought throughout this dissertation, but this distinction is intended as an illustrative tool rather than an accurate depiction of discrete classes of political science scholarship. The two schools of thought I describe do not embody all political science research, and in fact most contemporary work on political attitudes and behavior readily acknowledges both schools of thought. Nonetheless, these two schools of thought can be seen as reflecting the ends of a spectrum upon which models of political attitudes and behavior lay.

One branch of political science research, a school of thought I refer to as the “*Conscious Considerations*” School of Thought (CCSoT), is comprised of work that draws from a wide range of political science subfields and is centrally concerned with studying the conscious decisions individuals make when evaluating political objects (e.g., issues, candidates, parties, policies, or anything political that individuals may form an opinion on). The CCSoT encompasses research on topics such as candidate evaluations, elite decision making, international diplomacy, war, trade, judicial behavior, issue attitudes, vote choice, political participation, interest group behavior, and many others, but the central tenet of the CCSoT is that individuals are to be viewed as self-aware actors who consider information from their environment and behave accordingly. This school of thought is heavily influenced by work in rational choice and economics, as the conscious considerations studied under this school of

thought often amount to what is referred to as a “utility function” in econometrics literature. Yet it is not limited to these perspectives. Rather, sociological and psychological work on the influence of social movements, stereotypes and culturally or historically imbued knowledge could also fall under the umbrella of the CCSoT.

The CCSoT has been an integral part of understanding political behavior since before political science was recognized as a discipline, and it largely remains central to the field, but work building on neuroscience as well as social and evolutionary psychology has gained a significant amount of traction in the field over recent decades. This body of political science research, which I refer to as the “*Nonconscious Motivations*” School of Thought (NMSoT), is also comprised of work in a fairly wide range of political science subfields, but is predominantly made up of work in political psychology, political behavior, and more recently, biology and politics and political neuroscience. Scholarship advancing the NMSoT has primarily (although not entirely) centered on a narrower range of topics than work on the CCSoT – topics such as political ideology, candidate evaluations, vote choice, issue stances, and political participation – largely because of the relative youth of this school of thought and the direct relevance of these topics to psychology (for counter-examples, see e.g., Gibson 1981; Klein and Mitchell 2010; McDermott 2004; Mutz and Kim 2016). The NMSoT is primarily concerned with the deep-seated, evolutionarily imbued, often nonconscious neural and biological factors that drive people’s political attitudes and behavior. A cornerstone of this school of thought is that the bulk of factors driving human behavior lay outside of conscious awareness, and that the effects of conscious considerations such as those studied under the CCSoT pale in comparison to the power of primal, nonconscious motivations. In fact,

some scholars subscribing mostly to the NMSoT have suggested conscious attempts to “control” one’s attitudes and behaviors are merely post-hoc rationalizations of mental processes that have occurred long before the individual is aware of the issue (e.g., Eagelman 2011; Kahneman 2011; Lodge and Taber 2013; Pinker 2002; 2008).

A good deal of empirical work has supported both the CCSoT and the NMSoT in relation to politics, and as mentioned earlier, most contemporary scholars acknowledge both. In line with the CCSoT, individuals do seem to consciously weight costs and benefits pertaining to politics in some circumstances (e.g. Aldrich 1995; Downs and Rocke 1994; Fearon 1994; Simon 1982). Walter Lippmann spoke explicitly of the “images in our heads” that we construct to evaluate the political world (1946), and a tremendous amount of work has since shown consciously held perceptions acquired through culture, the media, or parental socialization to be integral to political cognition (Iyengar and Kinder 2010; Jennings, Stoker, and Bowers 2009; Zaller 1992), including race-related political cognition (Hurwitz and Peffley 1997; Peffley, Hurwitz, and Sniderman 1997; Kinder and Sears 1981).

However, in line with the NMSoT, “hot” affective responses to political stimuli often predict political attitudes and behaviors better than “cool” deliberative thought (Lodge and Taber 2005; 2013; Redlawsk 2002). Activity in the sympathetic nervous system (electrodermal activity, or EDA) independently predicts a range of political attitudes (Dodd et al. 2012; Gruszczynski et al. 2013; Mutz and Reeves 2005). Baseline levels of cortisol – a hormone associated with stress – predict voter turnout (French et al. 2014; Neiman et al. 2015). As already mentioned, political ideology is estimated to be between 25% and 60% genetically heritable (e.g., Alford, Funk, and Hibbing 2005; Hatemi et al. 2010; Hatemi et al. 2014; Settle et al. 2010).

With regard to race, environmental cues need not be explicitly tied to race to evoke prejudiced responses (Hurwitz and Peffley 2005; Mendelberg 2001; Valentino, Hutchings, and White 2002; but see Huber and Lapinski 2006). Critically, however, most of this work advancing the NMSoT has not done so by ignoring the CCSoT, and it has often been that case the scholars studying nonconscious motivations argue that the manifestations of nonconscious motivations are often shaped by contextual forces and conscious processes (e.g. Jost et al. 2003; Mondak 2010; Smith et al. 2011).

A particularly important strand of research on nonconscious motivations is work on implicit racial attitudes. Implicit measures of racial attitudes utilize computer tasks to gauge the associations people have between racial groups and positive or negative words or images. Implicit measures such as the Implicit Association Test (IAT; Greenwald, McGhee, and Schwartz 1998) and Affect Misattribution Procedure (AMP; Payne et al. 2005) tap automatic mental processes that occur before individuals have the chance to edit their responses consciously, and have been shown to be valid independent predictors of a wide range of race-related variables (Fazio, Jackson, Dunton, and Williams 1995; Greenwald, Poehlman, Uhlmann, and Banaji 2009; Olson and Fazio 2009).

Ultimately, there is overlap between the CCSoT and NMSoT, but a critical set of tenets of each school of thought nonetheless differentiate the factors emphasized by the two. In their extreme forms, the two schools of thought may be best thought of as two incomplete regression equations. The CCSoT focuses centrally on conscious factors influencing human attitudes and behavior and tends to leave nonconscious factors to the error term. For example, research falling strictly under the purview of the CCSoT might aim to understand attitudes and behavior by focusing on factors

related to the values and beliefs by which individuals aim to abide, the stereotypes upon which people form beliefs, the weighing of costs and benefits that goes into a decision, or conscious reactions to life events or events in the world (*Equation 1, E.1*). Conversely, factors related to the NMSoT might include genetic predispositions (or other prenatal factors), individual differences in personality traits, nonconscious physiological responses to environmental stimuli, or implicit associations (*Equation 2, E.2*).

$$E.1: \quad \textit{Attitude or Behavior} = \textit{Consciously Held Beliefs and Values} + \textit{Conscious Decisions} + \textit{Conscious Reactions to Life Events} + e$$

$$E.2: \quad \textit{Attitude or Behavior} = \textit{Genetic Predispositions} + \textit{Personality Traits} + \textit{Physiological Responses to Environment} + \textit{Implicit Preferences and Biases} + e$$

This is not an exclusive list of the factors emanating from each school of thought, but rather examples of factors that reflect each. Other factors may be studied within each school of thought, but in this dissertation a set of central tenets unites those within each incomplete equation. Table 1.1 contains the central tenets of each school of thought.

Table 1.1. Basic Tenets of Nonconscious Motivations School of Thought (NMSoT) versus Conscious Considerations School of Thought (CCSoT)

<i>NMSoT</i>	<i>CCSoT</i>
<ul style="list-style-type: none"> • Not necessarily human-specific • Based largely in the limbic system • Evolutionarily “older” brain regions and neural networks • Manifests as attitudes/behaviors that are fairly ubiquitous across contexts and domains • Most ubiquitous features in relation to race and politics: ingroup favoritism/group biases 	<ul style="list-style-type: none"> • Human-specific • Based largely in the frontal cortical regions • Evolutionarily “younger” brain regions and neural networks • Manifests as attitudes/behaviors that vary across contexts and domains • Most ubiquitous features in relation to race and politics: existential thought/ideological principles

To reiterate, the tendency for research to acknowledge one school of thought while neglecting the other is not universal and the distinction between the CCSoT and NMSoT is mainly illustrative, as many studies do indeed model the effects of both conscious and nonconscious motivations (e.g., Balzer and Jacobs 2011; Ditonto, Lau, and Sears 2013; Greenwald et al. 2009; Smith et al. 2011; Wagner et al. 2015). However, even work taking into account the direct effects of both conscious and nonconscious factors mostly fails to account for the interaction between them, and instead focuses primarily on comparing effect sizes to determine which exerts a stronger influence. The missing component here is the unique contribution of the processes that occur when automatic, often nonconscious factors are translated into conscious considerations, beliefs, and ideological principles.

The model I develop in this dissertation seeks to draw on both the CCSoT and the NMSoT in the context of race and politics, as there are clear implications for each school of thought in this literature. Most of the political science work on the topic stems from the CCSoT, focusing mainly on how historically imbued knowledge about racial conflict, consciously held racial stereotypes, and egalitarian values affect people's attitudes toward race-targeted government assistance. However, psychology work has been more likely to acknowledge NMSoT factors, and more recent political psychology research on race-related political attitudes has embraced the NMSoT in a variety of ways, such as comparing the effects of self-reported racial attitudes to those of implicit measures of racial attitudes when predicting an array of race-related political attitudes (e.g., Ditonto, Lau, and Sears 2013; Greenwald et al. 2009; Payne, Krosnick, Pasek, and Leikes 2010). However, all of this work nonetheless stops short of acknowledging that deep-seated motivations to favor ingroups, higher-order

thinking aimed at inhibiting those deep-seated motivations, and ideological principles such as a belief in limited government all interact with one another. How do people evaluate policies such as affirmative action for Blacks when their gut-level group biases, conscious beliefs about racial equality, and ideological values are all pulling in different directions? To rephrase the first question posed in this chapter, how much influence, if any, do the conscious beliefs and intentions highlighted by the CCSoT have over people's evaluations of race-related political objects in spite of the deep-seated motivations emphasized by the NMSoT?

The “Principles or Prejudice” Literature in Political Science

It is uncontroversial in the majority of academic literature that despite societal norms of racial tolerance and notable decreases in overt racism, racial prejudice still exists at substantial levels in the United States. Yet what has been less clear is how, exactly, such prejudice manifests itself in light of pressures toward political correctness and egalitarian values, and the degree to which prejudice subverts pressures toward political correctness and egalitarianism to influence policy opinions. A lengthy debate in political science has occurred over the past several decades regarding whether people oppose government assistance to Blacks because the recipients are Black or because such policies are perceived as violating individualistic values. Both sides of the debate fall mainly under the CCSoT, but the early work in this debate was conducted before the NMSoT was even beginning to be integrated into political science and more recent work has taken into account NMSoT factors to some degree.

CCSoT Work on “Principles versus Prejudice”

Symbolic racism was among the first and most influential concepts introduced to explain opposition to race-based policies as a function of covert racial prejudice that has developed in the public sphere as a reaction to decreasing social acceptability of explicit racism. It is hypothesized to be a blend of negative affect toward Blacks and conservative individualistic values, and consists of the beliefs that 1) racial discrimination is no longer a barrier for Blacks, 2) Blacks' disadvantages are due to their own reluctance to take personal responsibility for their lives, and 3) as such, efforts to aid Blacks are unwarranted (Kinder and Sears 1981; Sears and Henry 2003; Sears, van Laar, Carrillo, and Kosterman 1997). The same basic logic underlays the development of several other concepts such as *modern racism* (McConahay 1986), *racial resentment* (Kinder and Sanders 1996), and *laissez-faire racism* (Bobo and Smith 1998), among others. Proponents of this general viewpoint suggest that racial prejudice is the key catalyst of opposition to policies like affirmative action, and that prejudice and ideological principles like individualism cannot be entirely disentangled because contemporary individualism serves merely as a socially acceptable vehicle for expressing prejudice. According to Kinder and Mendelberg (2000), “today...prejudice is expressed primarily in the language of individualism; today individualism is part of racism” (p. 61). It is thus suggested that when the recipients of government assistance are Black, racial prejudice guised as individualism will lead to opposition to such assistance.

This viewpoint, however, was challenged on both methodological and theoretical grounds. Sniderman and colleagues primarily led the charge in proposing that individualistic values associated with conservatism are distinct from race-specific

attitudes and that when both are taken into account, nonracial individualistic values account for the majority of variance in race-based policy opinions (e.g., Sniderman and Carmines 1997; Sniderman et al. 2000; Sniderman and Tetlock 1986). The Symbolic Racism Scale (Kinder and Sears 1981; Henry and Sears 2002), according to Sniderman and colleagues, is a poor indicator of prejudice because it is confounded with conservative principles. Further, opposition to affirmative action based on individualistic principles is higher among conservatives, and conservatives tend to be relatively unaffected by whether the targets of a policy are Black or of some other particular group (e.g., Sniderman & Carmines, 1997; Sniderman et al., 1991; Sniderman et al., 1996). Accordingly, the reason for opposition to race-based policies is not the recipients' race by itself, but rather conservative, individualistic values. In other words, unless the target group of government assistance is perceived as violating individualistic values, which Blacks so often are perceived as doing, the race of the target group alone will not yield opposition to assistance.

Proponents of both sides of the “principles or prejudice” debate would fall largely under the umbrella of the CCSOT because the primary factors proposed to underlay race-based policy attitudes all involve consciously held beliefs. The difference between the two sides of the debate boils down to whether race-targeted policy support is driven by beliefs about discrimination and the work ethic of Blacks or beliefs about how involved government should be in ensuring group equality. Scholars engaged in this literature have examined the roles of ideological values, stereotypes, and beliefs but an account of stable personality traits, implicit preferences, or any other factor highlighted in the NMSOT remains absent throughout early work.

Principles or (Implicit?) Prejudice: Bringing in the NMSoT

Despite the lack of acknowledgement of NMSoT factors by most work on the “principles or prejudice” question, scholars studying racial attitudes more broadly have incorporated measures of deep-seated, automatic group biases into their understandings of prejudice for decades. The seminal work on *aversive racism* by Dovidio and Gaertner showed that although direct, explicit measures of racial attitudes suggest prejudice has decreased substantially since the Civil Rights era, when people have the opportunity to rationalize or explain away prejudiced attitudes or behavior, high levels of prejudice and discrimination are unveiled (Gaertner and Dovidio 1977; 1986; Dovidio and Gaertner 2000). Since this research was conducted, various indirect measures of racial attitudes have been developed to gauge racial preferences in a way that avoids influence from people’s conscious attempts to tailor or edit their responses. For example, it has been shown that physiological responses to members of racial outgroups independently predict racial policy attitudes (Smith et al. 2011), fear responses to males of one’s same race become extinct quicker than fear responses to males of another race (Olsson, Ebert, Banaji, and Phelps 2005), and activation in the startle and frown muscles in the face is greater when responding to racial outgroup members than racial ingroup members (Dambrun, Despres, and Guimond 2003; Vanman et al. 2004). The implication of these findings is that when it comes to racial attitudes, individuals have a motivation to edit or hide their opinions, and are often even unaware of their deep-seated preferences, and so simply asking for their opinions through survey self-reports can be incredibly problematic.

The primary workhorse when it comes to indirect measures of racial attitudes has been implicit association measures. As described earlier, implicit association

measures are an accessible and feasible means of gauging automatic biases with large-N survey and experimental samples. These measures are central to the NMSoT because they tap reflexive biases that are presumably unaltered by conscious deliberation and effortful thought (see Olson and Fazio 2009), and some work has even shown that implicit racial biases but not explicit racial biases are associated with activation in limbic regions of the brain such as the amygdala (Phelps et al. 2000). With regard to the relationship between racial attitudes and political opinions and behaviors, the use of implicit association measures has been critical. Anti-black implicit biases measured via the IAT have been shown to predict a range of political outcomes such as negative attitudes toward former President Barack Obama and opposition to the Affordable Care Act (Greenwald, Smith, Sriram, Bar-Anan, and Nosek 2009; Knowles, Lowery, and Schaumburg 2010). The IAT has also been used to show that Anti-Latino implicit biases predict conservative immigration attitudes (Pérez 2010; Pérez 2016). The AMP has also been instrumental in shedding light on how implicit prejudice influences political attitudes, as anti-Black implicit biases using the AMP have been shown to predict lower probabilities of having voted for Barack Obama in the 2008 presidential election, negative feelings toward Barack Obama as well as the idea of a Black president more broadly, and opposition to race-targeted policies such as affirmative action (Ditonto, Lau, and Sears 2013; Finn and Glaser 2010; Greenwald et al. 2009; Pasek et al. 2009; Payne, Krosnick, Pasek, and Leikes 2010; Segura and Valenzuela 2010).

Mediation of Implicit Measures by Explicit Measures

The research using implicit racial attitude measures to predict political outcomes has shown that automatic group biases predict political opinions in a

manner consistent with the broader tenets of the NMSoT. However, crucially, much of this work has also shown that the role of implicit attitudes is often mediated by the role of explicit attitudes (Ditonto, Lau, and Sears 2013; Finn and Glaser 2010; Payne, Krosnick, Pasek, and Leikes 2010), suggesting the political attitudes and behaviors political scientists are often interested in studying may be better predicted by measures of conscious, controlled processing than by measures of quick, reflexive preferences. This makes sense if the political variables being studied are conceptualized and measured in a way that they are influenced by the controlled processing that implicit measures are specifically designed to ignore, such as social desirability or consciously held stereotypes. In other words, CCSoT factors may matter more than NMSoT factors in predicting political outcomes because the psychological processes that occur when reporting political opinions or behaving politically are more similar to the processes that occur when reporting racial attitudes consciously than they are to the processes that underlay implicit racial attitude measures.

The work described above has not been interpreted in terms of the CCSoT or NMSoT, or in terms of “principles or prejudice” for that matter. Instead, this research has been mainly framed as a demonstration of the utility of implicit attitude measures compared to explicit attitude measures, or a methodological contribution to the study of prejudice. But is the full extent of these findings methodological, or are there substantial theoretical implications of the dominant role of explicit racial attitudes (rather than implicit racial attitudes) in predicting race-related political outcomes? If we are really only interested in public expressions of political opinions, do we really need to be concerned with implicit preferences? I suggest even in the case that the

political outcomes we are interested in are overt and explicit, implicit, automatic processes are still crucial. I suggest this work reflects the interaction between NMSoT and CCSoT factors.

From the viewpoints of social neuroscience and evolutionary psychology, it makes sense that measures of controlled processes such as explicitly reported racial preferences are more correlated with political outcomes than IAT or AMP scores because policy attitudes, vote choice, and self-reported racial preferences all presumably involve a relatively high level of deliberation and effortful thought. However, in this dissertation I suggest the interaction between automatic and controlled processes strongly influences the overt political expressions that people ultimately exhibit. Further, the work mentioned above on implicit racial attitudes and political outcomes has yet to incorporate implicit measures into the “principles or prejudice” framework. Specifically, how do ideological values, or principles, fit into the picture?

Group-Based Principles

There is some existing work in the race and politics literature that I suggest acknowledges the role of ideological values in conjunction with the interaction between automatic and controlled processes without necessarily knowing it. Work on Social Dominance Orientation (SDO; Pratto, Sidanius, Stallworth, and Malle 1994) has become a cornerstone of the literature on how group-based attitudes influence political opinions, and represents psychological motivations to endorse hierarchy in society and strive for a system in which some groups dominate others (Pratto, 1999; Pratto et al., 1994; Sidanius, 1993; Sidanius & Pratto, 1999; Sidanius et al., 1996; Sidanius, Levin, & Federico, 2001; but see more recent work on SDO as discussed in

Chapter 2). It has been offered as a theory to explain racial prejudice from a broad standpoint that acknowledges multiple levels of analysis and the various forces that ultimately create group-based inequality in society, including societal institutions that reinforce hierarchy, “real” conflict between groups over scarce resources, learned stereotypes about different groups, and psychological motivations to favor ingroups and derogate outgroups. Some work has even addressed the potential for evolutionary selection processes to drive individual differences in SDO (e.g., Pratto, Sidanius, and Levin 2006), although most SDO work does not put much attention on evolutionary factors. As such, the construct of SDO, though measured as a set of survey items, can be thought of as representing a combination of what would be referred to in this dissertation as CCSOT and NMSOT factors.

SDO research also addresses the role of ideological values. Individual differences in SDO are thought to correspond to differences in the “legitimizing myths” people subscribe to in order to rationalize or justify group-based inequality. Legitimizing myths are defined as “consensually held values, attitudes, beliefs, stereotypes, and cultural ideologies” (Pratto, Sidanius, & Levin 2006, p. 275). One of the most prominent legitimizing myths in contemporary society, according to Pratto, Sidanius, and colleagues, is individualism. Based on this hypothesized relationship between SDO and individualism, SDO has made its way into the conversation on “principles or prejudice.” It has been shown that SDO is highly predictive of race-targeted policy attitudes and also explains the relationship between principled arguments against race-targeted policies and racial prejudice (Sidanius, Pratto, & Bobo, 1996). Further, Federico and Sidanius (2002) provided evidence that principled

objections to race-targeted policies are more associated with group-based dominance than they are with nonracial individualism.

SDO may therefore reasonably be seen as emanating from the interaction between deep-seated group biases and higher-order, learned values such as individualism. Humans have been shown to be driven largely by automatic, often innate motivations to favor ingroups. However, such motivations are not expressed without being filtered through societal norms and institutions, and so higher-order “principles” – i.e. legitimizing myths, which can be hierarchy-enhancing or hierarchy-attenuating – interact with automatic group biases to form broad group-based belief systems, such as SDO. Thus, SDO is not purely an indicator of innate motivations to favor ingroups or higher-order ideological values, but rather it is the product of the interaction between the two. The implications of the role of SDO are therefore distinct from the implications of the roles of “principles” or “prejudice” separately. Unlike with “prejudice” alone, high levels of SDO imply support for group-based inequality in a way that does not necessarily favor the ingroup, but instead applies a uniform set of preferences across all groups. Yet unlike with “principles” alone, high levels of SDO also do not imply a lack of group-based discrimination, but instead entail support for a system in which some groups dominate others. I refer to constructs like SDO, which seem to lay at the intersection of innate group biases and higher-order values, as *group-based principles*. The central thesis of this dissertation is that it is group-based principles, which represent the interaction between automatic group biases and higher-order principles, rather than automatic group biases or ideological principles alone that drive opinions toward race-targeted government assistance.

Although it is uncontroversial that SDO reflects an interaction between group-based biases and ideological values, the SDO literature nonetheless does not engage with the distinction between automatic and controlled processes. Sidanius and colleagues draw on sociological models such as the *group positions model* (Blumer 1961), the *expectation states model* (Berger et al. 1970), the *realistic group conflict model* (Bobo 1983; 1988; Jackman 1991), the *racial oppression model* (Turner, Singleton, and Musick 1984), and *neoclassical hegemony models* (Gramsci 1976; Mosca 1939; Pareto 1943; Scott 1990) to postulate that:

(a) to one degree or another, almost all human societies are viewed as group-based hierarchies in which at least one dominant group enjoys a disproportionate share of positive social value (e.g., wealth, health, leisure time, education), and at least one subordinate group endures a disproportionate share of negative social value (e.g., social restrictions, poor health, low-status occupations, prison sentences). (b) At its core, politics can be viewed as an exercise in intergroup competition over scarce material and symbolic resources (e.g., wealth, high relative social status). (c) In this intergroup competition, groups will use ideological instruments such as notions of natural rights, national superiority, national destiny (e.g., manifest destiny, racism, the Protestant work ethic), and political ideology in an effort to legitimize each group's claims over these real and symbolic resources. (Sidanius, Pratto, and Bobo 1996, p. 477)

As such, the role of SDO in explaining the relationship between racial prejudice and individualistic values, as well as its role in predicting race-based policy opinions, is assumed by SDO theorists to emanate from perceived conflicts between groups and ideologies that are used deliberately to justify claims over contemporary real and symbolic resources. In other words, the roles of both group biases and ideological principles are thought to be conscious and intentional. Therefore, although the SDO framework identifies a more universal set of causes for race-based policy opposition than other work on the "principles or prejudice" question, it nonetheless pinpoints people's conscious beliefs about how modern groups in society should be

organized as the cause of race-based policy opposition. The goals of this dissertation therefore include not only establishing that group-based principles such as SDO underlay race-targeted policy attitudes, but also broadening the range of factors thought to underlay group-based principles by examining how automatic and controlled processes interact to influence these attitudes.

Evolution, Ingroup Bias, Conscious Control, and the Development of Group-Based Principles

Although this dissertation builds largely on the theoretical foundations of existing work on “principles versus prejudice,” which primarily stem from sociology and social psychology, one of its central contributions is fitting existing work into an evolutionary psychology framework. The reason for this is that the primary distinction between the automatic and controlled processes examined in this dissertation emanates in evolutionary theory – specifically, when the neural regions underlying automatic versus controlled processes are thought to have evolved and for what functions. Further, I suggest that the psychological processes driving the factors mainly focused on within the NMSoT are best characterized as evolutionarily “older” than the psychological processes driving the factors focused on within the CCSoT, yet these processes continually interact with one another. As is most often the case with work building on evolutionary models, the overarching hypothesis that the processes studied in this dissertation have evolved through natural selection is not directly tested. Nonetheless, placing the empirical tests of this dissertation in the context of evolution is essential because it is from an evolutionary perspective that hypotheses are derived. Many of the hypotheses tested in this dissertation would be substantially different or non-existent if not informed by evolutionary theory, because many of the

hypotheses are based on theories of where automatic biases and “higher-order” thought come from and what their adaptive functions are.

It is increasingly common among scholars studying political psychology for evolutionary theory to be used to derive hypotheses (e.g. Alford, Funk, and Hibbing 2005; Fowler and Schreiber 2008; Haidt 2012; Hatemi and McDermott 2011; Hibbing, Smith, and Alford 2014; Marcus 2013; Petersen 2014; Petersen and Aarøe 2012; Sidanius and Kurzban 2003). In the following section, I summarize evolutionary theory on the selection processes thought to underlay automatic versus controlled processes and their interaction. In doing so, I elaborate on the difference between automatic and controlled processes as well as clarify how *both* emanate from the process of natural selection.

The “Lizard Brain” in Politics

All social primates, including humans, exhibit behaviors that might reasonably be called “politics” (Alperson-Afil et al. 2009; Brown 1991; Buss 2005; de Waal 1996). In fact, group conflict, including within- and between-group conflict, has been around for millions of years and likely predates the appearance of *Homo sapiens* and tracks back to small-scale group living arrangements among human ancestors (e.g., Boehm 2000; Buss 2005; Wrangham and Peterson 1996; Petersen 2014; Petersen and Aarøe 2012). In other words, to a nontrivial degree, the political and group-based conflict we see today – clashes between national or subnational coalitions, disputes over resources, trade, rules enforcing hierarchy, punishments for rule breakers, and decisions about how resources are distributed between groups within countries – are hardwired consequences of being social animals.

The “lizard brain,” or “reptilian brain,” as some have referred to it (e.g.

LeDoux 1998; Lieberman 2009; MacClean 1973), consists of the millennia-old, core psychological adaptations humans carry with them as they navigate the contemporary world, such as the “fight or flight” system of the autonomic nervous system (ANS) or motivations to eat food and fornicate. The reference to reptiles stems from the fact that a reptile brain consists only of the most evolutionarily old brain regions necessary for survival. As such, the use of the term to reflect social processes is somewhat misleading, as the social processes discussed here, although evolutionarily old and basic, are not shared by reptiles. Nonetheless, the phrase has been used often simply to delineate between primal, innate motivations and higher-order executive functions that are unique to humans.

A major difference between the group conflict that took place millennia ago and the group conflict of current-day politics is the size and scope of group interactions. Whereas the group conflict of the vast majority of human history took place within and between small groups, clans, and tribes of 25-200 people (Kelly 1995), contemporary politics spans a much broader range of levels of analysis, including small groups just as before but also including national and international groups made up of millions (and often billions) of individuals who most often do not know each other. Importantly, mass-scale society has not been around until relatively recently in human history (within the last 500 years; Diamond 1998), and so the most powerful and ubiquitous adaptations humans have for interacting with one another in mass-scale society evolved in response to small-scale interactions (Petersen 2014; Petersen and Aarøe 2012; Tooby and Cosmides 1992). As such, the psychological tools humans bring to the table when interacting with today’s political world were not “designed” for mass-scale politics. This implies political attitudes and behaviors are

largely driven by the momentum of millions of years of evolutionary forces that we are not only unaware of but also blind to their purpose. Much NMSoT work is grounded in these principles of evolutionary theory.

Political scientist Michael Bang Petersen and colleagues have done groundbreaking research exploring the political implications of the mismatch between mass-scale society and the (often “outdated”) psychological tools humans have to deal with it. By deriving hypotheses from evolutionary theory, Petersen and others have developed studies showing a range of empirical findings that might seem quite strange if not understood from the perspective that political conflict is based in the small-scale group conflict of human ancestors. For example, in one experiment, participants who consumed a sugary soft drink were less likely to support social welfare than participants who consumed a drink with artificial sweetener, which makes sense if in small-scale group living situations, humans were more likely to rely on sharing a key resource, food, to survive during periods of temporary hunger (Aarøe and Petersen 2013). In another study, it was shown that the flexed bicep circumference of a man’s dominant arm – a key morphological indicator of upper-body strength – is associated with support for economic redistribution among low-SES males and opposition to economic redistribution among high-SES males (Petersen et al. 2013). According to Petersen and colleagues, this is because historically, in small-scale group living situations, the physically formidable males were the ones most able to acquire and defend resources in line with their self-interest, and so it should be expected that in contemporary society these individuals will be most likely to support a political system that benefits their own interests – i.e. high levels of economic redistribution among low-SES males and low levels of

economic redistribution among high-SES males. The overarching implication of the model developed by Petersen and colleagues is that most (if not all) of the group-related political attitudes we like to think of as the product of conscious considerations, beliefs, values, and the weighing of costs and benefits (i.e. CCSoT factors) are actually driven by primitive, evolved psychological tools designed for small-scale group interactions (i.e. NMSoT factors).

One of the most dominant adaptations evolved from small-scale group interactions that still influences political attitudes and behavior today is ingroup favoritism (e.g., Arceneaux 2012; Brewer 1999; Fiske 2000; Petersen 2012; Yamagishi and Mifune 2004). Indeed, for decades, psychologists have found the tendency for individuals to express bias in favor of members of their perceived ingroups to be a powerful and ubiquitous human phenomenon (e.g., Brewer 1986; Hogg and Abrams 2007; LeVine and Campbell 1972; Tajfel and Turner 1979; Turner et al. 1981). From an evolutionary perspective, ingroup favoritism has been adaptive because it allows individuals to rely on others without indiscriminately trusting everyone (Brewer 1997; 1999; Insko, Schopler, and Sedikides 1998; Tooby, Cosmides, and Price 2006). As Marilyn Brewer puts it:

Social differentiation and clear group boundaries provide one mechanism for achieving the benefits of cooperative interdependence without the risk of excessive costs. Ingroup membership is a form of contingent altruism. By limiting aid to mutually acknowledged ingroup members, total costs and risks of nonreciprocation can be contained (1999, p. 433).

Thus, as long as humans have been living socially, it has been strongly adaptive to demarcate individuals along group lines and favor ingroup members over outgroup members. Ingroup favoritism is therefore substantially hardwired into the human brain. Importantly, which individuals comprise ingroup members versus

outgroup members is decidedly *not* hardwired. Instead, the boundaries of group membership are fluid and context-dependent. Groups are *perceived*, and so the individuals people identify as members of their ingroup can depend highly on whatever demarcating characteristic of individuals is salient in the current environment (Efferson, Lalive, and Fehr 2008; Kurzban, Tooby, and Cosmides 2001; Tajfel and Turner 1979; Turner et al. 1981). As such, the tendency to favor one's ingroup is not bound to particular types of groups or periods of time; ingroup favoritism can manifest in any situation involving multiple individuals. If the previously stated premise is true – that humans engage in 21st century politics with prehistoric minds – we should expect most, if not all, of the political conflicts and disputes over resources that characterize contemporary politics to be driven primarily by sheer ingroup favoritism.

Perhaps unsurprisingly, ingroup favoritism and social identity motives have been shown to underlay a tremendous range of politically-relevant attitudes and behaviors (Fowler and Kam 2007; Green, Palmquist, and Schickler 2004; Greene 2004; Huddy 2001; Huddy and Khatib 2007; Iyengar, Sood, and Lelkes 2012; Kinder and Kam 2009; Mutz and Kim 2016), including attitudes toward affirmative action (Lowery et al. 2006). Further, it is evident that ingroup bias is automatic and occurs at a nonconscious level (Cunningham et al. 2004; Molenberghs et al. 2013; Van Bavel, Packer, and Cunningham 2008; Volz et al. 2009). The implication of these findings in conjunction with an evolutionary model of ingroup favoritism is that group biases are instantiated deep enough into the human psyche that they are triggered well before anything reasonably called a “conscious consideration” can take place.

Despite the predominance of group biases, there are also many instances in which ingroup favoritism does not manifest, or is even the exception rather than the rule. In various circumstances, many people embrace group equality, diversity, and globalism while rejecting hierarchy, racism, and nationalism. Further, ingroup favoritism is in many cases the antithesis of commonly held egalitarian values, and humans seem capable (at least sometimes) of adhering to such values; it is not uncommon for individuals to be somewhat self-aware regarding the pitfalls of ingroup favoritism. How is this possible? The picture painted by considering ingroup favoritism alone – a world where who’s team you are on is the defining characteristic of you as a person – seems partly accurate but does not leave any room for genuine principles such as egalitarianism, or individualism for that matter.

Controlled Processing, Rules, and Higher-Order Thought

To understand how primal motivations such as ingroup favoritism might be overridden, it is necessary to understand the development of the inhibition systems of the brain. Alongside motivations to form coalitions and delineate between ingroup and outgroup members automatically, humans have also evolved to be able to regulate and inhibit their primal motivations – or as Kahneman (2011) would put it, “think slowly.” Whereas automatic group biases evolved to allow individuals to quickly distinguish between trustworthy and untrustworthy individuals, “controlled” processes – also often called “executive function” – evolved so that individuals could interact with and respond to novel environments in a finer-grained, albeit slower fashion. Controlled processes allow individuals to refrain from applying blanket rules to every situation (e.g. “always favor ingroup members over outgroup members”), and instead to push an inhibitory lever that adjusts decision-making based on

circumstances (e.g. “*sometimes* favor ingroup members over outgroup members, but sometimes don’t”; see Tooby and Cosmides 1992). To put it simply, controlled processes let individuals adapt to complex environments by allowing nuance in the degree to which more primal “rules” for behavior are followed.

Consistent with this understanding of the role of controlled processes, the development of controlled processes throughout the lifespan is thought to be central to how humans learn and abide by increasingly complex rules. Specifically, the development of regions of the brain associated with controlled processing – e.g. the prefrontal cortex, or PFC – is believed to correspond to increasing complexity in rule usage by humans such as going from using a single rule, to switching between rules, to switching between pairs of incompatible rules (Bunge and Zelazo 2006; Zelazo, Frye, and Rapus 1996; Zelazo, Muller, Frye, and Marcovitch 2003). It is through activity in the PFC, in particular, that humans are able to engage in higher degrees of self-reflection, or “levels of consciousness” (LOCs), and inhibit automatic reactions to environmental stimuli (Kerr and Zelazo 2004; Ortner, Kilner, and Zelazo 2007; Zelazo 2004). If we consider, for a moment, higher-order beliefs and principles to be complex sets of rules or guidelines for shaping behavior, it makes sense that whereas “lizard brain” processes might be more responsible for sheer ingroup favoritism in humans, lofty values such as individualism and egalitarianism may be driven more by executive function and controlled processes.

Interactions between Controlled and Automatic Processes

Up to this point, I have outlined the “older” evolutionary forces such as ingroup favoritism that are ubiquitous in human behavior today, as well as the “younger” evolutionary forces that allow humans to inhibit primal motivations and

abide by more complex rules and principles. However, it would be misleading to suggest that this means there are essentially two major forces working against one another and that one represents variables such as ingroup favoritism and the other represents variables such as individualism and egalitarianism. A central argument of this dissertation is instead that these forces continually interact with one another to drive the full range of attitudinal variables, including ingroup favoritism, ideological principles, and ultimately attitudes toward race-based government assistance. All of these variables exist on a spectrum from “more influenced by automatic processes” to “more influenced by controlled processes.” Where attitudes exist on this spectrum may be more difficult to decipher than some might imagine, as evidence exists suggesting automatic and controlled processes interact constantly.

The most basic manifestations of controlled processing, such as directing attention, coordinating motor functions, and accessing memory, are not unique to humans and are not all that “young” evolutionarily. Rudimentary controlled processes are evident in monkeys (Fuster 1989) as well as rats (Kolb and Tees 1990), and the PFC is thought to have evolved at least 175 million years ago when the ancestors of all present-day mammals first appeared (Jerison 1997), which is “younger” than most brain regions associated with the “lizard brain” but still not recent by any means. In other words, if we are looking for the evolutionary basis of higher-order principles such as individualism or egalitarianism, it is unlikely to be found simply in the appearance of the PFC in human evolution. Further, contrary to much prior thinking (e.g. Deacon 1997; Holloway 1998), the human PFC is not proportionately larger than that of other mammals (Jerison 1997; Semendeferi, Damasio, Frank, and Van Hoesen 1997; Uylings and Van Eden 1990), suggesting the uniqueness of controlled

processing in humans – the set of neural processes that underlay existential thought, ideals, and belief systems – is not due to the mere existence of certain brain regions absent in other species, but instead is most likely due to *connectivity* within and across areas of the brain (de Schotten, Dell’Acqua, Valabregue, and Catani 2011; DeFelipe, Alonso-Nanclares, and Arellano 2002; Fuster 2002; Holloway 2002; Ramnani et al. 2006). That is, the unique executive functions of the human brain have likely evolved because “selectional pressures that bring about brain evolution do so by acting on interconnected systems rather than simply on individual brain areas” (Ramnani et al. 2006, p. 817). Humans have evolved so that quick, automatic and nonconscious motivations are in constant communication with controlled, reflective thought.

The Iterative-Reprocessing (IR) model of the neural bases of evaluation provides a roadmap for understanding how automatic and controlled processes continually interact with one another (Cunningham, Haas, and Jahn 2011; Cunningham and Zelazo 2007; Cunningham, Zelazo, Packer, and Van Bavel 2007). According to the IR model, environmental stimuli trigger an iterative evaluative process between automatic and controlled processes in the brain whereby the stimuli are interpreted repeatedly with regard to increasingly complex representations (Cunningham and Zelazo 2007). Critically, as Cunningham and Zelazo note, “Whereas evaluations that are based on few iterations of the evaluative cycle are relatively automatic, in that they are obligatory and might occur without conscious monitoring, evaluations based on additional iterations and computations are relatively reflective” (p. 97). Time is therefore a crucial element in this model because evaluations become more “complex” as controlled processes reinterpret, or reprocess, the stimuli.

Bringing it back to Race-Targeted Policies

The implication of the IR model for understanding the variables of interest in this dissertation is therefore fairly straightforward. The range of factors thought to drive opinions toward race-targeted government assistance, from primal group biases to lofty ideals about individualism and the proper role of government, exists entirely on a spectrum from “more influenced by automatic processes” to “more influenced by controlled processes”, and no factor exists solely as a function of automatic or controlled processes alone. In line with the growing scholarship that acknowledges both NMSoT and CCSoT factors, I suggest the variables most influential in driving support for race-based government assistance are neither completely automatic nor entirely conscious. As such, the existing “principles or prejudice” literature, which is predominantly based in the CCSoT but has recently embraced some aspects of the NMSoT, is theoretically limited. Instead, I suggest race-targeted attitudes are driven primarily by the interaction between automatic and controlled processes. I propose that group-based principles directly reflect this interaction because they reflect the translation of innate group biases into broad guidelines for thinking about contemporary politics – i.e. principles. I suggest group-based principles will be the dominant predictor of these race-targeted policy attitudes because such policies exemplify the intersection between primal group-based motivations and ideological principles. Critically, I propose the mechanism of the dominant role of group-based principles is the iterative process that occurs when automatic reactions interact with controlled processes. As individuals have more time to evaluate race-based policies, their evaluations should transition from being based on automatic group biases to being based on broad principles.

Caveats

The distinction between automatic processes and controlled processes is crucial in this dissertation, and so it is worth elaborating. Basically, the distinction between automatic and controlled processes reflects the distinction between reflexive, often nonconscious neural processes that occur primarily in the limbic system and conscious, controlled processes that occur primarily in the frontal cortical regions of the brain. Whereas automatic processes are evolutionarily older and motivate individuals primarily toward basic survival, controlled processes are evolutionarily younger (although still quite old) and serve to allow individuals to interact with the complexities of the contemporary world. Automatic and controlled processes are both thought to underlay group biases *and* higher-order thought (i.e. ideological principles and values) to some degree, although group biases are theorized as emanating more from automatic processes whereas higher-order thought is theorized as emanating more from controlled processes. Below, I explain several important caveats regarding what the distinction between automatic and controlled processes, as conceptualized in this dissertation, does *not* represent.

1. *“Controlled” does not mean “superior.”* One way some readers might interpret the distinction between automatic and controlled processes is as suggesting controlled processes are “better” than automatic processes because they are more deliberate and aimed at accuracy. However, this would not be an accurate characterization because as explained above, both are integral to human survival, and both underlay even the most “reasoned” and “logical” thoughts humans are capable of to some degree. Further, automatic processes

are essential for making quick decisions and motivating individuals toward goals (see e.g., Eagleman 2011; Kahneman 2011).

2. *Controlled processing does not mean “controlled” in the metaphysical sense.*

The use of the term “controlled” in this dissertation is not meant to reflect a role of “free will.” As explained above, controlled processes are conceptualized as products of natural selection just as automatic processes are, and so the term “controlled” is used simply to denote the more complex processes of the frontal cortical regions of the brain, which individuals can be consciously aware of.

3. *The distinction between automatic and controlled processes is not simply a*

distinction between implicit and explicit attitudes. Although much work has used implicit and explicit attitude measures to tap automatic versus controlled processes, and indeed such measures are used several times for this purpose in this dissertation, these measures are not pure indicators of automatic or controlled processes. In line with the IR model described earlier, automatic and controlled processes are both thought to influence implicit and explicit attitude measures; the difference is the *degree to which* each measure is influenced by each type of process.

4. *In line with the above caveat, the distinction between automatic and*

controlled processes is not simply a distinction between quick reactions and evaluations made over a longer period of time. Although it is theorized, in line with the IR model, that time is a critical factor in determining the degree to which controlled processes play a role in evaluation, even the quickest evaluations (including those made in less than a second) are believed to be

influenced somewhat by controlled processes (Cunningham, Haas, and Jahn 2011). As such, although time plays a central role, the primary distinction is between regions in the brain rather than time.

5. *The distinction between automatic processes and higher-order thought is not a distinction between processes based in evolution and processes not based in evolution.* As explained earlier, both automatic and controlled processes are thought to have evolved millions of years ago, and although the “higher-order thought” examined under the CCSoT (i.e. principles, ideological values, and existential thought) is human-specific, controlled processes are not because they can be found in all mammals to some degree.
6. *The distinction between automatic and controlled processes is not due to when humans began interacting in mass scale society.* As explained earlier, mass-scale society has only been around for approximately 500 years (Diamond 1998) whereas both automatic and controlled processes have been around for millions of years. In that sense, this dissertation is not at odds with the work of Petersen and colleagues on the mismatch between contemporary politics and evolved human psychology. However, it is proposed in this dissertation that controlled processes are what allow humans to apply primal motivations to the complex political world of today.
7. *The focus of this dissertation is not simply an interaction term in a regression equation.* The interaction between gut-level processes and higher-level thinking can be operationalized in many different ways and is not simply an interaction between two particular variables in a regression. The idea is to acknowledge the iterative process by which higher-level “principles” are built

from the iterative reprocessing that occurs between limbic and cortical regions of the brain. This might be an interaction between two variables aimed at gauging these processes, or it might be a single variable that captures the combination of these processes.

Chapter Outline

In the remainder of this dissertation, I present empirical findings aimed at shedding light on different aspects of the model outlined in this chapter. First, in Chapter 2, I take a step back and examine the multitude of constructs that may embody the group-based and higher-level factors influencing attitudes toward race-related political outcomes. I do so using only self-report measures to survey items. Specifically, I conduct a survey experiment through Amazon's Mechanical Turk to examine the degree to which race-targeted policy attitudes are related to group biases, individualistic values, or group-based principles. To examine the role of group biases, I use survey items tapping social identity motives – i.e. identification with ingroups. To gauge the role of ideological principles and values, I measure adherence to individualistic values such as work ethic and self-reliance. Finally, to gauge group-based principles, I measure SDO. I find that race-targeted policy attitudes are explained mostly by preferences for group-based social hierarchy (the “Equality” subdimension of SDO). Further, manipulating the work ethic of Whites versus Blacks has no effect on policy attitudes. These findings suggest race-targeted policy opposition is not purely “principled” but also does not simply boil down to ingroup favoritism. Rather, group-based principles seem to be the primary predictor of race-based policy attitudes. However, these results are constrained to policies assisting

Blacks only, and so it is impossible to know if the role of SDO-E is masking ingroup favoritism or attitudes specific to Blacks.

Therefore, in Chapter 3, I address the concern that social desirability might account for the seemingly negligible role of ingroup favoritism by utilizing the Minimal Groups Paradigm (MGP), in which participants are divided into arbitrary, meaningless groups within a laboratory setting. This is crucial because the methods used up to this point have likely been influenced by the desire of respondents to provide socially acceptable answers that indicate racial tolerance or egalitarianism. To tap the role of ingroup favoritism more directly, I strip away the “priors” people have about specific racial groups by having them interact with meaningless groups for which societal pressures are minimal. In this way, the MGP allows me to look at the conditions under which race-specific attitudes and stereotypes are absent. I also employ an implicit attitude measure to tap automatic processes underlying evaluations of the minimal groups. However, I am still able to introduce opportunities for individuals to rely on broad individualistic principles. To do so, I manipulate the work ethic of the minimal groups so as to vary the degree to which participants are cued to evaluate the groups based on “merit” rather than ingroup/outgroup distinctions. By looking at the degree to which ingroup favoritism versus merit-based evaluation in the MGP predicts real-world race-targeted policy attitudes, I can gauge the degree to which the “principles” people use to judge race-based policies are indeed principled as well as the degree to which such judgments are based in ingroup favoritism. I find that the only significant predictor of real-world race-based policy attitudes is *implicit* ingroup favoritism in the MGP. I also find that group-based principles are significantly associated with both ingroup favoritism (implicit and, to some degree,

explicit) and merit-based evaluation. These results suggest that a) the role of individualistic principles in predicting race-based policy opinions is inherently tied to group-based evaluations, b) the role of ingroup favoritism in evaluations of race-based policies is based primarily in automatic rather than controlled processes, and c) group-based principles embody some combination of automatic group biases and conscious, “principled” responses to environmental cues.

In the chapters thus far, I will have rigorously tested the degree to which race-based policy opinions can be declared “principled” versus “group-based,” and concluded that the answer is *both* – i.e. group-based principles play the dominant role in predicting race-based policy opinions, principles that are not group-based seem to be negligible (except for the role they play in influencing group-based principles), and the role of ingroup favoritism seems to emanate from automatic processes. In Chapter 4, I turn my attention to further drilling down on the role of automatic versus controlled processes. I demonstrate the utility of acknowledging the interaction between automatic preferences and controlled processes by examining the impact of implicit racial ambivalence – or discrepancies between implicit and explicit racial attitudes – on a variety of race-related political outcomes. Implicit racial ambivalence is not necessarily a direct operationalization of the interaction between automatic and controlled processes, but it nonetheless reflects the gap between individuals’ conscious considerations (“more controlled” factors) and implicit preferences (“more automatic” factors). Using the 2008 ANES data, I show that individual differences in implicit racial ambivalence regarding Whites and Blacks have a substantial impact on a range of race-related political outcomes, and that implicit racial ambivalence moderates the influence of higher-order principles such as egalitarianism and political

ideology on these outcomes. Therefore, the interaction between automatic and controlled processes seems not only to influence race-based policy attitudes directly, but also to moderate the role of higher-order principles. Returning briefly to the data used in Chapter 3, I demonstrate further that implicit ambivalence is negatively associated with ingroup favoritism.

In Chapter 5, I directly examine the neural processes that have thus far been assumed to underlay the roles of automatic and controlled processes by using fMRI. By using fMRI, I am able to determine the degree to which brain regions associated with controlled processing such as the prefrontal cortex (PFC) and anterior cingulate cortex (ACC) relate to racial biases in support for government assistance as well as reliance on group-based principles. In the experiment, I introduce participants to hypothetical individuals applying for government assistance and ask whether they support or oppose assistance to that individual. I experimentally manipulate the race of the applicant as well as whether the applicant's race is known consciously or only primed subliminally. This way, I can compare what happens when individuals are only implicitly made aware of the race of the target of government assistance versus when individuals are consciously aware of the target's race. According to the model laid out in this dissertation, we should see greater activity in the cortical regions of the brain (the PFC and ACC, to be specific) when individuals are consciously aware of the applicant's race compared to when they are only nonconsciously aware of the applicant's race. Further, individuals' evaluations should fall more in line with their broader ideological principles (specifically, SDO) when race is consciously known than when race is only subconsciously known because controlled processes can then exert greater influence. I find increased pro-Black biases in support for government

assistance in trials where race is consciously primed, and these increased pro-Black evaluations are associated with being low in SDO-D, but not SDO-E. More centrally for the purposes of this dissertation, I also find that brain activation patterns in the ACC and PFC largely behave as expected, although primarily with regard to the orbitofrontal cortex (OFC) and SDO-E.

Finally, in Chapter 6, I summarize the empirical findings of the previous chapters and assess the implications of these findings for the overarching framework outlined in this dissertation. Specifically, I summarize the results of my empirical investigations and discuss how the findings of the studies presented suggest the primal, gut-level factors studied under the NMSoT necessarily interact with controlled processes in order to yield higher-order principles and attitudes toward politics. The model developed in this dissertation holds substantial implications for how we understand a range of popular concepts, such as social desirability bias, symbolic racism, and system justification. With regard to race-based policy attitudes, specifically, these findings hold substantial implications for advancing and in some cases revising the “principles or prejudice” literature. Importantly, these findings also hold implications for predicting the utility of interventions designed to reduce the role of racial prejudice in contemporary politics.

CHAPTER 2

Disentangling “Principles” and “Prejudice”: Ingroup Favoritism, Individualism, and Group-Based Principles

“In the social scientific imagination, it is as if the advantaged are relentlessly looking to cash in on their dominance and the disadvantaged are proud revolutionaries-in-waiting. Both types of groups are seen as primarily self-interested, and overt conflicts of interest are assumed to be endemic.”

- John T. Jost, Mahzarin R. Banaji, and Brian A. Nosek²

Disputes over the legitimacy of policies like affirmative action are often framed as emanating from principled disagreements over the proper role of government, with those opposing such policies arguing the government should not be selectively assisting *any* person, regardless of their group memberships. Others argue such justifications for opposing race-targeted policies are really just rationalizations for deep-seated racial prejudice or group biases. As outlined in the previous chapter, this conflict is reflected in the mixed empirical research on the topic, with evidence building in support of the roles of both prejudice (e.g. Henry and Sears 2002; Kinder and Sears 1981; Sidanius, Pratto, and Bobo 1996) and individualism (e.g. Sniderman and Tetlock 1986; Sniderman et al. 1996) depending on how individualism and prejudice are operationalized. I suggest these mixed findings are partly due to a critical source of confusion: we don't fully understand how the multi-faceted constructs of individualism and group-based attitudes relate to one another, and how they interact to yield attitudes toward race-targeted policies. In other words, existing work in the “principles or prejudice” literature has yet to thoroughly define what the roles of “principles” or “prejudice” represent, let alone make a distinction between ingroup favouritism and group-based principles.

² Jost, J. T., Banaji, M. R., & Nosek, B. A. (2004). A decade of system justification theory: Accumulated evidence of conscious and unconscious bolstering of the status quo. *Political Psychology*, 25(6), 881-919.

Existing work has predominantly evaluated whether individualism or prejudice underlay race-based policy attitudes by gauging the role of single attitudinal constructs or testing whether or not racial biases exist in support for government assistance policies. However, it remains unknown which sorts of individualistic values underlay the role of individualism, as well as whether the role of prejudice conveys ingroup favoritism or broad group-based ideologies. If the central thesis of this dissertation – that group-based principles rather than automatic group biases or individualistic principles alone drive opinions toward race-targeted government assistance – then the distinction between ingroup favoritism and group-based ideologies is a critical one. Further, how should we be thinking about “individualistic principles”? Are such principles most likely to manifest as a blanket opposition to government assistance to anyone, or do such principles allow for some to receive assistance so long as they are “deserving”? Shedding light on these questions is a critical first step in this dissertation, as it allows us to move beyond the question of *whether* principles or prejudice influence race-based policy attitudes and toward the question of *how* each influences such attitudes.

There are two primary goals of this chapter. First and most critically, this chapter unravels the often unclear and overlapping conceptualizations of the dimensions making up individualism and group-based preferences that beset the existing literature on race-based policy attitudes. Three primary categories of factors are examined: individualistic values, social identity motives, and group-based ideologies. I refer to “social identity motives” instead of “ingroup favoritism” because in this chapter, I simply examine the role of beliefs and motivations related to people’s ingroups rather than constructing some measure of bias against other groups.

I do so in order to directly gauge individuals' ingroup attachments rather than relying on a measure reflecting some manifestation of those attachments. The social identity literature has been key to understanding the functions of these attachments and motivations (e.g. Hogg and Abrams 2007; Tajfel and Turner 1979; Turner et al. 1981).

Each of these three primary dimensions are broken up further into several components. Individualism, in particular, is often assumed to be unidimensional. This study acknowledges distinct aspects of individualism such as self-reliance and work ethic. Regarding "prejudice," research has yet to disentangle whether social identity motives or group-based ideologies best represent the role of prejudice in predicting race-based policy attitudes. Further, work has shown both Social Dominance Orientation (SDO; Pratto et al. 1994) and ingroup identification are multidimensional (Ho et al. 2012; Roccas et al. 2008). This study pits the roles of individualism, social identity, and social dominance – as well as the multiple dimensions comprising each – against one another to shed light on which ideological and group-based processes are most central to race-based policy attitudes.

The second goal of this chapter is to test the proposition that group-based attitudes underlay the effects of individualism on race-targeted policy attitudes while acknowledging the multi-dimensional nature of both group-based attitudes and individualism. Prior work has shown the effects of SDO on race-targeted policies are often mediated by individualism-related constructs (Federico and Sidanius 2002; Sidanius, Pratto, and Bobo 1996), suggesting "principled" objections to policies like affirmative action are more "prejudiced" than they claim to be (also see, e.g. Reyna et al. 2005). This assertion has significant implications for the questions tackled in this

dissertation because if the role of individualism is entirely explained by group-based motivations, it is strong evidence in favor of the dominant role of NMSoT factors over CCSoT factors, and contrary to what I propose in this dissertation, we may not need to account for higher-order principles at all when predicting race-based policy attitudes. However, it remains unknown which aspects of group-based attitudes, whether a particular dimension of SDO or some aspect of group identity, explain which aspects of individualism (e.g. self-reliance, work ethic) and thus account for this mediation. Thus, if group-based factors do underlay the role of individualism, we still do not know whether this is because of innate motivations to favor ingroups or group-based principles, which entail some degree of higher-order thought.

Competing Conceptualizations of Individualism

The traditional conservative stance on government assistance is often summed up as a belief in “pulling yourself up by your bootstraps.” But what does it mean to pull oneself up by one’s bootstraps? Is the value being placed on the fact that people are doing it on their own—independently and without help? Or, does the intrinsic value of such an act hinge on the hard work that inevitably goes into overcoming an obstacle? Ambiguity surrounding the concept has plagued the existing literature, with measures of individualism ranging from questions roughly approximating the Protestant Work Ethic (Furnham 1990; e.g. Kinder and Mendelberg 2000) to sheer self-reported political conservatism (e.g. Sniderman et al. 1996). Despite some grappling with this conceptual problem (e.g. Sniderman, Crosby, and Howell 2000), we still lack an empirical investigation of how to best operationalize individualism as it pertains to race-targeted policy attitudes. In this chapter, I acknowledge the potentially distinct roles of two factors underlying individualism: work ethic and self-

reliance. In both cases, these are consciously held values consistent with what is called higher-order thought in this dissertation, but there are important differences in the implications associated with each. If the role of individualism is explained by self-reliance, higher-order thought manifests in the formation of race-based policy opinions as equal rejection of government assistance to individuals regardless of group membership. However, if the role of individualism is explained by work ethic, higher-order thought should be conceptualized as reliance on deservingness cues to determine who may receive government assistance.

A common conceptualization of individualism is the Protestant Work Ethic, or the belief that hard work is the key to success. According to some, adherence to such a work ethic is thought to “fuse” with racial prejudice to form *symbolic racism* as explained in the previous chapter (Kinder and Sears 1981; Henry and Sears 2002; Rabinowitz, Sears, Sidanius, and Krosnick 2009; Sears and Henry 2003; Sears et al. 1997). Work ethic is posited to be inherent to constructs like symbolic racism, in which Blacks are seen as not sufficiently hardworking to warrant assistance. Kinder and Mendelberg (2000) utilize a measure of individualism developed by Feldman (1988), which closely approximates the values related to work ethic, and find that symbolic racism outperforms sheer individualism in predicting race-based policy attitudes.

The lack of conceptual clarity regarding this brand of individualism has been identified by Sniderman, Crosby, and Howell (2000) in a critical analysis of *symbolic racism*. According to Sniderman et al., the items used in the symbolic racism scale “have nothing to do with individualism” (p. 243). Sniderman et al. defer to an Emersonian view of individualism as an ethic of self-reliance in one’s own life. As such, individualism is conceptualized as a high valuation of independence regardless

of how hardworking an individual is. They proceed to use political ideology as a more apt proxy for individualism as it represents preferences for limited government and thus the degree to which people prefer a society unassisted by government intervention. They find preferences for limited government to outperform racial attitudes in predicting policy attitudes.

Yet even if one agrees with the proposition that symbolic racism muddies the concept of individualism, is Emersonian self-reliance the “true” individualism that we should be concerned with, and is political ideology a sufficient proxy for such individualism? Further, there is still ambiguity regarding whether the key aspect of individualism underlying race-targeted policy attitudes is self-reliance or a high valuation of hard work. Kinder and Mendelberg (2000) explicitly discuss individualism in terms of both Tocqueville’s ideas about the American citizen being isolated from large-scale society (p. 45) and Weber’s ideas about work ethic (p. 46). However, no work to date explicitly sets out to disentangle the effects of these alternative notions of individualism.

Perhaps most concerning is the fact that the measures of political ideology used in existing work constitute a wide range of factors beyond preferences for limited government, some of which are explicitly related to group-based attitudes (e.g. Graham, Haidt, and Nosek 2009; Jost et al. 2003). In other words, research that uses ideology measures as proxies for individualism are likely conflating individualism with a variety of other factors that may include the very group-based attitudes that the role of individualism is being compared to.

Social Dominance, Group Identity, and the Multiple Facets of Each

As with individualism, the role of “prejudice” in existing literature has also

been construed quite broadly. This is because the primary goal of existing research has been to determine whether people's attitudes toward race-based policies are due to racial biases or nonracial principles. As such, this literature has gauged the relative influence of individualism versus prejudice by testing the explanatory power of scales measuring racial prejudice against scales measuring individualism (Bobo 1998; Henry and Sears 2002; Kinder and Mendelberg 2000; Sears and Henry 2003; Sniderman, Crosby, and Howell 2000) or by examining whether biases exist in support for government assistance to racial minority groups compared to nonracial groups (Kinder and Mendelberg 2000; Kuklinski et al. 1997; Reyna et al. 2005; Sears et al. 1997; Sniderman et al. 1996; Sniderman, Crosby, and Howell 2000; Sniderman et al. 1991). Although these efforts have made tremendous headway in advancing understandings of how racial attitudes compete with individualism in predicting race-related political outcomes, in order to more precisely determine the specific mechanisms of these relationships, it is necessary to break down "prejudice" into its constituent parts. This study considers two distinct aspects of group-based attitudes that might reasonably comprise the role of "prejudice" yet have unique implications for understanding how, exactly, group-based preferences influence policy attitudes.

One component of group-based attitudes examined in this study is the set of beliefs individuals have about how groups in society should be organized. Such beliefs, defined in the previous chapter as *group-based principles*, are distinct from social identity motives because they do not necessarily entail favoring the groups one belongs to. Instead, group-based principles are broad ideological beliefs about how groups should interact with one another irrespective of the groups one belongs to.

This chapter focuses on one particular form of group-based principles that has been at the center of research on race-based policy attitudes: Social Dominance Orientation.

SDO researchers have made great strides in showing that broad preferences regarding how groups in society are organized significantly predict attitudes toward policies assisting minority groups (e.g. Federico and Sidanius 2002; Pratto, Sidanius, and Levin 2006; Sidanius, Pratto, and Bobo 1996). Critically, social dominance theory differs from psychology work centered on social identity because it strives to encompass the wide array of factors associated with the universal tendency for societies to organize hierarchically in groups beyond just individuals' motivations to identify with and favor ingroups. These factors include cultural ideologies, institutions, "real" conflicts between groups over finite resources, personality differences between individuals, and evolutionary mechanisms, among others (Pratto, Sidanius, and Levin 2006). Accordingly, individual differences in SDO are thought to reflect differences in the *legitimizing myths* – or "consensually held values, attitudes, beliefs, stereotypes, and cultural ideologies" (Pratto, Sidanius, and Levin 2006, p. 275) – that people rely upon to either enhance or attenuate group-based hierarchy in society. In contrast to social identity forces, which motivate individuals to favor their ingroup, both dominant and subordinate group members can latch onto the same legitimizing myths, and so it is often the case that subordinate group members subscribe to hierarchy-enhancing myths and dominant group members subscribe to hierarchy-attenuating myths (Pratto, Sidanius, and Levin 2006; see also Sidanius and Pratto 1999).

SDO has been shown to consist of two distinct dimensions: SDO-Dominance (SDO-D), which represents individuals' support for active dominance by superior

groups over inferior groups, and SDO-Equality (SDO-E), which represents individuals' rejection of equality between groups via subtle, often indirect means (Ho et al. 2012). Thus, whereas SDO-D predicts variables such as old-fashioned racism, nationalism, and support for immigrant persecution, SDO-E predicts variables such as opposition to welfare, affirmative action, and other redistributive policies (Ho et al. 2015; Jost and Thompson 2000; Kugler, Cooper, and Nosek 2010). Essentially, whereas manifestations of SDO-D tend to be overt and easily recognizable, manifestations of SDO-E tend to be covert and guised as unrelated to group preferences.

Although studies have largely parsed out the differential effects of SDO-D and SDO-E on race-targeted policies, the explanatory power of the dimensions of SDO (or any group-based principle for that matter) relative to other group-based preferences remain unknown. Specifically, as suggested throughout this dissertation, people's attitudes toward how groups in society should be arranged hierarchically are not necessarily the same as their attitudes toward the groups they belong to. Several studies find that of the two dimensions of SDO, it is SDO-D that seems to be most related to ethnocentrism and other social identity motives (Jost and Thompson 2000; Kugler, Cooper, and Nosek 2010). However, Ho et al. (2015) find that *neither* SDO-D nor SDO-E are related to ethnic identity among disadvantaged minority groups in a way that would be consistent with ingroup favoritism. Thus, despite the preponderance of research showing that SDO is a substantial predictor of race-based policy attitudes, it remains unknown how the explanatory power of SDO compares to social identity motives to favor ingroups.

Indeed, as summarized in the prior chapter, seminal work in psychology has

shown the tendency to favor ingroups to be a fairly universal and powerful force in driving human attitudes and behavior (Tajfel and Turner 1979; Turner et al. 1987), yet there is substantial variation in the degree to which people identify with ingroups and derogate outgroups in general (Brewer 1999, 2001; Sumner 1906). Kinder and Kam (2009) have shown broad identification with groups, or ethnocentrism, varies quite a bit across individuals and predicts a wide array of political issue attitudes (see also Huddy 2001). Specifically with regard to attitude toward affirmative action, Lowery et al. (2006) showed ingroup favoritism to predict opposition to affirmative action independent of concerns regarding outgroups. This work is crucial because it suggests sheer ingroup favoritism, rather than broad preferences for group-based hierarchy, may be a driving factor behind race-based policy attitudes. However, existing work is still lacking because it has yet to directly pit these distinct forces against one another, and also because even social identity motives are not unidimensional.

Roccas et al. (2008) test a multi-dimensional conceptualization of ingroup identification. According to this work, group identity can vary along four dimensions: Importance, or the degree to which individuals feel that the groups they belong to are important to who they are; Commitment, or the degree to which individuals aim to benefit and help the groups they belong to; Superiority, or the degree to which individuals view the groups they belong to as superior to other groups; and Deference, or the degree to which individuals believe it is important to honor and revere their groups' norms, symbols, and leaders. They show that SDO (modeled as a single dimension) is most related to the Superiority dimension of group identification, suggesting even if SDO-D is related to social identity motives as suggested by other studies (Jost and Thompson 2000; Kugler, Cooper, and Nosek 2010), it nonetheless

only reflects a single dimension of identity concerns. In order to fully gauge the role of identification with ingroups relative to group-based principles in predicting race-targeted policy attitudes, this study acknowledges the multidimensionality of group identity.

Is Individualism Just an Excuse for Prejudice?

A critical goal of this chapter is to shed light on a central finding of existing work on race-based policy attitudes: that “principled” individualistic arguments against race-targeted policies are often explained by group-based preferences. This suggests individualistic values are primarily used as an “excuse” for evaluating race-based policies based on what group receives assistance from the policy. In this chapter, this is referred to as the “principled mediation” hypothesis, and its proponents generally suggest contemporary individualism often serves as a socially acceptable vehicle for expressing prejudice.

Empirical work addressing this issue has generally concluded that although measures of individualism tend to largely predict race-based policy opposition, their roles are largely explained by group-based preferences. Correlational analyses have suggested SDO explains principled objections to affirmative action (Federico and Sidanius 2002; Sidanius, Pratto, and Bobo 1996), and experimental evidence suggests the link between conservatism and affirmative action attitudes is explained by stereotypes about Blacks (Reyna et al. 2005). However, a test of this proposition that acknowledges a distinction between group-based principles and social identity motives, and the multiple dimensions of each, has yet to be done. Do group-based preferences explain the role of work ethic or self-reliance? Is the relationship between

individualism and opposition to race-based policies explained by broad hierarchy-enhancing ideologies or by ingroup favoritism?

The distinction between the role of social identity motives and group-based principles is key because if the role of individualism is entirely explained by ingroup favoritism, it suggests higher-order thought has little, if anything, to do with race-based policy evaluations. However, if the role of individualism is explained by group-based principles, it suggests higher-order thought plays a role but we have simply been misunderstanding it. Then again, the principled mediation hypothesis has not been tested with the array of measures used in this study, and so it is also possible that with more direct measures of the distinct dimensions of “principles” and “prejudice,” the principled mediation finding of past research does not replicate.

Work building on evolutionary theory actually supports the idea that individualistic principles may be deeply rooted and are not necessarily just vehicles for the expression of group biases. Petersen and colleagues have argued that reliance on deservingness heuristics to evaluate whether or not individuals should receive assistance is evolutionarily rooted in within-group interactions in small-scale group living among human ancestors. Specifically, it was adaptive for humans only to share resources with “hardworking” others because according to the “logic of social exchange,” these individuals were the most likely to be able to reciprocate with favors or resources. Conversely, “lazy” individuals or “cheaters” who exploited the good will of others were less likely to reciprocate (Petersen 2012; Petersen, Sznycer, Cosmides, and Tooby 2012; see also Cosmides and Tooby 1992; 2005; Price, Cosmides, and Tooby 2002). In the context of present-day politics, it is suggested these evolved psychological tools are applied to evaluations of social welfare

programs because, through aspects of the environments such as media or socialization, individuals learn to apply this deep-seated mechanism of determining one's deservingness to mass-scale politics (Petersen 2012; Petersen et al. 2012). Accordingly, individualistic values – or at least work ethic because work ethic allows for the possibility of assisting others as long as they are deserving – may be expected to influence race-based policy attitudes net any effects from group-based attitudes.

Overview and Hypotheses

The empirical tests of this chapter were done using data from an online survey, and include correlational analyses using Structural Equation Modeling (SEM) in conjunction with an experimental manipulation of the degree to which different groups are seen as hardworking.

Self-Reliance versus Work Ethic

Two conceptualizations of individualism are measured: 1) self-reliance, i.e. a belief that individuals should strive to be self-sufficient and avoid dependence on others, and 2) work ethic, i.e. a belief that hard work is a necessary and sufficient prerequisite for success. Most existing literature interprets the role of individualism as a combination of self-reliance and work ethic, and so there is no strong empirical basis for expecting one to play a role rather than the other. However, most work, including work based in evolutionary theory by Petersen and colleagues, makes the implicit assumption that there are instances in which most people will support assistance to certain individuals even if they are opposed to government assistance in general, which would favor the role of perceptions of deservingness and work ethic rather than blanket opposition to any government assistance in line with self-reliance. If deservingness heuristics are indeed, as Petersen and colleagues claim, evolved

mechanisms for detecting cheaters, people should be willing to assist individuals seen as hardworking. As such, the experimental manipulation targets the role of work ethic by varying the degree to which Blacks versus Whites are seen as hardworking.

Nonetheless, the ideals associated with living on one's own, isolated from broader society, are prominent in the United States. Tocqueville's observations of American society are often reflected in libertarian rhetoric as well as popular media such as television shows about the charm of frontier living.

Group-Based Principles, Social Identity Motives, and the Dimensions of Each

I disentangle “prejudice” by including separate measures of group-based principles (SDO-D and SDO-E) and social identity motives (ingroup importance, ingroup commitment, ingroup superiority, and ingroup deference). Most existing work on the role of “prejudice” in influencing race-based policy attitudes has focused on racial prejudice without aiming to distinguish between group-based principles and social identity let alone the subdimensions of each, and so remains agnostic as to which is the primary factor driving the role of prejudice. However, some research points decidedly at one versus the other. On the one hand, SDO research suggests opposition to race-targeted government assistance is driven by hierarchy-enhancing ideologies that can be embraced by both dominant- and subordinate-group members, and so sides with the role of group-based principles (Federico and Sidanius 2002; Sidanius, Pratto, and Bobo 1996; Ho et al. 2015; Jost and Thompson 2000). If enough controlled processing occurs when people evaluate race-targeted policies, we should expect deep-seated ingroup bias to be translated into higher-order principles that apply to how groups should be organized broadly. On the other hand, the role of sheer ingroup favoritism has been noted by several studies as well (e.g. Dovidio, Mann, and

Gaertner 1989; Lowery et al. 2006). Ingroup favoritism is fairly deep-seated in human attitudes and behavior and occurs automatically, and scholars have increasingly argued that quick, automatic preferences dominate people's political evaluations (e.g. Lodge and Taber 2013). Given that I expect sufficient controlled processing to occur when people evaluate race-targeted policies, I nonetheless expect SDO to outperform ingroup favoritism.

Explaining “Principled Mediation”

Mediational analyses using SEM are used to disentangle the dimensions of individualism and group-based attitudes that explain the “principled mediation” finding of past research. Most prior work seems to suggest individualism masks ingroup favoritism, because people use individualistic arguments to abide by social norms that frown upon the self-interest inherent in ingroup favoritism. However, it would be consistent with the proposed model of this dissertation if individualism were mediated by group-based principles – i.e. SDO – because it is hypothesized throughout this dissertation that group-based principles will be the dominant predictor of race-based attitudes. With regard to the dimensions of SDO, some research would suggest SDO-E is the main factor underlying the role of individualism because it reflects the subtle, hierarchy-enhancing ideologies often used to justify inequality and thus should be most related to how individualism is used as a rationale for opposing redistributive policies.

There is little a priori theory for predicting which dimension of individualism group-based attitudes underlay. On the one hand, group-based attitudes may underlay self-reliance because a laissez-faire approach to government allows dominant groups to remain on top and subordinate groups to remain on the bottom. The caveat here is

that if social identity motives are the primary factor underlying self-reliance, we would expect this mediation to occur only among members of dominant groups. If group-based principles underlay self-reliance, this mediation should occur regardless of group membership. On the other hand, according to work suggesting stereotypes about Blacks being lazy are key to race-based policy opposition (Henry and Sears 2002; Reyna et al. 2005), work ethic may be the dimension of individualism most driven by group-based attitudes because such values can be relied upon selectively as a means of assigning deservingness to groups.

Data and Measures

Eight hundred and twenty-four adults were recruited through Amazon's Mechanical Turk (MTurk) between January 7th and January 16th 2015 in exchange for \$1 in compensation. Participants were required to have successfully completed at least 500 MTurk Human Intelligence Tasks (HITs), have been approved for at least 95% of the HITs they completed, and had to reside in the United States. Due to incorrect responses on attention questions implemented throughout the survey, 199 participants had to be dropped from the sample.³ Further, only White respondents were analyzed for the purposes of this study due to the large amounts of expected heterogeneity in support for race-targeted government assistance across races (see e.g. Kinder and Sears 1981; Sears and Henry 2003). The remaining sample of 513 adults was 40.9% male, and had a median age of 37 years old. Politically, the sample was 53.2% liberal, 20.1% moderate, and 26.7% conservative as measured by a five-point

³ For the attention questions, participants were asked to mark a specific response option on a 5-point scale from "Strongly Disagree" to "Strongly Agree" for each of 3 items dispersed throughout the survey. Participants who marked the wrong option for any of these 3 items were dropped from analyses. This fairly strict criterion may explain the somewhat high rate of failure on these questions.

self-report measure. MTurk has been shown to be reasonably representative of the U.S. population when compared to student samples but less representative – specifically, more educated and more liberal – than national surveys using random sampling (e.g., Berinsky, Huber, and Lenz, 2011; Buhrmester et al., 2011; Clifford, Jewell, and Waggoner, 2015). Thus this survey displayed similar biases to a typical MTurk sample (barring education, which was not measured).

Individualism. The development of questions gauging self-reliance and work ethic was influenced by widely-used measures of Protestant Work Ethic (PWE; Furnham 1990), but were specifically adopted from prior work on the multidimensionality of PWE, which divides items into one of several dimensions including self-reliance and work ethic (Miller, Woehr, and Hudspeth 2001).

Social identity motives. Broad identification with ingroups were measured via a modified version of a scale created by Roccas et al. (2008), which measures four dimensions of group identity as described earlier: Commitment, Importance, Deference, and Superiority.

Social Dominance Orientation. Social Dominance Orientation (SDO) was measured via the original 16-item scale (Pratto et al. 1994) and consisted of two dimensions: SDO-Dominance (SDO-D) and SDO-Equality (SDO-E), which gauge preferences for the domination of some groups over others and rejection of equality between groups, respectively (see Ho et al. 2012).

Race-targeted policy attitudes. Participants indicated on 7-point scales from “Strongly Oppose” to “Strongly Support” how they felt toward three redistributive race-targeted government assistance policies (affirmative action for Blacks in hiring for jobs, increased federal education spending in neighborhoods predominantly

populated by Blacks, and federally funded college scholarships for Blacks). These three items were used as indicators of a latent construct reflecting support for race-based government assistance ($\alpha = 0.92$).

Work Ethic Manipulation. Participants were split randomly into one of five conditions: a Control condition, the Blacks High Merit condition, the Blacks Low Merit Condition, the Disabled High Merit condition, or the Disabled Low Merit condition. In every condition, participants were asked to read several pages of information that were collected in March 2014 as part of a scientific report aimed at looking at how various demographic groups are similar to or different from one another. This information was not shown to participants until all questions had been answered except the manipulation check items (described below) and items aimed specifically at group-based policy attitudes, which the manipulations were intended to influence.

The information shown was all fabricated, which participants were told in a debriefing letter sent out after all data collection was complete. In all conditions, three sets of results from the fictitious study were presented to participants with short summaries of the result, why the topic was investigated, a graph displaying each result, and a few sentences explaining each graph. The content of these three results of the fictitious report was held constant across conditions and was related to the geographic distribution of Mormons in the United States, the age distribution of voters in the 2008 Presidential Election, and gender differences in perceptions of who is the “head of the household” – a male or a female.

In each of the four conditions besides the control condition, the three findings from the report described above were preceded by an additional finding related to the average amount of hours per week spent working across two demographic groups. In

the Blacks High Merit condition, participants were shown the (fictitious) finding that on average, Blacks work more hours per week (47 hours) than Whites (30). In the Blacks Low Merit condition, these numbers were simply reversed so that Whites were shown to work more hours on average per week. The Disabled High Merit and Disabled Low Merit conditions reflected these exact same conditions except that the groups being compared were disabled persons and non-disabled persons. The reason for including manipulations aimed at disabled persons, as well as the results regarding these manipulations, are explained in the next chapter. For this chapter, the manipulations of interest involve Blacks and Whites. As explained above, non-White participants were omitted from analyses so that Blacks are always the outgroup.

Following the presentation of information from the fictitious report, participants were asked questions gauging their perceptions of a variety of traits among a variety of groups including Blacks and disabled persons. Perceptions of how “hardworking” each group was (on a 7-point scale from “not hardworking” to “hardworking”) were included in each battery as a manipulation check to ensure that the information they read did, on average, influence perceptions of the work ethic of the relevant groups. The survey concluded with participants answering a series of questions about their levels of support for or opposition to group-based policies as described above.

Each model in the analyses controls for the effects of age, sex, religion (Christian or non-Christian), and political ideology (on a 7-point scale from “Strongly Liberal” to “Strongly Conservative”).⁴ The exact items used for each measure as well as the materials for the work ethic manipulation can be found in the Appendix.

⁴ The effects of control variables are not discussed in this chapter but are available upon request.

Results

Confirmatory Factor Analyses

Confirmatory factor analyses (CFAs) using robust maximum likelihood (MLR) estimation were performed in order to disentangle the potentially distinct dimensions comprising both individualism and group-based attitudes. The MLR estimation procedure adjusts standard errors to take into account the degree to which the data are not normally distributed, and is recommended for data that involve Likert-scale items such as those used in this study (e.g. Rhemtulla, Brosseau-Liard, and Savalei 2012). A scaling correction factor is estimated through the use of MLR to reflect the degree to which the data are not normally distributed, with numbers farther from 1 indicating greater skewness.

CFAs were first run separately for each individualism- and group-related construct to ensure the dimensions underlying each were measured adequately, and then jointly in a single model to assess relationships between dimensions and test whether each dimension was indeed distinct. Error covariances were added between individual factor items to account for large positive residual covariances (which indicate left over covariance between items not accounted for by the factor) and thus improve model fit. This resulted in three error covariances for the self-reliance factor of individualism and two error covariances for the work ethic factor of individualism. For group-based attitudes, one error covariance was estimated for the commitment dimension of group identity, one for SDO-D, and two for SDO-E. Error covariances between indicators were estimated only if they were theoretically appropriate (in all cases, they involved items from the same factor that had very similar wording) and significantly improved model fit. Items for which error covariances were estimated

are noted in the Appendix. The best-fitting model estimated each dimension separately, lending support to the idea that group-based principles, social identity motives, and individualism, as well as their subdimensions, are distinct. The correlations between each factor are shown in Table 2.1 and model fit for the joint CFA is shown in the caption. Full results from each CFA are available upon request.

Table 2.1 shows the individualism factors were significantly related to one another as expected. However, it should be noted that although both of the individualism factors were significantly related, they were not redundant. When we turn to group-based attitudes, it is evident that just as with individualism, a unidimensional conceptualization of group-based preferences is not appropriate. Although there was a high degree of correlation within the sub-dimensions making up group identity and group-based principles, the correlations between factors representing group identity and factors representing SDO were quite low. Although many correlations were in the expected direction, many of the factors assumed to be related in past work were only modestly correlated if at all. In particular, the correlations between each of the four dimensions of group identity were only modestly correlated with SDO-D (except for Commitment, which was uncorrelated) and the only two correlations with SDO-E were quite weak. The largest correlation between SDO and the group identification constructs existed between the deference dimension of group identity and SDO-D, despite the fact that prior work has suggested that SDO should be most related to the superiority dimension. These results thus illustrate the multi-dimensionality of group-based attitudes and a distinction between SDO (i.e. group-based principles) and ingroup-related attitudes (i.e. social identity motives).

Table 2.1. Correlation Matrix for All Latent Factors of Individualism, Social Identity Motives, and Group-Based Principles

	Self-Reliance	Work Ethic	Importance	Commitment	Superiority	Deference	SDO-D	SDO-E
Self-Reliance	1.00							
Work Ethic	0.37*	1.00						
Importance	-0.08	0.33*	1.00					
Commitment	-0.04	0.43*	0.83*	1.00				
Superiority	0.12 [‡]	0.41*	0.72*	0.75*	1.00			
Deference	0.13*	0.51*	0.60*	0.64*	0.69*	1.00		
SDO-D	0.15*	0.19*	0.14*	0.09	0.20*	0.33*	1.00	
SDO-E	0.09	0.23*	0.03	0.02	0.12*	0.12*	0.72*	1.00

$\chi^2 = 2051.267$, $df = 952$, $p < .05$; Log Likelihood = -24177.154, scaling correction factor = 1.363; CFI = 0.918; RMSEA = 0.049 (0.046-0.052); N=513; * $p < .05$, [‡] $p < .10$.

Of particular interest for this chapter are the crossover relationships between individualism, social identity motives, and SDO. Both individualism factors were positively related to superiority and deference, although the correlations for work ethic were substantially stronger than the correlations for self-reliance. Importance and commitment were also positively related to work ethic, but unlike any of the other group-based attitude factors, they were unrelated to self-reliance. Finally, turning to group-based principles, SDO-D was weakly positively correlated with both individualism factors and SDO-E was positively associated with work ethic but unrelated to self-reliance.

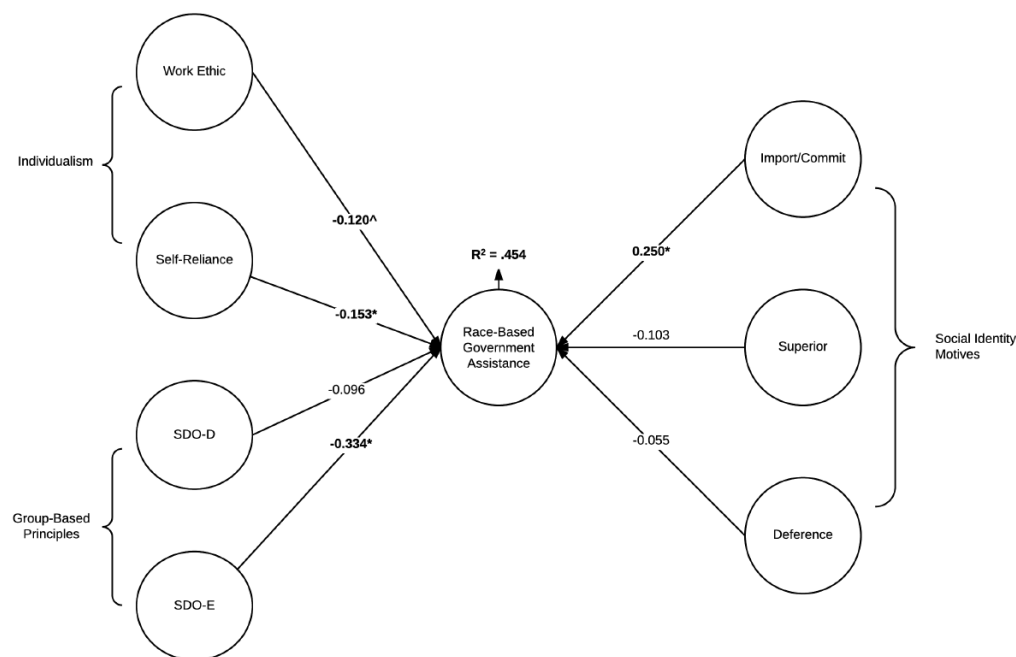
Structural Equation Models

The primary analyses of this chapter involve using the multiple dimensions of individualism, social identity motives, and group-based principles to predict race-based policy attitudes as well as to explain the relationship between individualism and race-based policy attitudes. Analyses were first conducted to estimate the simultaneous effects of all individualism and group-related constructs on race-based policy attitudes. Then, based on these full model results, subsequent models tested the degree to which the relationship between individualism and race-based policy attitudes was explained by either social identity motives or group-based principles. SEMs were run with the same estimation procedures as with the CFAs. Due to high collinearity between the commitment and importance dimensions of group identity, these factors were combined in this and all subsequent models.⁵

Figure 1 illustrates the results of an SEM predicting support for race-based government assistance with all factors comprising individualism, social identity motives, and group-based

⁵ The results did not change substantially when using these factors individually instead.

principles. It should be noted that even with just these variables in the model, nearly half of the variation in support for race-based government assistance was explained. This amount of explained variance was approximately the same when control variables were excluded from the model ($R^2 = .506$ without control variables).⁶



Model estimates are standardized MLR regression coefficients; model also controlled for all other effects of ideology, age, sex, and whether the respondent identified as Christian; * $p < .05$, ^ $p < .10$.

Figure 2.1: SEM Predicting Race-Based Policy Attitudes with Individualism, Social Identity Motives, and Group-Based Principles

SDO-E was the most predictive factor out of all categories of constructs, followed by importance of/commitment to ingroups, which is *positively* associated with support for race-based government assistance, and finally self-reliance. Work ethic was only marginally related to support for government assistance. SDO-D, superiority, and deference all showed no significant

⁶ This slight increase in the amount of variance explained when control variables are excluded is likely due to the fact that the unique variance explained by each primary independent variable increases when the collinearity between those variables and the control variables is excluded from the model.

relationship with race-based policy attitudes when the effects of the other constructs were simultaneously estimated. The results of this model are notable for several reasons. First, group-based principles – specifically, SDO-E – were dominant in predicting race-targeted policy support.⁷ However, SDO-E was not the only dimension that mattered. Social identity motives were influential, but in the *opposite* direction than expected. The importance and commitment factors were actually positively related to support for race-targeted policies assisting Blacks.⁸ It thus seems that the effects of group identity on policy attitudes are far from unidimensional. Although other social identity motives such as believing one’s ingroups are superior to other groups or believing ingroup norms and leaders should be obeyed are unrelated to race-based policy support, merely valuing membership in ingroups is associated with greater support for government assistance to Blacks. Finally, contrary to the “principled mediation” hypothesis, there was evidence of significant independent effects of individualism alongside the effects of SDO and group identity, although only self-reliance was significant at the $p < .05$ level.

The full model results therefore provide evidence suggesting the different dimensions of individualism and group-based attitudes have distinct effects on race-targeted policy attitudes but the dominant portion of explanatory power comes from SDO-E. Nonetheless, with independent effects from individualism even in the face of group-based attitudes, it seems there is still some chance the individuals are able to evaluate government assistance programs based on the perceived deservingness of the target group rather than whether the target group is an ingroup or outgroup, or where the target group stands hierarchically in society.

⁷ It should be noted that the lack of effects from SDO-D is likely due to collinearity with SDO-E, as SDO-D negatively predicts policy attitudes as expected when SDO-E is removed from the model.

⁸ When importance and commitment are included as separate constructs in the model, it is evident that importance is predominantly driving this relationship but commitment has a consistent relationship in the same direction.

Putting Merit to the Test: Effects of Work Ethic Manipulations

Table 2.2 shows the results of OLS regressions illustrating the manipulation checks testing whether the information presented in each condition influenced participants' perceptions of how "hardworking" the target groups of the manipulations were. The Black High Merit condition led to a significant increase in how hardworking participants saw Blacks but had no effect on how hardworking participants saw Whites. Conversely, the Black Low Merit condition led to a significant increase in how hardworking participants saw Whites but had no effect on their perceptions of Blacks. As such, it can be said that although the manipulations did not make the group depicted as working less hours per week seem less hardworking, they did make the group depicted as working more hours per week seem more hardworking. Thus, the manipulations did have an effect on evaluations of how hardworking Blacks and Whites are *relative to one another*.

Table 2.2: Manipulation Check – Perceptions of “Hardworking” by Condition

	Blacks	Whites
(Intercept)	4.355* (0.144)	5.094* (0.113)
Blacks High Merit Condition	0.557* (0.206)	0.131 (0.161)
Blacks Low Merit Condition	0.062 (0.206)	0.327* (0.161)
R ²	0.029	0.016
F	3.818* (df = 4; 508)	2.068 (df = 4; 505)
N	513	510

Model estimates are unstandardized OLS regression coefficients with standard errors in parentheses; The control group is the reference category; Only results for manipulations regarding Blacks and Whites shown - see Chapter 3 for results regarding disabled persons; *p < .05, †p < .10.

Yet turning to Table 2.3, it is evident that despite the effects of the conditions on perceptions of how hardworking Blacks and Whites are, there was no discernible effect of the

conditions on policy attitudes. These results call into question the independent effects of work ethic seen earlier, but corroborate the overall finding that the effects of individualism are limited nonetheless.

Table 2.3: Effects of Work Ethic Manipulations on Race-Based Policy Attitudes

	Race-Based Policy Support	
Blacks High Merit	0.066	
	(0.199)	
Blacks Low Merit	0.062	
	(0.203)	
Age	0.013	
	(0.005)	
Male	-0.064	
	(0.133)	
Christian	0.020	
	(0.141)	
Ideology	-0.615***	
	(0.049)	
R ²	0.396	
N	513	
SEM Diagnostics		
χ^2		15.794
df		16
Log Likelihood		-7042.312
Scaling Correction Factor		1.103
RMSEA	.000 (.000 - .040)	
CFI		1.000

Model estimates are standardized MLR regression coefficients with standard errors in parentheses; The control group is the reference category; Only results for manipulations regarding Blacks and Whites shown - see Chapter 3 for results regarding disabled persons; Model fit is perfect because model is fully saturated; *p < .05, †p < .10.

Testing the “Principled Mediation” Hypothesis

A critical question remains: is the relationship between individualism and support for race-based government support explained by group-based preferences? In other words, is there support for the “principled mediation” hypothesis and can we identify the dimensions of individualism and group-based attitudes responsible for it? The full model results shown in Figure 2.1 suggest individualism maintains an independent relationship with policy attitudes despite the roles of group-based principles and social identity motives, but the lack of effects of

the work ethic manipulation calls into question what the work ethic survey items really measured. Further, it is still possible that a significant portion of what would otherwise be interpreted as individualism is actually explained by group-based attitudes.

Table 2.4 compares the results of the full model with the results of a model in which SDO-E is removed from being a predictor of race-based policy attitudes. SDO-E is removed because the full model results suggest it is the only group-based attitude variable predicting race-based policy attitudes in the expected direction, and so is the only factor that can reasonably explain the role of individualism. The only individualism coefficient that changes notably when SDO-E is removed from the full model is the coefficient for work ethic. The relationship between work ethic and race-based policy support goes from being only marginally significant to being significant at the $p < .05$ level. This suggests that of the two individualism factors, work ethic shows the most evidence of potentially being mediated by group-based attitudes. Indeed, as shown in Table 2.1, SDO-E is unrelated to self-reliance but significantly positively related to work ethic.

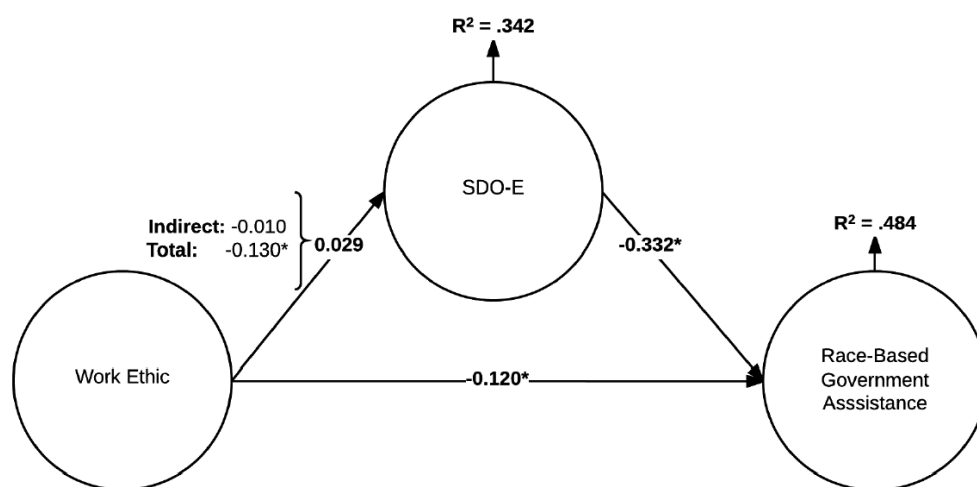
Table 2.4: Predicting Race-Based Policy Attitudes With and Without SDO-E

	With SDO-E	Without SDO-E
<i>Social Identity Motives</i>		
Importance/Commitment	0.250 (0.180)*	0.281 (0.206)*
Superiority	-0.103 (0.300)	-0.191 (0.371)
Deference	-0.055 (0.241)	0.110 (0.266)
<i>Group-Based Principles</i>		
SDO-D	-0.096 (0.140)	0.336 (0.098)*
SDO-E	-0.334 (0.132)*	—
<i>Individualism</i>		
Work Ethic	-0.120 (0.102)‡	-0.196 (0.102)*
Self-Reliance	-0.153 (0.139)*	-0.107 (0.148)*
<i>Control Variables</i>		
Age	0.012 (0.005)	-0.010 (0.005)
Male	0.014 (0.128)	0.017 (0.129)
Christian	0.013 (0.131)	0.006 (0.134)
Ideology	-0.402 (0.063)*	-0.482 (0.060)*
R ²	0.454	0.450
N	513	513
SEM Diagnostics		
χ^2	2975.589	2995.290
df	1282	1283
Log Likelihood	-29854.869	-29864.681
Scaling Correction Factor	1.342	1.348
RMSEA	.053 (.050 - .055)	.053 (.051 - .055)
CFI	.890	.889

Model estimates are standardized MLR regression coefficients with standard errors in parentheses; *p < .05, ‡p < .10.

The change in the coefficient for work ethic when SDO-E is removed from the full model suggests SDO-E may at least partially mediate the relationship between work ethic and race-based policy support (the term “partially” is used because even in the full model the coefficient for work ethic is marginally significant). However, to properly test the degree to which the effect of work ethic is explained by SDO-E, a model was tested that replicates the full model but also makes SDO-E endogenous by regressing it on work ethic, self-reliance, and the control variables that are already predicting race-based policy attitudes. Further, indirect and direct effects were calculated for work ethic through SDO-E to quantify the degree to which the relationship

between work ethic and race-based policy support is explained by SDO-E (see Sobel 1988). The direct relationship between work ethic and support for race-based government assistance as well as work ethic's indirect effect through SDO-E are illustrated in Figure 2.2.



Model estimates are standardized MLR regression coefficients; model also controlled for all other effects included in the full SEM model shown in Table 2.4, as well as the effects of self-reliance, ideology, age, sex, and whether the respondent identified as Christian on SDO-E; * $p < .05$, † $p < .10$.

Figure 2.2: Testing “Principled Mediation” of Work Ethic by SDO-E

As shown in Figure 2.2, the direct relationship between work ethic and race-based policy support is again significant, and with the control variables in the model predicting SDO-E, work ethic no longer exhibits a significant relationship with SDO-E. In conjunction with the lack of an indirect effect of work ethic through SDO-E, these results suggest the effects of work ethic on support for race-based government assistance were not mediated by SDO-E. Work ethic exhibits a fairly weak but independent relationship with race-based policy support that is not explained by group-based attitudes.

Taking Stock of What “Really” Matters

The findings presented in this chapter provide a much needed clarification to understanding what factors are “really” at play when people evaluate race-targeted policies. The existing literature is beset by conflicting empirical findings regarding the roles of “principles” versus “prejudice.” I have suggested this is largely due to conceptual ambiguity. This study not only clarifies the roles of various constructs in predicting race-based policy attitudes but also allows for a second look at previous findings on the topic. At a broad level, it is evident that both individualism and group-based attitudes are complex, multi-dimensional constructs and the dimensions of each have distinct relationships with race-based policy attitudes that are at times contrary to what would be expected given unidimensional understandings of the role of each.

Is There Room for Non-Group-Based Principles?

The values commonly thought of as underlying the role of individualism such as work ethic and self-reliance have been largely assumed in research grappling with the philosophical theories of Tocqueville, Emerson, and Weber. The findings in this chapter suggest the effects of individualistic values seem best encapsulated by self-reliance rather than work ethic, although both constructs showed evidence of consistent, independent (albeit fragile in the case of work ethic) relationships with policy attitudes in the expected direction. However, the null effects of the work ethic manipulation call into question what measures of these values – work ethic in particular – are really gauging. Do independent effects of self-reliance and work ethic suggest there is room, after all, for reliance on higher-order principles that are *not* fused with group-based attitudes? Or, perhaps when people report a strong adherence to the virtues of working

hard, they are really exposing some more deep-seated group-based preference not measured in these data.

The Dominant Role of Group-Based Principles

With regard to group-based attitudes, the distinction between group-based principles and social identity motives seems key. SDO-E is the most dominant predictor of race-based policy opinions above and beyond group identity and individualistic values. This finding should not be taken lightly, as much of the existing research on race-based policy attitudes makes the implicit assumption that ingroup favoritism is a major driver of opposition to government assistance. The results presented here suggest that on the contrary, broad ideological beliefs about how groups should be organized in society are more central to people's attitudes about government assistance to racial minority groups. Yet the results also show SDO is not the only game in town, and the relationships between people's attitudes toward groups they belong to and how they think groups should be organized broadly is more complex than suggested by prior research. Group identity variables seem more correlated with SDO-D than SDO-E, as expected (Jost and Thompson 2000; Kugler, Cooper, and Nosek 2010), but these correlations are rather weak and inconsistent. Further, although SDO-E is the most dominant predictor of race-based policy attitudes, group identity variables exhibit unique relationships with policy attitudes alongside and independently from SDO-E.

Crucially, importance of and commitment to ingroups is *positively* associated with support for race-based policies. This finding is unexpected given the decades of work in psychology and political science on the relationship between group identity and ingroup favouritism (see e.g. Hogg and Abrams 2007), as well as research specifically suggesting

ingroup favoritism leads to opposition to race-based policies (Lowery et al. 2006). It seems given the findings of this chapter that positive attitudes toward ingroups in the form of deference to ingroup norms and leaders or beliefs that one's ingroups are superior to all other groups are fairly irrelevant, but those who simply value their group memberships and seek to help members of their ingroups are often actually more likely to support race-based government assistance. How could this be? One possibility is that by controlling for the "negative" aspects of group identity such as the superiority and deference dimensions, the effects of the importance and commitment dimensions of group identity reflect broader communitarianism rather than positive feelings aimed only at ingroups. However, this possibility is not testable in the current data, nor is it a primary focus of this dissertation, and so I leave further exploration of this unexpected finding to further research.

Individualism: Not Just an Excuse for Prejudice

"Principled mediation" was not evident in the data. Although the relationship between work ethic and race-based policy support was slightly influenced by whether or not SDO-E was included as a covariate, further analysis suggested the effects of work ethic were not significantly explained by SDO-E. Work ethic maintained an independent relationship with race-based policy opinions, and so the proposition that principled objections to race-targeted policies are a facade for prejudice was not supported. This calls into question existing work suggesting SDO explains the role of individualism in predicting race-based policy attitudes (e.g. Federico and Sidanius 2002; Sidanius, Pratto, and Bobo 1996). There are multiple differences between this study and the existing work on the topic that may explain this discrepancy in results, but the primary distinction this study focused on was the operationalization of individualism. Whereas prior

research showing evidence of “principled mediation” has operationalized individualism as political conservatism or examined how SDO mediates the role of individualistic arguments against race-based policies, this study used direct measures of the individualistic values that the “principles or prejudice” literature has centered on. Thus, it is possible that some other form of individualism is at the core of “principled mediation.” Or, it is also possible that when measured directly, the doctrines associated with “pulling yourself up by your bootstraps” play a genuine role, albeit one that pales in comparison to the role of group-based principles and is not reflective of work ethic as generally conceived (according to the null effect of the manipulation).

But Have We Really Ruled Out Ingroup Favoritism?

All of the analyses presented in this chapter have predicted attitudes toward government assistance to Blacks, but given the broad nature of what I am trying to address in this dissertation – that is, the interplay between evolutionarily rooted group biases and broad higher-order principles – shouldn’t these findings hold up for a broad array of target groups besides Blacks? Most of the “principled or prejudice” literature examines attitudes toward specific racial groups (mainly Blacks) without considering how people view group-based resource distribution broadly. This is sufficient when the question is simply whether “principles” or “prejudice” underlay policy attitudes, but it does not say much about the mechanisms of these relationships. By limiting investigations to policies aimed at a particular racial group, it does not properly allow for disentangling ingroup favoritism from group-based principles because it remains possible that the role of, for example, SDO is explained by the fact that Whites are answering the SDO questions with Blacks in mind. If this is the case, it would suggest it *is* still about ingroups and outgroups. This concern is addressed in Chapter 3.

CHAPTER 3

Ingroup Bias in the Raw: The Minimal Groups Paradigm

“There is no doubt that under some conditions all men can and do display hostility towards groups other than their own, be they social, national, racial, religious or any other. There is also no doubt, however, that under other conditions this hostility either does not appear or can be modified.”

- Henri Tajfel⁹

In 1954, Muzafer Sherif gathered 22 White, middle-class, twelve-year-old boys from Protestant backgrounds on a 200-acre Boy Scouts of America camp in the Robbers Cave State Park in Oklahoma. He split the boys into two groups, which the boys named the “Eagles” and the “Rattlers.” Over the course of three weeks, he had the boys engage in group activities to reinforce their group membership and then compete in various tasks against the other group such as baseball, swimming, and tug-of-war. He found that the “Eagles” and the “Rattlers” each developed their own hierarchical structures and that competition between the groups bred ingroup pride and negative attitudes and hostile behaviors across group lines (Sherif 1954; 1958; Sherif et al. 1961). Sherif’s somewhat unsurprising findings and simple research design – flawed in many ways relative to contemporary research standards – demonstrated what decades of research in social psychology would identify as a universal human tendency: ingroup favoritism. It is not surprising, then, that a prominent explanation for why people disagree on policies that dictate how societal resources should be distributed across groups – policies such as affirmative action and race-based education spending – is that people simply prefer whatever policy favors the groups they belong to (e.g. Dovidio and Gaertner 1996; Dovidio, Mann, and Gaertner 1989; Lowery et al. 2006).

⁹ Tajfel, H. (1981). *Human groups and social categories*. New York, NY: Cambridge University Press.

Yet the idea that people evaluate government assistance policies solely on the basis of whether or not their group benefits from the policy runs directly counter to arguments that such evaluations are not racial or group-based, but are instead based on individualistic values. According to this viewpoint, people who oppose race-targeted government assistance do so because they believe everyone is capable of “pulling themselves up by their bootstraps” if they so desired, and so group-based assistance unfairly favors some individuals over others. Despite much evidence to the contrary (Henry and Sears, 2002; Kinder and Mendelberg 2000; Kinder and Sears, 1981), the perspective that individualism rather than prejudice drives race-based policy attitudes is to some degree backed up with empirical evidence (e.g., Sniderman and Carmines, 1997; Sniderman et al., 2000). However, so far in this dissertation, I have found evidence in favor of *neither* the role of individualism nor the role of ingroup favoritism.

In the previous chapter, I found group-based principles – specifically, SDO-E – rather than ingroup bias or individualism, to be the foremost predictor of attitudes toward race-targeted government assistance. Although this dissertation is far from the first study to show the dominant role that SDO plays in race-based policy attitudes (e.g. Federico and Sidanius 2002; Ho et al. 2015; Sidanius, Pratto, and Bobo 1996), it is the first study to pit its effects against the effects of ingroup favoritism and individualism simultaneously. The fact that SDO outperforms social identity motives and individualism suggests race-based policy attitudes are not driven simply by whether or not one’s own group is perceived as benefiting, but they are also not driven solely by lofty principles about work ethic and self-reliance. Instead, race-based policy opinions seem group-based yet principled at the same time. People favor certain groups over others, but in a

manner that reflects broader rules regarding group hierarchy that should be applied regardless of which groups one actually belongs to.

These findings may be surprising given the ubiquitous nature of ingroup favoritism, which scholars since Sherif have demonstrated in a wide variety of settings using a wide range of methods (Hogg and Abrams 2007; Tajfel and Turner 1979; Turner et al. 1981). Yet the central argument of this dissertation is that group-based principles rather than ingroup favoritism or individualistic values alone drive race-based policy opinions because evolutionarily speaking, humans have developed controlled neural processes that allow for higher-order thought to interact with innate, automatic group biases. In other words, the simple set of behavioral guidelines that automatic ingroup favoritism provides – a simple rule set that was developed primarily for small-scale group interactions – can be modified in such a way that addresses the broad specter of mass-scale society. Group-based principles allow individuals to translate innate group biases into generalized belief systems applicable to society as a whole.

I argue that although ingroup favoritism is a dominant force in many aspects of human behavior, it takes a back seat to group-based principles when people evaluate race-based policies because race-based policy opinions occur at a high enough level of consciousness (“LOC” in terms of social neuroscience research; Zelazo 2004) that there is ample time for automatic group biases to interact with controlled processes and higher-order thought. Therefore, I suggest the dominant role of group-based principles is due to controlled processes that interact with automatic group biases when people evaluate race-based policies.

Despite this theoretical framework, and the findings presented in Chapter 2, it is possible I have not adequately ruled out the roles of sheer ingroup favoritism and individualistic values.

Prior studies, including my analyses in Chapter 2, have been specific to particular races (almost entirely, Blacks – but see Akrami, Ekehammar, and Araya, 2000; Bahr and Chadwick, 1974; and Hurwitz and Peffley, 1992 for examples with other races). But this is problematic because examinations of “real-world” groups and policies come with the limitation that history and context inherently influence perceptions of such groups and policies. Thus, a stereotype such as that Blacks are lazy, for example, makes it difficult if not impossible to distinguish whether the role of group-based principles is due to where groups broadly are perceived as “belonging” in the social hierarchy or perceptions that Blacks, in particular, have violated individualistic principles and thus deserve lower standing in the social hierarchy. Both mechanisms would result in racial biases in support for government assistance and significant effects of group-based principles but they have substantially different implications. The former mechanism reflects group-based principles as conceptualized in this dissertation, but the latter reflects an application of individualistic values. Indeed, these sorts of conceptual problems have been at the core of disagreement in much of the “principles or prejudice” literature (see e.g., Sniderman, Crosby, and Howell 2000).

In addition, by only examining the attitudes of Whites toward Blacks, it is possible that group-based principles still act as a proxy for ingroup-outgroup dynamics. The items comprising the SDO survey battery refer to “group” in a general sense. However, if the participants are all White and the primary group they have in mind when answering the SDO questions is Blacks, it is possible the SDO measures are picking up on ingroup favoritism and social identity motives rather than broad principles about how groups should be organized regardless of one’s ingroup.

In this chapter, I use the Minimal Groups Paradigm (MGP) to examine the degree to which ingroup favoritism in a Minimal Groups (MG) setting, in which group distinctions are entirely arbitrary and have no history or context attached to them, is predicted by opposition to race-based government assistance. In doing so, I isolate ingroup favoritism in its purest form to determine the degree to which it is associated with race-based policy attitudes while avoiding confounds related to specific racial groups. Further, to isolate the role of individualistic values, the perceived merit of each minimal group is experimentally manipulated. As such, I artificially apply deservingness stereotypes to otherwise meaningless groups, and examine the degree to which individuals adjust their attitudes and behavior in accordance with the relative merit of the groups, as well as the degree to which sensitivity to merit manipulations is predicted by opposition to race-based policies.

I also utilize both implicit and explicit attitude measures within the context of the MGP in order to gauge automatic versus controlled processes as related to ingroup favoritism in the MGP as well as race-based policy attitudes. A major argument I make in this dissertation, which has thus far been untested, is that group-based principles are a dominant predictor of race-based policy opinions because they represent the interaction between automatic group biases and higher-order ideological values. Two components of this argument are tested in this chapter: 1) the relationship between ingroup favoritism and race-based policy attitudes is due to automatic processes, and 2) group-based principles are associated with automatic/implicit group biases as well as controlled/explicit higher-order principles (in this experiment, sensitivity to the relative merit between groups).

Ingroup Bias: A Versatile Tool

The details of how and why ingroup favoritism manifests have been augmented substantially since the Sherif experiments that showed how prejudice could evolve from competition between groups. Importantly for the purposes of this dissertation, it has been shown that “real” competition between groups over resources is not necessary for perceived group distinctions and ingroup favoritism to develop (Tajfel and Turner 1979; Turner et al. 1981). Social Identity Theory (Tajfel and Turner, 1979) and Self-Categorization Theory (Turner et al., 1987) have been the most dominant models of intergroup attitudes and behavior for decades, and suggests that in an effort to cultivate structure in the social world and build self-esteem, individuals will construct group distinctions by utilizing available cues from the environment (e.g. Hogg and Abrams 2007; Tajfel and Turner 1979; Turner et al. 1981). In other words, the human motivation to identify as members of a group is universal, but the precise group distinctions at play in any given situation depend on what traits and characteristics are salient. Therefore, group distinctions are fluid, and individuals can find traits to use to delineate between ingroups and outgroups in any social context. Indeed, using the MGP, research has shown that people will favor ingroups and express bias against outgroups even when group membership is arbitrary (e.g. Tajfel and Turner 1979).

Recent research in social neuroscience has corroborated the innate tendency to favor ingroups by identifying neural components associated with group categorization beyond mere racial categorizations (Van Bavel, Packer, and Cunningham 2008; 2011; Ratner, Kaul, and Van Bavel 2013). The central implication of this research is that people tend to categorize themselves

into ingroups and outgroups at an automatic level and so prejudice can manifest spontaneously across a variety of groups beyond just those laden with historically imbued negative stereotypes.

By constraining the role of prejudice in the “principles or prejudice” literature to being aimed at Blacks, it also makes it impossible to parse out the role of ingroup favoritism from the role of consciously held beliefs and stereotypes about groups. For instance, if we constrain the role of prejudice to attitudes toward Blacks, prior work showing that people oppose assistance to other groups besides Blacks (e.g. Sniderman and Carmines, 1997; Sniderman et al., 1991; Sniderman et al., 1996) may seem like evidence of individualism when in reality it could be evidence of broad ingroup favoritism. In order to fully determine whether learned stereotypes are a necessary condition for opposing government assistance to a group, it would be ideal to experimentally manipulate whether stereotypes are attached to the group.

Overview and Hypotheses

In this chapter, I use the Minimal Groups Paradigm (MGP) to examine the degree to which attitudes and behavior in a MG setting predict “real-world” policy attitudes – namely, attitudes toward race-based government assistance. By manipulating perceptions of merit between two minimal groups – one of which is the participant’s ingroup and the other of which is the participant’s outgroup – it will be possible to examine the degree to which the perceived merit of groups influences attitudes and behaviors toward those groups, as well as if these attitudes and behaviors in a MG setting differentiate supporters of race-based policies from those who oppose such policies. Further, I utilize both implicit and explicit measures of attitudes toward each of the minimal groups in order to account for the distinction between automatic and controlled processes. I conduct my analyses in three main stages.

First, I use the same survey measures that were used in Chapter 2 to see if my results from Chapter 2 replicate in the smaller laboratory sample. The only difference between the measures used in Chapter 2 and those used in this chapter is that the ingroup identification measures – i.e. the measures of social identity motives – are all directed specifically at the minimal ingroup in this experiment as opposed to ingroups in general. I expect that as in Chapter 2, group-based principles will predict race-based policy attitudes more strongly than individualism or social identity motives.

Second, I estimate the degree to which participants favored their ingroup and were influenced by the merit manipulation across three MG bias variables: an implicit MG bias variable, an explicit MG bias variable, and a behavioral MG bias variable determined by decisions in a point distribution task. I interpret greater ingroup favoritism when the ingroup is high in merit (relative to the outgroup) and less ingroup favoritism when the ingroup is low in merit (relative to the outgroup) as evidence of an influence of individualism. However, ingroup favoritism and effects of the merit manipulation in the context of the MGP alone are not the primary concern of this stage of analyses. Instead, I go on to determine whether ingroup favoritism and sensitivity to the merit manipulation (in the expected direction) is predicted by opposition to race-based policies. My hypotheses for this stage of analyses depend on the measure of MG bias. I expect that only implicit MG bias (rather than explicit or behavioral MG bias) will be associated with opposition to race-based policies because the role of ingroup favoritism should, according to my theoretical model, be limited to automatic processes. On the other hand, I expect that only explicit and behavioral MG bias will be influenced by the merit manipulations, as merit-based considerations and the role of individualism more broadly should

be mainly constrained to controlled processes. Thus, sensitivity to the merit manipulations should only be predicted by race-based policy opposition with regard to the explicit and behavioral MG bias variables.

In the third and final stage of analyses, I examine the relationships that SDO – my measure of group-based principles – has with ingroup favoritism and sensitivity to the merit manipulations. I have argued that the reason group-based principles will be the dominant predictor of race-based policy attitudes is because they represent the interaction between automatic group biases and higher-order thought. Accordingly, I expect implicit (but not explicit or behavioral) MG bias to be associated with greater levels of SDO, but the impact of explicit and behavioral (but not implicit) MG bias on SDO to be moderated by the merit manipulations. In other words, at the implicit level, I expect ingroup bias to have a direct relationship with group-based principles, but at the explicit level, I expect group-based principles to predict sensitivity to merit manipulations.

Data and Methods

Ninety undergraduate students from a mid-sized Midwestern university enrolled in various Political Science courses participated in the study for course credit. Data collection took place over the course of 3 semesters between April 2015 and March 2016. The sample was 53.3% male, 86.7% White, 4.4% Latino, and 66.7% Christian. Forty-four percent of the sample identified as Democrats or leaning Democrat, 40% identified as Republicans or leaning Republican, 12.2% identified as Independent, and 3.3% identified as something else. The sample was 45.6% liberal, 16.7% moderate, and 37.8% conservative. All survey questions and experimental procedures took place in a computer lab where no more than 5 participants were

run through at a time. Partitions separated each participant to ensure maximum isolation during the experiment. The experiment consisted of four components: a survey component (see Appendix for exact items), minimal groups assignment via the Numerical Estimation Style Test (NEST), a resource distribution task (the Tajfel matrices), and measurement of implicit attitudes via the Affect Misattribution Procedure (AMP; Payne et al. 2005).

Race-based policy attitudes. Students reported attitudes toward the same three race-targeted government assistance policies as in Chapter 2, which were averaged together to form a composite variable ($\alpha = 0.82$).

Social Dominance Orientation. The same 16 SDO items as in Chapter 2 were used to construct variables indicating SDO-E ($\alpha = 0.85$) and SDO-D ($\alpha = 0.88$).

Social identity motives. Variables for the same four dimensions of ingroup identification as in Chapter 2 were used to construct measures of ingroup importance ($\alpha = 0.83$), commitment ($\alpha = 0.85$), deference ($\alpha = 0.79$), and superiority ($\alpha = 0.87$). Importantly, instead of referring to groups broadly, these items referred specifically to the minimal ingroup in the experiment.

Individualism. Measures of work ethic ($\alpha = 0.91$) and self-reliance ($\alpha = 0.89$) were constructed using the same items as in Chapter 2.

Explicit MG bias. Measures of explicit attitudes toward each of the minimal groups consisted of feeling thermometers ranging from 0-100 with labels above the 0 (“Very Cold”), 50 (“Neutral”), and 100 (“Very Warm”) marks. A value of “2” was used as a starting point for all participants to be better able to distinguish between those who intended to mark a “0” on the scale and missing data. A measure of explicit Minimal Groups (MG) bias was constructed by subtracting each participants’ feeling thermometer score for the outgroup from their feeling

thermometer score for the ingroup. Thus, higher scores on this variable indicate greater ingroup bias.

Minimal Groups Assignment and NEST

After completing the survey component of the study, participants completed the Numerical Estimation Style Test (NEST; Tajfel et al. 1971), which is a commonly used method for randomly assigning participants to groups in the MGP. Specifically, participants were shown a series of white screens with various numbers of black dots on each screen for 5 second each, and asked to estimate the number of dots on each screen. Traditionally, participants are told a bogus story about how dot-counting tendencies are significant personality traits and that people tend to either be “Overestimators,” individuals who tend to overestimate the number of dots on each screen, or “Underestimators,” individuals who tend to underestimate the number of dots on each screen. They are told that these tendencies are randomly distributed in the population. In this study, however, participants were randomly assigned to one of three conditions: the MGP condition, High Merit Ingroup condition, and Low Merit Ingroup condition.

In the MGP condition, participants were simply divided into Overestimators and Underestimators based on the same procedures as in the traditional NEST. In the High Merit Ingroup and Low Merit Ingroup conditions, the NEST was still used but instead of dividing participants into Overestimators and Underestimators, participants were divided into Accurate Estimators and Inaccurate Estimators. As feedback, participants were randomly assigned to be told that they were either relatively accurate or relatively inaccurate in their estimates compared to others who completed the same task, and thus were categorized for the duration of the experiment as Accurate Estimators or Inaccurate Estimators. Essentially, in these conditions,

rather than representing a “true” minimal groups paradigm in which group membership is meaningless, group membership indicated the degree of merit (accuracy) associated with group members (see Appendix for Instructions for NEST and group assignments). Upon being introduced to the minimal groups, participants were shown two abstract, colorful symbols and told that one is the logo for their ingroup and the other is the logo for the other group (see Appendix for logos). The logos were randomized between participants so that the ingroup logo for half of the participants was the outgroup logo for the other half.

Resource Distribution Task

After being assigned to their groups, participants completed the resource distribution task, which took the form of the “Tajfel matrices” originally developed by Henri Tajfel and colleagues to measure ingroup favoritism and outgroup discrimination as behavior resulting from minimal group assignments in the Minimal Groups Paradigm (e.g., Tajfel and Turner, 1979; Bourhis, Sachdev, and Gagnon, 1994). This task involved participants choosing among options in which points are distributed between their own group and the outgroup. Importantly, participants were explicitly told that the points were not going to them personally, but rather to the group as a whole. Participants completed 6 versions of this task in order to yield three dependent variables: one that measures *ingroup favoritism*, or the degree to which a participant distributed more points to the ingroup instead of choosing to maximize points for both groups, one that measures *maximum differentiation*, or the degree to which a participant favored the greatest difference in points between groups in favor of the ingroup rather than maximizing profit for either the ingroup or both groups, and one that measures *parity*, or the degree to which a participant preferred equal distribution of points between groups rather than favoring the

ingroup (see Appendix for exact instructions and task). During each resource distribution trial, participants were shown the minimal groups logos at the bottom of the screen with labels above each to serve as a reminder of which represented which group.

Implicit MG Bias

Measures of implicit attitudes toward each of the two minimal groups were obtained via the Affect Misattribution Procedure (AMP; Payne, Cheng, Govorun, and Stewart 2005), in which participants were shown abstract, meaningless symbols preceded by a quickly shown prime and asked to rate whether the abstract symbol was “pleasant” or “unpleasant.” Participants completed 96 trials of the AMP. For each trial, a fixation cross was shown for 306 ms, followed by the prime for 20 ms, then a black and white noise mask for 94 ms, the abstract symbol for 106 ms, and finally another black and white noise mask until the participant registered a response. In total, aside from the time taken for the participant to respond, the stimulus presentation occurred over 526 ms. The inter-stimulus interval (ISI) between the participants’ responses and onset of the next trial was 1 s. Although the original AMP employs primes that are shown for around 75 ms and without a first noise mask following the prime, more recent work has experimented with various modifications to the protocol and suggested that a shorter stimulus presentation time with a mask immediately following the prime yields effects that isolate affect misattribution processes more appropriately (Rohr, Degner, and Wentura 2015).

The procedure is a measure of implicit biases that individuals have toward the primes, as the affect attributed to the prime has been shown to pour over into evaluations of the target symbols. This procedure has been used in a variety of research to measure implicit bias and prejudice (Hofmann and Baumert 2010; Imhoff, Schmidt, Bernhardt, Dierksmeier, and Banse

2011; Payne et al. 2005; Payne et al. 2010; Rohr, Degner, and Wentura 2015). The primes were the group logos associated with the minimal ingroup and outgroup that the participants were assigned to. Prime images were set in a random order across trials with the constraint that no three primes of the same type (Ingroup or Outgroup) occurred consecutively.

Results

Replicating Chapter 2 Results

The results of an OLS regression mostly replicated the results discussed in Chapter 2. Specifically, the dominant predictor of race-based policy opposition was SDO-E (SDO-D was significant when SDO-E was removed from the model, just as in Chapter 2). Unlike the results of Chapter 2, however, neither individualism factor predicted race-based policy attitudes. Further, In Chapter 2, I unexpectedly found positive relationships between ingroup identification and support for race-based policies. However, these unexpected relationships were not evident in these data. As would normally be expected, the importance dimension of ingroup identification was related to opposition to race-based policies. This suggests the unexpected positive relationship between ingroup importance/commitment and race-based policy support found in Chapter 2 may have indeed been spurious or due to the measures being aimed at groups in general rather than particular, relevant ingroups. The results of this analysis are shown in Table 3.1.

Table 3.1: Predicting Race-Based Policy Opinions with Group-Based Principles, Social Identity Motives, and Individualism

	Race-Based Policy Support
<i>Social Identity Motives</i>	
Importance	-0.546 (0.328) [‡]
Commitment	0.119 (0.313)
Superiority	-0.106 (0.312)
Deference	0.425 (0.303)
<i>Group-Based Principles</i>	
SDO-D	-0.342 (0.209)
SDO-E	-0.573 (0.261)*
<i>Individualism</i>	
Work Ethic	-0.685 (0.421)
Self-Reliance	0.335 (0.340)
<i>Control Variables</i>	
Male	-0.132 (0.284)
Non-Christian	0.157 (0.330)
Not Born-US	-1.809 (0.918) [‡]
Non-English 1st Language	1.693 (1.172)
Nonwhite	1.364 (0.599)*
Latino	-1.268 (0.712) [‡]
R ²	0.429
N	89

Model estimates are unstandardized OLS regression coefficients with standard errors in parentheses; *** $p < .001$, ** $p < .01$, * $p < .05$, [‡] $p < .10$.

Preliminary MGP Results

Analyses were first conducted to examine baseline levels of ingroup favoritism versus merit-based evaluations across these outcome measures. The difference between feeling thermometer scores for the ingroup ($M = 68.30$, $SD = 19.16$) and feeling thermometer scores for the outgroup ($M = 55.88$, $SD = 17.12$) was positive and significant ($t = 5.32$, $df = 82$, $p < .001$), suggesting explicit ingroup bias in the aggregate. Similarly, there was evidence of significant implicit ingroup bias in the aggregate as well, as a greater proportion of symbols following ingroup primes were rated as pleasant ($M = 0.31$, $SD = 0.23$) than symbols following outgroup

primes ($M = 0.26$, $SD = 0.22$; $t = 2.74$, $df = 81$, $p < .01$). Finally, there was mixed evidence of ingroup favoritism in the aggregate with regard to resource distribution in the Tajfel matrices. One-sample t -tests showed that participants were significantly more likely than chance to express *ingroup favoritism* over maximum joint profit ($M = 2.51$, $SD = 4.91$; $t = 4.78$, $df = 86$, $p < .001$) and *maximum differentiation* over maximum ingroup/joint profit ($M = 1.82$, $SD = 4.81$; $t = 3.52$, $df = 86$, $p < .001$). However, participants were also significantly more likely than chance to express *parity* over ingroup favoritism ($M = 8.12$, $SD = 5.39$; $t = 14.05$, $df = 86$, $p < .001$). These somewhat contradictory descriptive statistics for the three resource allocation variables is not surprising given prior work on how much variation exists in the strategies people use for these types of tasks across contexts (e.g. Bornstein et al. 1983a; 1983b; Forgas and Fiedler, 1996). Given that in this chapter the interest is in variation in resource allocation decisions rather than descriptive levels of these variables in the aggregate, this concern is of little consequence for tests of my hypotheses.

A series of OLS regressions were run to estimate the main effects of condition on explicit MG bias and resource allocation decisions. All models controlled for the effect of ideology (5-point scale from “very liberal” to “very conservative”) and included dummy control variables for being male, nonwhite, Latino, non-Christian, being born outside the US, and not speaking English as one’s first language. There was a marginal effect of condition on explicit MG bias such that there was marginally less explicit ingroup bias in the Low Merit condition than in the High Merit condition ($b = -11.088$, $SE = 5.635$, $t = -1.970$, $df = 72$, $p = .053$), suggesting weak evidence for merit-based evaluation in the expected direction. However, neither the Low Merit condition nor the High Merit condition showed a significant difference from the MGP condition.

With regard to resource distribution via the Tajfel matrices, the merit conditions had weak effects, if any, and in ways that only slightly resemble merit-based evaluation as hypothesized. There were no effects of condition on *ingroup favoritism* or *maximum differentiation*. There was marginally more evidence of *parity* in the Low Merit condition than in the MGP condition ($b = 3.018$, $SE = 1.532$, $t = 1.970$, $df = 76$, $p = .053$), as would be expected for merit-based evaluation, but there were no other significant differences between conditions.

To analyze implicit MG bias, multilevel logit models were run with each AMP trial as a separate case and allowing intercepts to vary at the participant-level. Thus, these models estimate the probability of any given symbol being rated as “Pleasant” (rather than “Unpleasant”) as a function of which group logo (ingroup or outgroup) preceded the symbol, experimental condition, and the control variables (which were the same as in the regression models above). An empty linear multilevel model grouping by participant yielded an ICC that suggested 16% of the variation across trials can be explained by participant. In the multilevel logit model, condition had a significant main effect on the probability of rating symbols as pleasant. Participants were significantly more likely to rate symbols as pleasant in the Low Merit condition than in the MGP condition ($b = 0.823$, $SE = 0.327$, $z = 2.519$, $p < .05$) or the High Merit condition ($b = 0.662$, $SE = 0.320$, $z = 2.069$, $p < .05$). However, there was no significant interaction between condition and group, and so implicit ingroup bias did not differ as a function of merit condition.

Minimal Ingroup Bias and Race-Based Policy Attitudes

There was evidence of ingroup favoritism and a lack of merit-based evaluation in the MG setting, but the hypotheses of this study have to do with the relationship between attitudes and behavior in the MG setting and “real-world” race-based policy attitudes. In these analyses, I

model attitudes and behavior in the MG setting as my outcome variables and individual differences in support for race-targeted policies as a predictor (along with the same control variables as before). To restate my hypotheses, I expect implicit, but not explicit or behavioral MG bias to be predicted by race-based policy opposition.

Table 3.2 shows the coefficients for explicit MG bias as well as each of the three behavioral MG bias variables regressed on race-based policy support. Also shown are the coefficients for interactions between race-based policy support and experimental condition, which were calculated in separate regression models. All models controlled for the same variables as in prior models. Race-based policy support showed no significant relationship with any of the explicit or behavioral MG bias variables. Further, there were no significant interactions between policy support and condition.

Table 3.2: Race-Based Policy Attitudes, Merit, Explicit MG Bias, and Behavioral MG Bias

	Explicit MG Bias	Tajfel Resource Allocation		
		Ingroup Favoritism	Maximum Differentiation	Parity
R-B Policy Support	0.324 (1.951)	-0.019 (0.455)	-0.194 (0.441)	0.072 (0.503)
R-B Policy Support*Low Merit	5.644 (4.599)	-0.048 (1.109)	0.670 (1.082)	0.080 (1.250)
R-B Policy Support*High Merit	0.167 (4.267)	1.327 (1.007)	1.382 (0.983)	-0.202 (1.135)

Estimates are unstandardized regression coefficients with standard errors in parentheses; the MGP condition is the reference group in interactions; Main effect coefficients are from models with no interaction term; ***p < .001; **p < .01; *p < .05; †p < .10.

Table 3.3 shows the results of analyses testing the same hypothesis as above except with regard to implicit attitudes. Multilevel models allowing intercepts to vary by participant were again run with interaction terms between the dummy variable for outgroup primes and the race-based policy attitude variable. These interactions indicate the extent to which the negative relationship between outgroup prime and the probability of being marked “Pleasant” (which is significant in the aggregate) changes across levels of support for race-targeted policies. Separate

models were also run to calculate three-way interactions between the outgroup prime variable, race-based policy attitudes, and condition. However, due to model convergence failure, the three-way interaction models are simple logit models.

Table 3.3: Race-Based Policy Attitudes and Implicit MG Bias

	Pr(Rate Symbol as ‘Pleasant’)
Low Merit Condition	0.797 (0.328)*
High Merit Condition	0.141 (0.330)
Outgroup Prime	-0.555 (0.050)***
R-B Policy Support	0.002 (.115)
Outgroup Prime*R-B Policy Support	0.126 (0.037)***
Outgroup Prime*R-B Policy Support*Low Merit Cond.	-0.083 (0.098)
Outgroup Prime*R-B Policy Support*High Merit Cond.	-0.043 (0.091)
N(trials)	7776
N(participants)	81
Log Likelihood	-4692.00
AIC	9412.10

Estimates are unstandardized logit coefficients from a multilevel model predicting the probability of a symbol being marked as “Pleasant” using AMP trials as observations and allowing intercepts to vary at the participant-level; Standard Errors are in parentheses; Three-way interactions with condition are calculated from separate models, and are calculated using OLS because of model convergence failures; Effects of control variables not shown; *** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$.

The interaction between the outgroup prime variable and race-based policy support was significant and positive, indicating that at greater levels of support for race-targeted policies, the gap between ingroup primes and outgroup primes was less negative than at lower levels of support for each policy. Put simply, support for race-targeted policies was associated with lower levels of implicit MG bias. Figure 1 illustrates how the effect of the outgroup prime variable (i.e. implicit MG bias) varies across race-based policy support, from one standard deviation below the mean to one standard deviation above the mean.

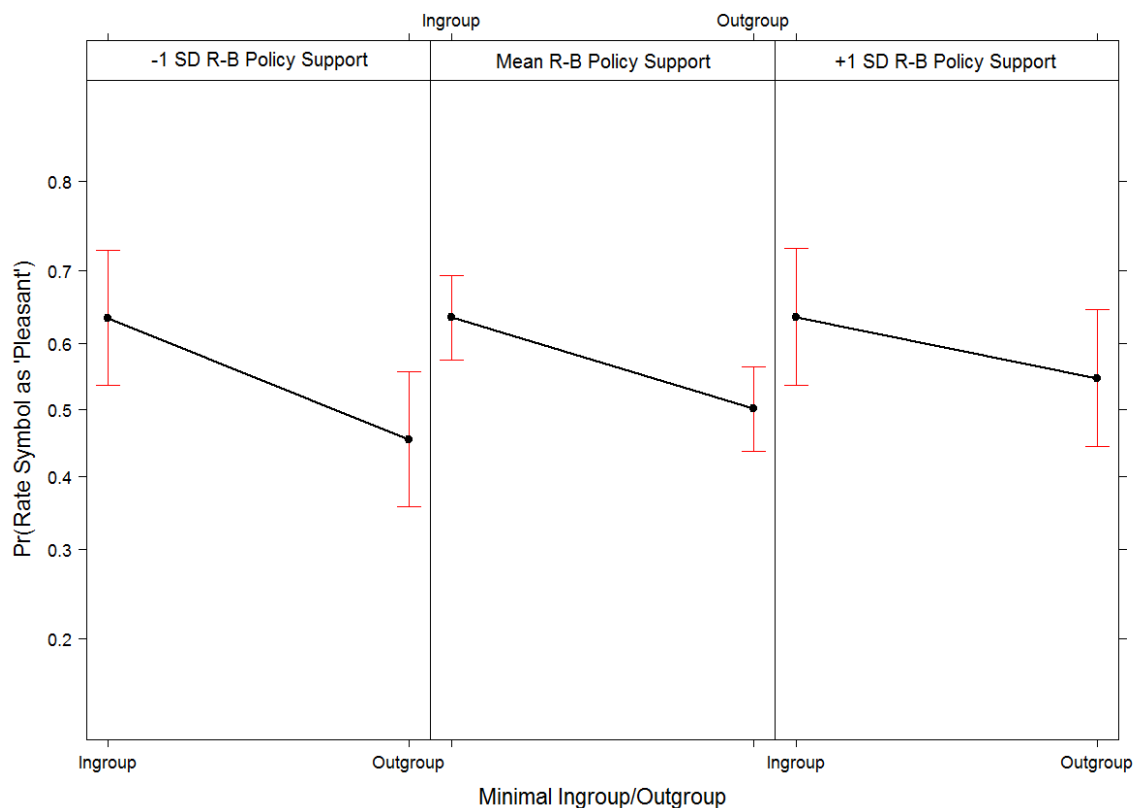


Figure 3.1: Implicit MG Bias across Levels of Support for Race-Based Policies

Table 3.3 also shows that there were no significant three-way interactions with condition, suggesting sensitivity to the merit manipulations in terms of predicting implicit MG bias did not vary as a function of race-based policy support. Those who oppose race-targeted policies were no more or less likely to exhibit merit-based evaluations than those who support race-targeted policies.

Ingroup Bias, Merit, and SDO

The final set of analyses in this chapter concerned the relationships between attitudes and behavior in the MGP and group-based principles. Table 3.4 shows the results of OLS regression models regressing explicit MG bias and each of the behavioral MG bias variables on SDO-E and

SDO-D as well as their interactions with the conditions. SDO-E and SDO-D are always included in separate models to avoid collinearity problems. Further, the interaction terms shown in Table 3.4 are estimated in separate models from those used to estimate the main effects of SDO. I hypothesized that at the explicit level (i.e. for the explicit and behavioral MG bias variables), sensitivity to the merit manipulations rather than ingroup favoritism alone would be associated with SDO. The findings shown in Table 3.4 offer mixed support for this hypothesis.

Table 3.4: SDO, Merit, Explicit MG Bias, Explicit MG Bias, and Behavioral MG Bias

<i>SDO-E</i>				
		Tajfel Resource Allocation		
	Explicit MG Bias	Ingroup Favoritism	Maximum Differentiation	Parity
SDO	2.717 (4.105)	1.948* (0.945)	1.984* (0.910)	-1.781 [‡] (1.035)
SDO*Low Merit	11.656 (7.905)	-1.169 (1.938)	-1.092 (1.873)	-1.981 (2.076)
SDO*High Merit	19.583* (8.381)	-1.857 (2.053)	-0.822 (1.984)	-4.604* (2.198)
<i>SDO-D</i>				
		Tajfel Resource Allocation		
	Explicit MG Bias	Ingroup Favoritism	Maximum Differentiation	Parity
SDO	5.358 (3.083)	2.035** (0.703)	1.986** (0.679)	-0.975 (0.798)
SDO*Low Merit	-1.423 (6.856)	-2.342 (1.792)	-2.530 (1.725)	-0.656 (2.036)
SDO*High Merit	21.725** (6.711)	-0.663 (1.737)	-0.849 (1.672)	-2.390 (1.974)

Estimates are unstandardized regression coefficients with standard errors in parentheses; the MGP condition is the reference group in interactions; Main effect coefficients are from models with no interaction term; Effects of control variables not shown; ***p < .001; **p < .01; *p < .05; [‡]p < .10.

As expected, the effects of the merit manipulations depended significantly on participants' levels of both SDO-E and SDO-D. Although there were no cases in which the Low Merit condition led to decreased ingroup favoritism compared to the MGP condition, the High Merit condition led to increased ingroup favoritism, as was hypothesized, but only significantly among those high in SDO. This interaction is illustrated in Figure 3.2 with SDO-D. From observing Figure 3.2, it can be seen that as SDO-D increases, the effects of the merit

manipulations begin to look increasingly as they are expected to if individualism is at play. This same pattern is evident for the interaction with SDO-E (not shown).

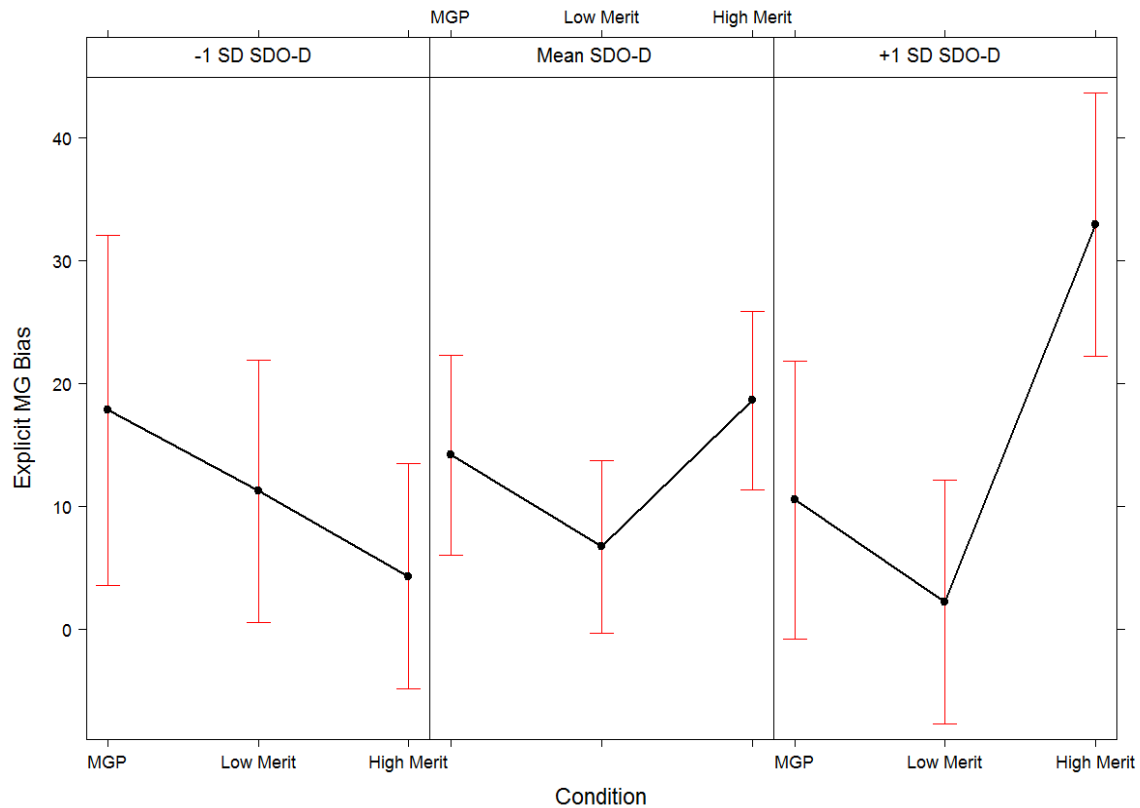


Figure 3.2: Effects of Merit Manipulations across Levels of SDO-D

Yet when it comes to behavioral MG bias, the effects of the merit manipulations do not vary across levels of SDO except in one case, where as expected individuals are less likely to exhibit parity (i.e. equal distribution between groups) in their point distribution patterns in the High Merit condition at high levels of SDO-E. But for the other behavioral MG bias measures and in all cases with regard to SDO-D, the relationship between behavioral MG bias and SDO seems direct. That is, contrary to expectations, *explicit* ingroup bias is associated with group-based principles when it comes to participants actually dividing up points between their group and the other group.

With regard to implicit attitudes in the MGP, it was expected that implicit MG bias would have a straightforward relationship with group-based principles without any dependence on the merit manipulations. Said differently, sensitivity to the merit manipulations in such a way that reflects the effects of individualism on implicit MG bias should not be dependent on SDO. As shown in Table 3.5, this hypothesis was mostly supported. The analyses used here used the same exact MLM strategy as was the case when looking at the relationship between implicit MG bias and race-based policy support. However, due to model convergence issues, all results showing the three-way interactions and all SDO-E results shown in Table 3.5 are from ordinary logit models.¹⁰ All interactions between SDO and each merit condition were significant, but not in a way that suggested greater reliance on individualism was associated with SDO. For both SDO variables, the interaction terms for both the High Merit condition and the Low Merit condition were in the same direction. Decomposing these interactions showed that implicit MG bias increased across levels of SDO primarily in the MGP condition but in both merit conditions, this relationship was dampened significantly. Put simply, the effects of the merit manipulations contained in the interactions shown in Table 3.5 do not seem to reflect individualistic behavior.

¹⁰ For the model estimating the two-way interaction between Outgroup Prime and SDO-E, the MLM allowing for random intercepts by participant converged normally with an un-centered but not a centered SDO-E variable, and so although the simple logit results are shown for this model so that the main effects are not misleading, the MLM results did not differ substantially.

Table 3.5: SDO and Implicit MG Bias

	SDO-E	SDO-D
Low Merit Condition	0.514 (0.059) ***	0.814 (0.326)*
High Merit Condition	0.039 (0.059)	0.146 (0.330)
Outgroup Prime	-0.470 (0.046) ***	-0.540 (0.051) ***
SDO	0.058 (0.058)	-0.007 (0.184)
Outgroup Prime*SDO	-0.097 (0.068)	-0.160 (0.065) ***
Outgroup Prime*SDO*Low Merit Cond.	0.623 (0.159) ***	0.348 (0.156) *
Outgroup Prime*SDO*High Merit Cond.	0.555 (0.179) **	0.367 (0.153)*
N(trials)	7776	7776
N(participants)	81	81

Main effects and the two-way interaction between SDO and Outgroup Prime for SDO-D are unstandardized logit coefficients from a multilevel model predicting the probability of a symbol being marked as “Pleasant” using AMP trials as observations and allowing intercepts to vary at the participant-level; Due to model convergence failures, all other estimates are unstandardized logit coefficients; Standard Errors are in parentheses; Three-way interactions with condition are calculated from separate models, and are calculated using OLS because of model convergence failures; Effects of control variables not shown; ***p < .001, **p < .01, *p < .05, ‡p < .10.

If we turn to the two-way interactions between Outgroup Prime and SDO, it is evident that in the case of SDO-D, the degree to which symbols following outgroup primes were rated as less pleasant than symbols following ingroup primes (i.e. implicit MG bias) was significantly moderated by SDO-D. Specifically, the negative effect of Outgroup Prime, and thus implicit MG bias, was significantly stronger at higher levels of SDO-D. This result is illustrated in Figure 3.3, and supports the hypothesis that automatic ingroup bias will be associated with group-based principles. This interaction was in the same direction but not significant with regard to SDO-E.

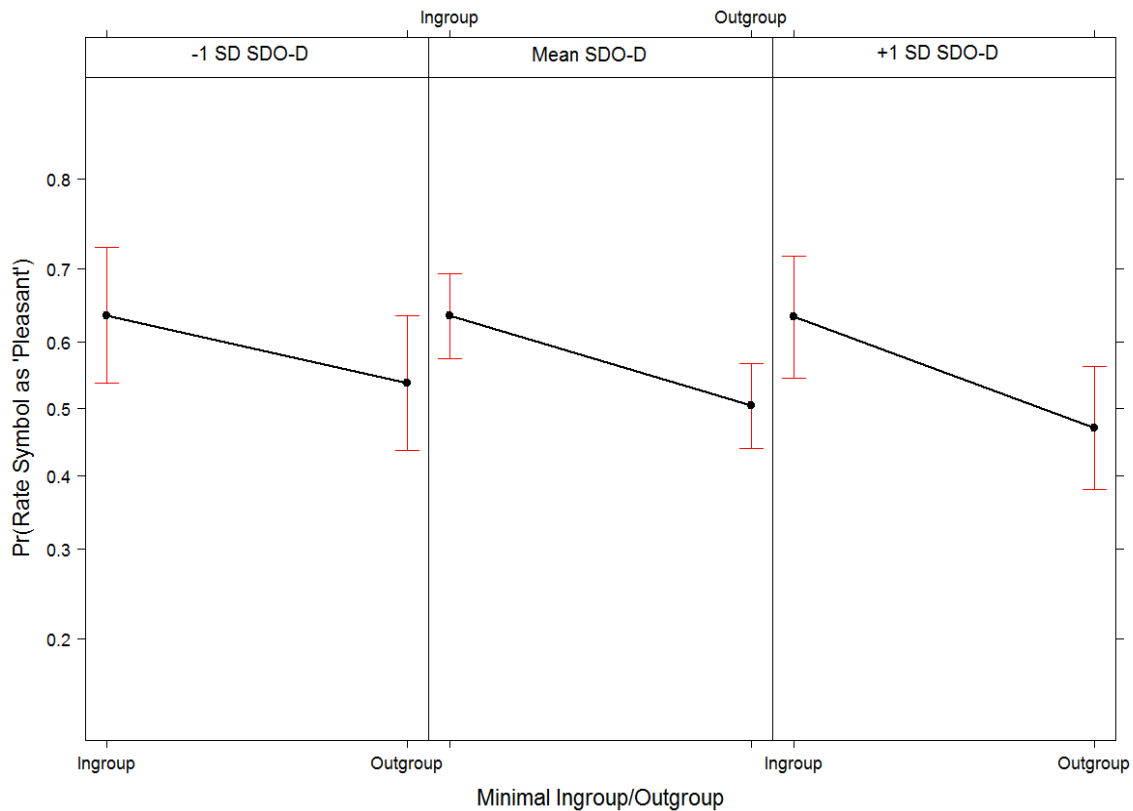


Figure 3.3: Implicit MG Bias across Levels of SDO-D

The Subtle Role of Ingroup Bias, and the Not-So-Subtle Role of Group-Based Principles

Before discussing results related to the MGP in this experiment, it is worth noting that correlational analyses basically replicated the results of Chapter 2. Group-based principles – SDO-E, in particular – predicted race-based policy attitudes better than individualistic values or social identity motives. Ingroup importance showed weak evidence of an independent *negative* effect, which suggests the unexpected positive association this construct (as well as ingroup commitment) had with race-based policy attitudes in Chapter 2 may have been spurious or due to the fact that the items did not target specific, relevant groups. Nonetheless, group-based principles were the foremost predictor of race-based policy opinions.

Turning to the results pertaining to the MGP, ingroup bias was fairly ubiquitous across explicit, implicit, and behavioral measures. Further, the ingroup bias that occurred in the MGP predicted opposition to race-targeted policies, but as hypothesized, *only* with regard to implicit attitudes. This suggests that in line with the theoretical foundations of this dissertation, there is a role for ingroup favoritism in evaluations of race-targeted policies, but it is a subtle role that results primarily from automatic rather than controlled processes. Across self-report measures, in which individuals have the time to deliberate and engage in controlled processing, the role of factors related to ingroup identification pales in comparison to factors related to group-based principles. However, ingroup bias in its purest form – that is, at the implicit level in a setting where groups have no context or prior beliefs associated with them – is significantly predicted by opposition to race-based government assistance.

Evidence of “Pure” Individualism? Not Really...

Some research using the MGP has found that manipulating the status or merit of groups has very limited effects on how the groups were treated (e.g. Bettencourt, Charlton, Dorr, and Hume, 2001; Commins and Lockwood 1979; Ellemers, Wilke, and Van Knippenberg, 1993; Tajfel and Turner, 1979). In the aggregate, the effects of the merit manipulations in this study corroborate these past findings. There were no substantial effects of the merit manipulations on implicit, explicit, or behavioral MG bias. There was also no evidence of sensitivity to the merit manipulations being predicted by race-based policy attitudes in the aggregate. Although participants in the Low Merit condition were less likely to mark symbols as pleasant in general, this effect was not moderated by whether the symbol was preceded by an ingroup or outgroup prime, and so implicit MG bias did not differ across conditions.

The Group-Based and “Principled” Foundations of Group-Based Principles

A key contribution of this chapter to the dissertation more broadly was the ability to examine whether automatic group biases and higher-order principles come together to form group-based principles as hypothesized. I have argued that evolutionarily speaking, group-based principles rather than ingroup bias or ideological values alone should predict attitudes toward race-targeted government assistance because they represent the confluence between quick, deep-seated motivations to favor ingroups and slower, controlled processes that allow individuals to adjust the “simple” guidelines provided by sheer ingroup favoritism and translate them into principles applicable to mass-scale society. The results presented in this chapter “mostly” align with this theoretical framework.

As expected, implicit ingroup bias was directly associated with SDO without being moderated by the role of merit in a way that indicated individualism, and explicit attitudes toward the minimal groups were associated with SDO in a way that indicated merit-based evaluation, or individualism. However, when it came to behavioral MG bias variables, there was weak evidence of a relationship between SDO and merit-based evaluation, as would be expected, but for the most part the relationship between behavioral ingroup bias and SDO was direct. This suggests that in certain instances – specifically, with regard to behavioral measures, which “should” be influenced heavily by controlled processes – group-based principles can represent ingroup bias that occurs at a fairly high level of consciousness. Therefore, this suggests the role of ingroup favoritism is not entirely constrained to automatic processing.

Implications for How We Think About Race and Politics

The findings presented in this chapter suggest three overarching conclusions. First, the role of ingroup favoritism in predicting race-based policy opinions seems mainly automatic. Second, reliance on individualism as evidenced by merit-based evaluations and use of deservingness cues seems extremely limited in the aggregate. Finally, group-based principles seem to comprise some combination of implicit *and* explicit group biases as well as sensitivity to deservingness cues, i.e. individualism. These findings are mostly in line with the theory I have presented in this dissertation, and suggest that people's attitudes toward hot-button race-related political issues such as affirmative action and race-targeted education spending are predominantly influenced by the interaction between people's deep-seated, primal motivations to favor the groups they belong to and higher-order principles that people consciously strive to abide by. But it is still reasonable to be somewhat puzzled by the fact that explicit and behavioral MG bias did not predict race-based policy attitudes. What is going on psychologically when people explicitly divvy up resources between groups in the MGP that is different from when they explicitly divvy up resources between groups in the "real-world"? The MGP is, after all, a few steps removed from "real-world" politics. In the next chapter, I go back to examining attitudes toward Whites and Blacks and explore the implications of gaps between people's implicit and explicit racial attitudes. I suggest it is the translation process between automatic and controlled processes that lead evaluations of "real-world" groups to differ from what goes on between meaningless groups in the laboratory.

CHAPTER 4

When Automatic and Controlled Collide

“His mind slid away into the labyrinthine world of doublethink. To know and not to know, to be conscious of complete truthfulness while telling carefully constructed lies, to hold simultaneously two opinions which cancelled out...to forget, whatever it was necessary to forget, then to draw it back into memory again at the moment when it was needed, and then promptly to forget it again...that was the ultimate subtlety; consciously to induce unconsciousness, and then, once again, to become unconscious of the act of hypnosis you had just performed.”

- George Orwell¹¹

In his seminal novel, *Nineteen Eighty-Four*, George Orwell introduced readers to a psychological phenomenon he referred to as “doublethink,” or “the power of holding two contradictory beliefs in one’s mind simultaneously, and accepting both of them” (214). Research has shown that what might be called doublethink does indeed exist across a variety of beliefs and attitudes, albeit in a less nefarious fashion than the government-induced doublethink Orwell spoke of. Individuals often harbor both positive and negative beliefs and attitudes toward an object concurrently – a phenomenon called ambivalence (Kaplan 1972), which has been studied extensively for its relevance to understanding political attitudes (Alvarez and Brehm 1997; Feldman and Zaller 1992; Lavine 2001; Rudolph and Popp 2007). Recent work in social psychology has shown that individuals need not be consciously aware of their ambivalence for it to exist and substantially influence how they think and process information (Petty, Tormala, Briñol, and Jarvis 2006). With regard to race, for example, individuals can think one way about a racial group at a conscious, or explicit, level and yet feel differently in their gut.

Work in political science and psychology has attempted to weigh the utility of implicit versus explicit measures of attitudes when predicting political outcomes, but this work remains

¹¹ Orwell, George. *Nineteen Eighty-Four*. London: Secker and Warburg, 1949.

largely focused on simply comparing the main effects of implicit and explicit attitude measures as predictors of political phenomena, thus neglecting the implications of individuals simultaneously holding contradictory implicit and explicit preferences. It is precisely this distinction between the implicit and explicit levels of evaluation that I am interested in. As discussed throughout this dissertation, automatic and controlled neural processes interact with one another over time, and it is this interaction that I suggest is key to understanding how people evaluate race-targeted government assistance policies.

In the previous chapter, I showed that whereas the role of ingroup bias in race-based policy attitudes seems primarily automatic, principles seem to enter the picture at a more controlled level. Yet in a minimal groups setting, people are perfectly willing to explicitly favor the ingroup, so what is different about how people process information in the MGP versus in the “real-world”? I suggest the difference lays in the translation process between automatic group biases and controlled thought that occurs when people evaluate “real-world” groups and policies relevant to mass-scale politics. As such, instead of evaluating the utility of implicit *versus* explicit attitude measures, I examine individual differences in the correspondence between implicit and explicit attitudes that exists *within* individuals and use these individual differences as a predictor of race-related political outcomes. Although the distinction between automatic and controlled processes is not perfectly reflected by the difference between implicit and explicit attitude measures, the difference between scores on these measures should indicate the gap between “more automatic” and “more controlled” evaluations.

I also address the potential concern that given the dominant role of group-based principles and the fact that group-based principles, along with other self-report measures, seem

to subsume the role of automatic group biases, measures of automatic processes might be unnecessary for understanding race-based policy attitudes. On the contrary, I suggest measures of automatic group biases are critical because the gap that exists between one's implicit and explicit attitudes can tell us the degree to which we should expect them to think ideologically rather than in terms of sheer ingroup favoritism in the first place. The observed relationship between SDO and race-based policy opinions in this dissertation and in other work, though substantial, has never been 1-to-1, and I have argued that group-based principles tend to predict race-based policy opinions because they reflect some combination of automatic group biases and higher-order principles. Therefore, in this chapter, I explore the idea that discordance between evaluations made implicitly and evaluations made explicitly can, in itself, provide useful information about the degree to which people are influenced by the "group-based" or "principled" components of group-based principles.

By examining conflict between implicit and explicit attitudes, I address several questions long overdue to be answered in the literature on race-related implicit cognition. What are the political implications when someone reports unbiased racial attitudes in a survey yet exhibits prejudiced implicit associations? Conversely, what about those who express prejudice explicitly but show no evidence of racial biases implicitly? The answers to these questions has substantial implications for understanding the preferences, biases, and motivations people bring to the table when evaluating political phenomena, and in particular policies that have to do with socially sensitive topics such as race-targeted government assistance.

In this chapter, I take advantage of the unique opportunity afforded by the the 2008 American National Election Study (ANES) Time Series data, which included participants'

responses to a race-centered version of the same implicit attitude measure used in the previous chapter, to examine implicit racial ambivalence, or the ambivalence created by the gap between implicit and explicit racial attitudes, as it exists in the United States. Further, I look at the effects of implicit ambivalence on several race-related political attitudes (including but not limited to support for race-based government assistance) and, crucially, its moderating effect on the role of ideology in predicting race-related political attitudes and behaviors. Therefore, in this chapter, I take a step back from looking at the role of group-based principles and instead focus on what I argue to be the mechanism underlying the dominant role of group-based principles –the interaction between automatic group biases, conscious racial attitudes, and ideological values.

Do We Really Need Implicit Attitude Measures?

Implicit and explicit racial biases can be observed in children as young as 6 years old, yet explicit bias decreases as children reach adulthood, suggesting that self-reported racial attitudes shift toward being less prejudiced as people become entrenched in the egalitarian norms of society (Baron and Banaji 2006). This finding highlights a critical aspect of attitudes that is increasingly recognized in psychology and political science and has been noted repeatedly throughout this dissertation: attitudes are comprised of a combination of both automatic and controlled as well as conscious and nonconscious processes, and as such, self-report measures are limited in their ability to encapsulate people's preferences (e.g. Fazio, Jackson, Dunton, and Williams 1995). When individuals have a motivation to obfuscate or edit their self-reported responses to survey items, or when people are simply unaware of their biases, explicit measures of racial attitudes can be especially problematic.

Contemporary work on implicit racial biases uses a wide array of “indirect” measures to gauge racial attitudes in ways that subvert more controlled processes that would otherwise lead individuals to tailor self-reports. A relatively straightforward and feasible way to tap implicit racial biases in large sample survey data is to measure people’s implicit associations. These methods, described in Chapter 1, primarily involve measuring individuals’ quick, automatic associations between visual stimuli and positive or negative words or images to assess deep-seated and reflexive biases, and have been extremely useful in assessing racial attitudes. These measures tap evaluations that are presumably less controlled and more automatic, and thus less susceptible to the influence of social desirability (Greenwald, McGhee, and Schwartz 1998; Olson and Fazio 2009).

A nontrivial amount of research has been done examining the relationship between implicit racial attitudes and political preferences. As referenced earlier in this dissertation, research has used the Implicit Association Test (IAT; Greenwald, McGhee, and Schwartz 1998) to predict attitudes toward President Barack Obama, the Affordable Care Act, and immigration policy (Greenwald, Smith, Sriram, Bar-Anan, and Nosek 2009; Knowles, Lowery, and Schaumberg 2010; Pérez 2010; Pérez 2016). The same measure of implicit biases used throughout this dissertation, the Affect Misattribution Procedure (AMP; Payne, Cheng, Govorun, and Stewart 2005), along with the same data used here (as it is rare to find nationally representative data with implicit measures), have been used in several articles examining how scores on the AMP relate to vote choice in the 2008 presidential election, support for affirmative action, attitudes toward Barack Obama, and other racial attitudes (Ditonto, Lau, and Sears 2013; Finn and Glaser 2010; Greenwald et al. 2009; Pasek et al. 2009; Payne, Krosnick, Pasek, and

Lelkes 2010; Segura and Valenzuela 2010). However, although most of this research finds implicit racial attitudes can indeed be used to predict political outcomes, all of these studies also find either that explicit attitude measures predict political outcomes just as well as implicit attitude measures, or that explicit attitudes mediate the effects of implicit attitudes. As such, the utility of measuring implicit attitudes has been far from self-evident with regard to predicting political outcomes.

Explicit measures may be more predictive of political outcomes than implicit measures because implicit measures often gauge learned associations and culturally imbued stereotypes rather than the racial preferences that ultimately drive political attitudes and behavior (Arkes and Tetlock 2004). Indeed, implicit measures have been shown to predict a range of behaviors indicative of racial discrimination no better than explicit measures (Blanton et al. 2009; Oswald, Mitchell, Blanton, Jaccard, and Tetlock 2013). This may be because political attitudes and behaviors are often influenced by the controlled processes that implicit measures are specifically designed to ignore, such as social desirability, which has been shown to influence attitudes such as those toward affirmative action (e.g. Wilson, Moore, McKay, and Avery 2008). Whereas implicit attitude measures are often best at predicting quick, automatic behaviors such as whether or not to shoot a potentially armed target (e.g. Correll, Urland, and Ito 2006; 2006; Petty, Fazio, and Briñol 2009), the political attitudes and behaviors often studied by political psychologists mostly take place over a period of time that allows for higher-level thinking, hesitation, inhibition, deliberation, and sometimes rationalization. Payne, Burkley, and Stokes (2008) showed that the correlation between implicit and explicit attitude measures depends heavily on the “structural fit” between measures, or the similarity between measures in task demands. For

example, the correlation between AMP scores and explicit evaluations of Black and White faces was significantly higher than the correlation between AMP scores and survey batteries such as the Modern Racism Scale (MRS; McConahay 1983). In other words, implicit attitudes tend to be most strongly related to outcome measures that involve similar levels of conscious processing, and political attitudes and behaviors tend to involve processing that is more deliberate than what goes on during measurement of implicit attitudes (see also Hofmann, Gawronski, Gschwendner, Le, and Schmitt 2005).

So does this mean we should abandon implicit measures as a means of understanding race and politics, or even political attitudes more broadly? I have argued that although race-related political attitudes are largely steeped in automatic processes, by the time they are reported in a survey or relied upon to cast a vote, they have interacted substantially with controlled processes and higher-order thought. So why do we need to actually measure automatic processes? I suggest that implicit attitude measures are actually central to understanding race-related political attitudes and shed much needed light on the mechanisms of political cognition broadly. I propose that focusing research around whether implicit *or* explicit attitudes better predict political outcomes is problematic because such research ignores the unique value of the *interplay* between implicit and explicit attitudes. By neglecting the spectrum upon which attitudes exist from automatic (early processes) to controlled (later processes), we may be missing a significant part of the picture and potentially building misleading models of when and how things like racial prejudice influence political attitudes.

More broadly, neglecting the interplay between implicit and explicit preferences means neglecting the full range of motivations and attitudinal forces that shape public opinion. Racial

attitudes, in particular, have been shown to be central to American public opinion, but they compete with other values and considerations in driving political attitudes and behavior. The results presented in this chapter suggest the degree to which race, rather than factors like political ideology, is a driving force in evaluations of political phenomena depends largely on how much individuals have used controlled processing to “resolve” their implicit and explicit racial preferences.

Implicit Ambivalence: The Implicit-Explicit Gap

Multiple studies have examined the correlations that exist between implicit and explicit attitude measures, and the findings of these studies indicate implicit attitude measures tend to significantly predict explicit attitudes, with average correlations ranging between about .25 and .35, suggesting they are correlated but far from redundant (Cunningham, Preacher, and Banaji 2001; Hofmann et al. 2005; Nosek 2006; 2007). Individual-level variation in the “gap” between automatic and controlled processes therefore reflects individual-level variation in the mental processes that mediate the relationship between implicit and explicit attitudes.

There is some disagreement as to what the moderate correlations between implicit and explicit measures mean – i.e. why implicit and explicit measures tend to covary yet remain distinct. Some scholars suggest a single attitude construct exists and low correspondence between implicit and explicit measures occurs only to the degree that explicit measures are altered by effortful, conscious processing (Fazio et al. 1995; Fazio and Olson 2003). Others suggest implicit and explicit measures tap structurally separable constructs represented by distinct mental processes, and that there are fundamental differences between the automatic, unintentional processes picked up by implicit measures and the more controlled, deliberate

processes picked up by explicit measures (Devine 1989; Dovidio et al. 1997; Greenwald and Banaji 1995; Hofmann et al. 2005; Nosek 2006; 2007; see also Cunningham et al. 2004 and Stanley, Phelps, and Banaji 2008 for neuroscientific evidence). However, regardless of whether implicit and explicit measures tap a single mental representation or distinct constructs, there is a near consensus regarding why scores on implicit and explicit measures often differ. People have quick, automatic reactions to attitude objects, which are mostly gauged by implicit measures, and controlled processes translate these automatic reactions into conscious evaluations reported in explicit measures. With regard to race, individuals have automatic reactions to racial outgroups and, over time, may inhibit, alter, or even rationalize their automatic racial biases to abide by learned values, norms, and principles.

The fact that effortful mental processes intervene between automatic reactions to an object and controlled evaluations of that object explains how people can feel positively toward an object explicitly yet negatively at an implicit level, or vice versa. *Implicit ambivalence* refers to ambivalence that manifests as a discrepancy between one's implicitly and explicitly held attitudes (Petty et al. 2006), and is to be expected with regard to attitudes that individuals are motivated to change, such as racial attitudes (Petty et al. 2006; Plant and Devine 1998). Some work has examined the ramifications of implicit ambivalence regarding race, specifically (Petty, Briñol, and Johnson 2012). As noted by Alvarez and Brehm, "racial policy wears all the signs of a policy debate that yields *internal* conflict, or ambivalence over policy choices" (1997, p. 345; also see Dovidio, Mann and Gaertner 1989).

Expectations for Implicit Ambivalence

Psychology work on implicit ambivalence suggests it leads to effortful thinking and deliberation related to the target of ambivalence. With regard to race, individuals whose implicit and explicit attitudes conflict – i.e. those high in implicit ambivalence – should deliberate more about race than unambivalent individuals. This is because implicit ambivalence leads individuals to dedicate cognitive resources to addressing the disparity between implicit and explicit evaluations. Indeed some work has shown that priming implicit racial ambivalence led to greater scrutiny of messages about hiring African American faculty at a university (Petty, Briñol, and Johnson 2012). Therefore, I expect race to be a more central factor in the evaluation of political targets among individuals whose racial attitudes are unresolved (i.e. high in implicit ambivalence) compared to those whose attitudes are resolved (i.e. low in implicit ambivalence). Greater controlled processing centered around race is therefore assumed to be the mechanism underlying the role of implicit ambivalence.

Need for Cognition, Education, and Resolving Conflicting Attitudes

Given that implicit ambivalence is thought to yield increased controlled processing and scrutiny over the target of ambivalence, it should be expected that individual differences in implicit ambivalence are associated with various measures of sophistication such as education and Need for Cognition (NFC; Cacioppo and Petty 1982), or the trait-like tendency to enjoy effortful thinking. However, the expected direction of these relationships is uncertain. On the one hand, it might be expected that because priming implicit ambivalence is associated with deliberation and effortful thought, measures of sophistication like education and NFC will be

positively associated with implicit racial ambivalence. Indeed, although more highly educated individuals are generally less likely to express overt racial prejudice (e.g. Bobo and Licari 1989; McClosky and Zaller 1984; Sniderman and Piazza 1993), implicit biases tend to be fairly impervious to conscious attempts at controlling prejudiced expressions (e.g. Baron and Banaji 2006; Greenwald et al. 2009; but see e.g. Fazio and Olson 2003 on the malleability of implicit attitudes). Thus, those higher in NFC and education may tailor their self-reports, causing their explicitly reported racial attitudes to diverge from implicit measures. Further, it has been shown that explicit ambivalence toward political parties and candidates is positively associated with NFC, largely because such individuals are more likely to interact with heterogeneous information networks (though this is mainly the case with weak partisans; Rudolph 2011; Rudolph and Popp 2007).

On the other hand, it may be the case that NFC and education are negatively related to implicit racial ambivalence because individuals who think more effortfully in general and are highly educated may be better equipped to resolve discrepancies between their implicit and explicit preferences. In other words, because implicit ambivalence is measured as an individual differences variable (as opposed to being primed through an experimental manipulation), it may be expected that *low* levels of implicit ambivalence indicate individuals who have, over time, dedicated greater controlled processing to resolving their ambivalence. This hypothesis is supported by an array of prior research findings. Explicit attitudinal ambivalence has predominantly been shown to be negatively related to NFC because individuals who are generally more cognitively sophisticated are better able to reconcile opposing beliefs and preferences (Jonas, Diehl, and Brömer 1997; Thompson and Zanna 1995; Thompson, Zanna, and

Griffin 1995). Work on motivated reasoning that suggests it is individuals with greater cognitive sophistication who are best equipped to manipulate information in order to justify preexisting biases (Lodge and Taber 2013; Taber and Lodge 2006). Further, the role of implicit attitudes in predicting policy preferences has been shown to be strongest among the highly educated (Pérez 2016). As such, individuals high in measures of sophistication may be the most adept at shifting explicit attitudes to line up with and reinforce implicit preferences. Importantly, unlike with explicit ambivalence, individuals are not necessarily aware of their implicit ambivalence, and so this hypothesis suggests either that individuals high in NFC and education are more likely to be aware of their implicit biases, or that NFC and education influence nonconscious processes that resolve ambivalence.

Implicit Racial Ambivalence and Race-Related Political Outcomes

The relationships between implicit ambivalence and measures of sophistication described above are critical with regard to the primary goal of this chapter – that is, understanding how the interaction between automatic and controlled racial evaluations might influence attitudes toward race-related political outcomes. Given psychology research on the mechanisms underlying implicit ambivalence, individuals high in implicit ambivalence – i.e. those whose implicit and explicit attitudes differ a great deal – should evaluate race-related political outcomes along a racial dimension to a greater degree than individuals low in implicit ambivalence. Said differently, race should play a weaker role, relative to other considerations, in the evaluation of political targets among individuals with the cognitive capacity to have resolved their ambivalence.

I do not hypothesize a main effect of implicit racial ambivalence, but I do expect it will moderate the influence of ideological values on political attitudes. As explained throughout this dissertation, racial attitudes are not the only dimension along which individuals evaluate race-related political targets, and the “principles or prejudice” literature in political science has revolved around whether they are even a key dimension in light of competing values such as individualism and preferences for limited government (e.g. Federico and Sidanius 2002; Kinder and Mendelberg 2000; Kinder and Sears 1981; Sidanius, Pratto, and Bobo 1996; Sniderman et al. 1996). I have also argued that whereas the influence of group biases tends to be automatic, the influence of higher-order principles such as individualism and egalitarianism is primarily driven by controlled processes. Therefore, I expect that individuals who have utilized controlled processing to resolve disparities between their implicit and explicit racial attitudes will rely more on ideological values and higher-order principles when evaluating race-related political objects than individuals who have not utilized controlled processing to resolve their implicit ambivalence. Conversely, as implicit ambivalence causes race to become the central dimension along which political targets are evaluated, other factors – ideological values and higher-order principles like ideology and egalitarianism – will be relied upon less for evaluating these targets. In sum, as implicit ambivalence increases, the relationship between ideological values and race-related political attitudes should be diminished.

Data and Measures

The analyses for this chapter were tested using data from the post-election interviews conducted as part of the American National Election Study (ANES) 2008 Time Series Study. These data provided a unique opportunity for testing the hypotheses of this study because at the

end of the post-election interview, respondents were given the Affect Misattribution Procedure (AMP), which has been shown to be a valid measure of implicit prejudice (Payne, Cheng, Govorun, and Stewart 2005) and was the same measure used to tap automatic group biases in Chapter 3. The ANES 2008 Time Series Study took place between November 5 and December 21, 2008, and was conducted via face-to-face interviews using computer-assisted personal interviewing (CAPI). Complex sampling procedures were used to ensure adequate generalizability to the adult U.S. population. This entailed a five-stage address-based sampling frame split by county, census tract, census block group, and then household. The data were weighted in all analyses to account for this complex sampling strategy. Black and Latino respondents were oversampled, but due to the racial nature of this study, only non-Hispanic Whites were included in the sample for analyses. The exclusion of Hispanics and non-Whites was done to control for the heterogeneity that would otherwise be expected in the effects of racial attitudes. This is common practice across studies focused on racial prejudice (see e.g. Kinder and Sears 1981; Pèrez 2010; Sears and Henry 2003). Respondents received monetary compensation for participation. The final sample used for analyses consisted of 1158 respondents (44% male; median age = 49; 41% with a college degree or higher).

Prior research examining implicit ambivalence informed its measurement in this study, and so was measured as the difference between implicit and explicit racial attitudes (Petty et al. 2006).

Explicit racial bias. Explicit racial attitudes were gauged via feeling thermometers for Blacks and Whites (coded so that higher values indicate more positive feelings; Blacks: $M = 66.28$, $SD = 19.46$, $n = 1037$; Whites: $M = 73.03$, $SD = 19.32$, $n = 1038$). An explicit racial bias

score was calculated by subtracting individuals' feeling thermometer score for Whites from their feeling thermometer score for Blacks, and dividing this score by 100 to yield a score between 0 and 1. Thus, higher values on these variables reflected greater pro-White/anti-Black explicit racial bias ($M = 0.07$, $SD = 0.19$, $n = 1033$).

Implicit racial bias. Implicit racial attitudes were gauged via the AMP, which consisted of 48 computerized trials in which a White or Black male face was shown for 75 milliseconds (ms) after brief exposure to a fixation cross, followed by a Chinese ideograph shown for 250 ms (although a technical error caused the first face image for each race not to appear for some participants).¹² Following the Chinese ideograph, a black and white noise mask (a screen with random black and white pixels) was shown until respondents indicated the ideograph to be either "pleasant" or "unpleasant." Respondents were explicitly instructed to avoid influence from the prior photos. The aggregate of binary responses to the ideographs following faces of each race has been used as a valid measure of individuals' implicit evaluations of racial groups, with lower scores (i.e. lower proportions of "pleasant" to "unpleasant" responses) for ideographs preceded by a particular racial group indicating greater prejudice toward that racial group (Payne et al. 2008). Due to the technical error mentioned above, the items corresponding to each of the first face images for each race were omitted from all analyses, yielding 23 indicators for each AMP construct (AMP-Blacks: $M = 0.51$, $SD = 0.30$, $n = 1025$, $\alpha = 0.92$; AMP-Whites: $M = 0.66$, $SD = 0.26$, $n = 1025$, $\alpha = 0.90$). An implicit racial bias score was calculated by subtracting individuals'

¹² The differences between this AMP and the AMP used in Chapter 3 are therefore that the primes were shown for longer (and thus more consciously perceived), the primes involved Black and White faces rather than minimal group logos, and the symbols were Chinese ideographs rather than randomly generated symbols.

AMP score for Whites from their AMP score for Blacks. Thus, higher values on these variables reflected greater pro-White/anti-Black implicit racial bias ($M = 0.16$, $SD = 0.30$, $n = 933$).

Implicit racial ambivalence. Implicit racial ambivalence was operationalized as the absolute value of the difference between individuals' standardized (z-scored) score on the implicit bias variable and their standardized (z-scored) score on the explicit bias variable. Standardized bias scores are used to account for the fact that levels of bias on the feeling thermometer variables are not necessarily equivalent to levels of bias on the AMP variables, and so implicit ambivalence is measured as the absolute value of the difference between individuals' relative points on the distributions of implicit and explicit bias ($M = 0.83$, $SD = 0.78$, $n = 1000$). Due to a substantial skewness value of 1.52, this variable was subjected to a square root transformation when included as a predictor in statistical models.

The direction of ambivalence is not expected to matter, but was taken into account in analyses to test for dependence on whether ones explicit attitudes were more biased than their implicit attitudes or vice versa. Separate strategies were used for bivariate versus multivariate analyses. For bivariate analyses, the implicit ambivalence variable prior to the absolute value transformation was used ($M = 0.01$, $SD = 1.14$). For multivariate analyses, a dummy variable for direction of ambivalence were created using the implicit ambivalence variable prior to the absolute value transformation such that individuals with greater levels of explicit prejudice than implicit prejudice were coded as 0 and those with greater implicit prejudice than explicit prejudice were coded as 1. Ambivalence was fairly evenly split in terms of direction, with 53.7% of respondents having more anti-Black explicit than implicit attitudes relative to the respective means of each measure.

Need for Cognition. NFC was gauged via two items from the NFC scale (Cacioppo and Petty 1982). Respondents were asked about the degree to which they like having responsibility for handling situations that require a lot of thinking and whether they prefer simple or complex problems, which were summarized into a single variable ranging from 0 to 1, with higher values indicating greater NFC ($M = 0.56$, $SD = 0.35$, $n = 1059$).

Education. Education was gauged via the education summary variable provided in the ANES data, which ranges from 0 (no high school diploma) to 7 (advanced degree).

Political Interest. In multivariate models, political interest was included to control for variation explained by NFC and education that is due to knowledge about politics, specifically. Political interest was measured via items asking about how often and/or closely the respondent follows government and politics ($M = 2.45$, $SD = 1.02$, $n = 1059$).

Race-Related Political Attitudes. Race-related political attitudes were denoted by seven variables. Symbolic racism was gauged via a composite variable created by averaging the four items used in the standard symbolic racism scale (Kinder and Sears 1981; $M = 3.59$, $SD = 0.93$, $n = 1059$).

Attitudes toward the influence of Blacks in politics were measured via a single item asking whether respondents thought Blacks have 1 “too much influence” ($n = 112$), 3 “just about the right amount of influence” ($n = 546$), or 5 “too little influence” ($n = 300$) in U.S. politics.

Attitudes toward Black elected officials were measured via a set of items ultimately formed into a composite scale ranging from 1 “White candidates are a great deal better suited to be an elected official than Black candidates” to 7 “Black candidates are a great deal better suited to be an elected official than White candidates” ($M = 3.89$, $SD = 0.50$, $n = 1039$). A second

variable was formed from items asking the same question except “in terms of intelligence” ($M = 3.87$, $SD = 0.55$, $n = 1043$).

Support for preferential hiring and promotion of Blacks was gauged via a composite variable made up of an initial item asking if respondents were for or against preferential hiring and promotion of Blacks, and a second item measuring strength of their position. The resulting measure contained the categories: 1 “strongly against” ($n = 681$), 2 “not strongly against” ($n = 198$), 3 “not strongly for” ($n = 52$), and 4 “strongly for” ($n = 47$).

Support for the federal government’s role in ensuring the fair treatment of Black people in jobs was gauged via a similar composite variable comprised of an initial item regarding the respondent’s position and a second item measuring opinion strength. A preceding variable, however, asked respondents if they had an interest in the topic, and so only respondents who indicated an interest in the topic were asked about their position. Thus, the number of respondents with scores for this variable is substantially lower than that of the first measure and support is fairly higher. Higher scores indicate greater support for the federal government ensuring fair treatment of Black people in jobs (“strongly against”, $n = 247$; “not strongly against”, $n = 60$; “not strongly for”, $n = 49$; “strongly for”, $n = 194$).

Finally, a variable measuring Obama vote was measured with a question asking respondents if they voted for Barack Obama in the general election (361 voted for Obama, 455 did not).

Higher-Order Principles. Higher-order principles were measured in two ways. Political ideology was estimated via a traditional 7-point self identification scale ranging from “extremely liberal” to “extremely conservative,” as well as a follow-up question asking those who initially

chose “moderate; middle of the road,” “don’t know,” or “haven’t thought much about it” if they would choose “liberal,” “conservative,” or “moderate” if they “had to choose.” These respondents were then coded as “slightly liberal,” “slightly conservative,” and “moderate,” respectively ($M = 4.41$, $SD = 1.47$, $n = 1038$).

Egalitarianism was estimated via the mean of 6 items asking about equality in society, each using 5-point scales from “Disagree Strongly” to “Agree Strongly.” The resulting composite variable was coded so that higher scores reflect greater levels of egalitarianism ($M = 3.39$, $SD = 0.77$, $n = 1059$, $\alpha = 0.70$). The wording for all questions is shown in the Appendix.

Results

It is first worth taking advantage of the unique opportunity offered by a nationally representative sample to note the degree to which implicit racial ambivalence exists in the United States. Figure 4.1 shows two histograms illustrating implicit racial ambivalence as it existed in 2008 based on these data. The top panel shows implicit ambivalence scores before the absolute value transformation, and so includes direction of ambivalence, whereas the bottom panel shows the variable after the absolute value transformation but still before the square root transformation. The implicit and explicit bias variables were both standardized, and so values on the x-axis indicate the difference in z-scores between one’s implicit and explicit attitudes.

The average implicit ambivalence score is less than a standard deviation difference between one’s implicit and explicit attitudes scores, which makes sense given the two types of measures are often correlated but non-redundant. Yet there is variation, and the sample is roughly evenly split between people whose implicit bias score exhibits a relatively more anti-Black bias than their explicit bias score and people whose scores exhibit the opposite pattern.

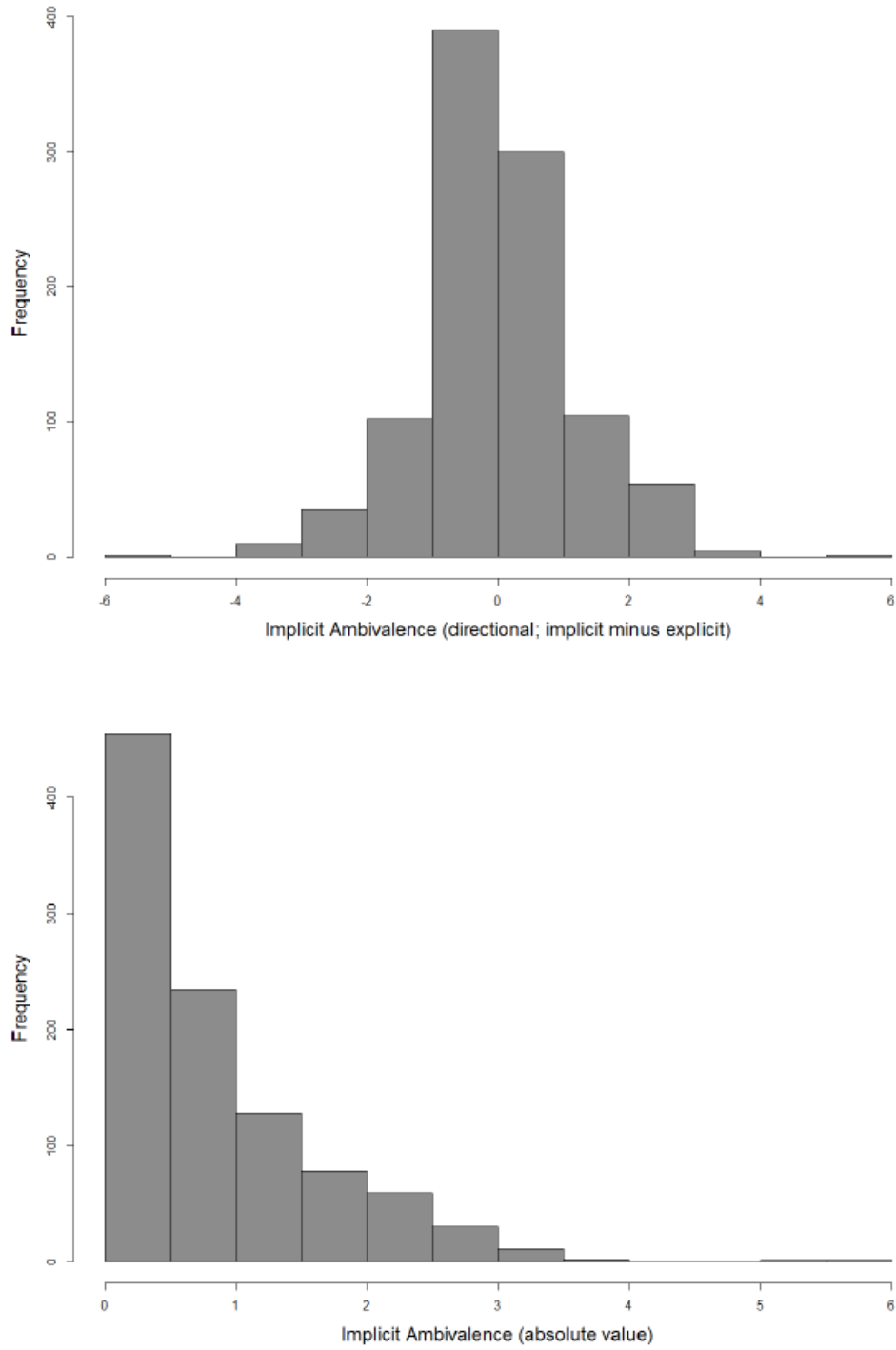


Figure 4.1: Implicit Racial Ambivalence in the United States (2008)

Moving on to the relationships implicit ambivalence has with other variables of interest, Table 4.1 shows the correlations between each of the racial bias variables (absolute value implicit ambivalence, directional implicit ambivalence, implicit bias, and explicit bias) and the constructs for which there were theoretical expectations. The implicit and explicit bias variables are included to compare the effects of implicit racial ambivalence to those of implicit and explicit bias when modeled separately.

Table 4.1: Correlations Between Racial Attitude Measures and Covariates of Interest

	NFC	Education	Symbolic Racism	Black Influence
Implicit Ambivalence (abs)	-0.059 [‡]	-0.102**	0.102**	-0.056 [‡]
Implicit Ambivalence (dir)	-0.022	-0.041	0.012	0.039
Explicit Prejudice	-0.097**	-0.150***	0.257***	-0.165***
Implicit Prejudice	-0.120***	-0.192***	0.269***	-0.117***
<i>(continued)</i>	Black Elected Officials 1	Black Elected Officials 2	Affirmative Action	Fair Job Treatment
Implicit Ambivalence (abs)	-0.218***	-0.155***	0.004	-0.101*
Implicit Ambivalence (dir)	0.189***	0.211***	-0.004	0.011
Explicit Prejudice	-0.380***	-0.426***	-0.111***	-0.236***
Implicit Prejudice	-0.192***	-0.173***	-0.123***	-0.230***
<i>(continued)</i>	Vote for Obama	Ideology	Egalitarianism	
Implicit Ambivalence (abs)	-0.083*	0.018	-0.027	
Implicit Ambivalence (dir)	-0.018	0.024	0.034	
Explicit Prejudice	-0.129***	0.077*	-0.136***	
Implicit Prejudice	-0.155***	0.100**	-0.099**	

Black Elected Officials 2 is the same as Black Elected Officials 1 except with the additional caveat of being “in terms of intelligence”. Cell entries for Vote for Obama are point-biserial correlations. *** $p < .001$; ** $p < .01$; * $p < .05$; [‡] $p < .10$.

Looking first at the non-racial variables, it is evident that implicit and explicit bias were both negatively correlated with NFC and education, as expected, and implicit racial ambivalence was also negatively associated with NFC and education (although only marginally with NFC), supporting the hypothesis that cognitive sophistication and education are associated with an increased ability of individuals to resolve discrepancies between implicit and explicit preferences. It should also be noted that the direction of ambivalence did not seem to matter for any of these variables, and so it seems to be the case that the relationship between implicit ambivalence and these variables holds whether one’s implicit attitudes are more biased relative

to the mean of implicit bias than their explicit attitudes are relative to the mean of explicit bias, or vice versa.

Turning to race-related political attitudes, it is first evident that implicit and explicit prejudice were consistently and unsurprisingly related to anti-Black positions on race-related political outcomes. More importantly for the purposes of this paper, and unexpectedly, implicit racial ambivalence also exhibited direct relationships with these political outcomes in a consistent manner. Individuals high in implicit racial ambivalence were consistently more likely to believe that Blacks have too much influence in U.S. politics (although only marginally), that Whites are better suited than Blacks to serve as elected officials, and that the federal government should not play a role in ensuring fair treatment of Blacks in jobs. Those high in implicit racial ambivalence were also less likely to report having voted for Barack Obama in the 2008 general election. The only exception was support for affirmative action, which implicit ambivalence did not predict. Direction of ambivalence only mattered for attitudes toward Black elected officials, such that individuals who exhibited more biased explicit than implicit attitudes (relative to the means of each measure) were less likely to believe Blacks are equally or better suited to serve as elected officials than individuals who expressed greater implicit than explicit bias (relative to the mean of each measure).

With regard to higher-order principles, both implicit and explicit bias were significantly related to ideological conservatism and anti-egalitarianism, but neither implicit ambivalence variable exhibited a relationship with these two variables. However, the hypothesized relationship of implicit racial ambivalence with higher-order principles is that implicit ambivalence will moderate the influence of principles on race-related political attitudes such that

the influence of principles is diminished at greater levels of implicit ambivalence. Table 4.2 summarizes the results of regression models examining this moderation hypothesis.

Table 4.2: Implicit Racial Ambivalence, Principles, and Race-Related Political Outcomes

	Symbolic Racism	Black Influence	Black Elected Officials 1	Black Elected Officials 2
Implicit Ambivalence (abs)	0.163*	0.919	-0.290***	-0.223***
Implicit Amb*Direction	-0.168	2.274**	0.491***	0.524***
Imp Amb*Dir*Ideology	0.178*	1.481*	<i>NS</i>	0.148**
Implicit Amb*Ideology	—	—	-0.046‡	—
Imp Amb*Dir*Egalitarianism	-0.304‡	<i>NS</i>	<i>NS</i>	<i>NS</i>
Implicit Amb*Egalitarianism	—	<i>NS</i>	0.102*	<i>NS</i>
<i>(continued)</i>	Affirmative Action	Fair Job Treatment	Vote for Obama	
Implicit Ambivalence (abs)	0.947	0.682*	-0.534**	
Implicit Amb*Direction	0.925	2.017‡	-0.057	
Imp Amb*Dir*Ideology	<i>NS</i>	<i>NS</i>	<i>NS</i>	
Implicit Amb*Ideology	1.380**	<i>NS</i>	0.425**	
Imp Amb*Dir*Egalitarianism	<i>NS</i>	2.484‡	<i>NS</i>	
Implicit Amb*Egalitarianism	<i>NS</i>	—	0.911**	

Cell entries represent regression coefficients in separate models controlling for NFC, education, political interest, age, and sex. Entries for Symbolic Racism, Black Elected Officials 1, and Black Elected Officials 2 are linear regression coefficients. Entries for Black Influence, Affirmative Action, and Fair Job Treatment are odds ratios from ordered logistic regressions. Entries for Vote for Obama are binomial logistic regression coefficients. Two-way interactions between implicit ambivalence and ideology/egalitarianism are only estimated if the three-ways including direction of ambivalence were not significant. *NS* = Not Significant; ****p* < .001; ***p* < .01; **p* < .05; ‡*p* < .10.

The results shown in Table 4.2 are the coefficients of interest from separate regression models. Each column represents a different dependent variable. The rows labeled “Implicit Ambivalence (abs)” indicate models examining the main effects of implicit racial ambivalence controlling for direction of ambivalence, ideology, egalitarianism, NFC, education, political interest, age, and sex. The “Implicit Amb*Direction” rows represent models interacting implicit ambivalence with the direction of ambivalence dummy variable to determine if direction of ambivalence matters. The other rows reflect the moderating effects of implicit ambivalence on ideology and egalitarianism, respectively. Three-way interactions between implicit ambivalence, ideology/egalitarianism, and direction of ambivalence were first tested to determine whether moderation of ideology or egalitarianism was dependent upon direction of ambivalence. If the

three-way interactions were insignificant (which is what I expect), two-way interactions were tested between implicit ambivalence and ideology/egalitarianism so that no higher-order interactions were included unnecessarily. Linear regressions were used to predict symbolic racism and attitudes toward Black elected officials. Ordered logistic regressions were used to predict perceptions of Black influence in U.S. politics, support for affirmative action, and support for the federal government ensuring fair job treatment for Blacks. Finally, a binomial logistic regression was used to predict the probability of reporting having voted for Obama.

Importantly, the intention of the implicit ambivalence variables is to estimate levels of discrepancy *between* implicit and explicit evaluations rather than focusing simply on one or the other. Thus, a significant influence of implicit ambivalence should indicate that even among individuals with equal levels of explicit prejudice, their political attitudes can be predicted by the distance between their implicit and explicit attitudes. However, a concern might be that the effects of implicit ambivalence are due simply to variation in one of the racial attitude variables (implicit or explicit). The considerable degree of variation on both implicit and explicit attitude measures alleviates this concern to some degree because it indicates that variation on the implicit ambivalence variable reflects variation across both implicit and explicit measures.¹³

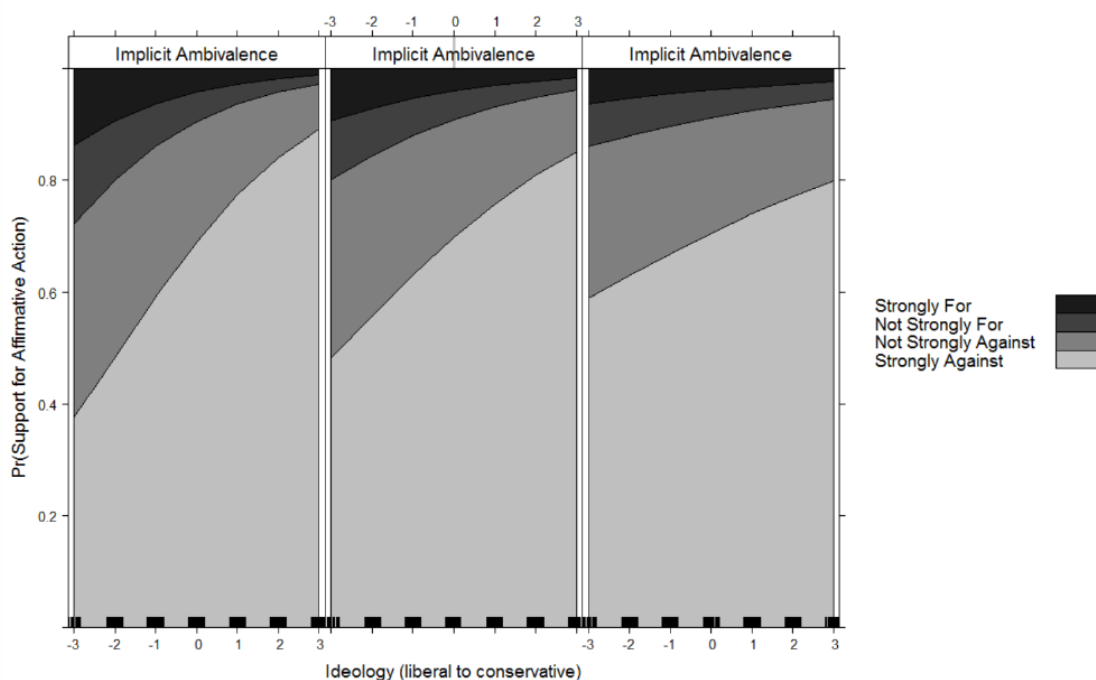
Corroborating the bivariate results described above, implicit racial ambivalence showed a direct relationship with six out of the seven dependent variables. In all cases except attitudes toward affirmative action, the discrepancy between individuals' implicit and explicit racial preferences played a statistically significant role in predicting race-related political outcomes. In

¹³ Nonetheless, models were run controlling for implicit and explicit racial attitudes (one at a time) to ensure that the effects of implicit ambivalence did not disappear when controlling for one racial attitude variable or the other (which could indicate that the effect of implicit ambivalence was driven mainly by variation in only one of the variables). The inclusion of these controls did not change results substantially and are available upon request.

two of these instances – symbolic racism and whether the respondent reported voting for Obama – the direction of ambivalence did not matter. However, in contrast to the bivariate analyses, direction mattered for four of the political outcomes examined. In all cases where direction mattered, implicit racial ambivalence was associated with anti-Black political attitudes among individuals whose explicit attitudes were more (relatively) biased than their implicit attitudes, but was unrelated to political attitudes among those whose implicit attitudes were more (relatively) biased than their explicit attitudes. For example, among those with more (relatively) biased explicit than implicit attitudes, implicit ambivalence was significantly associated with more negative attitudes toward Black relative to White elected officials ($b = -0.535, SE = 0.055, p < .001$); yet among those with more (relatively) biased implicit attitudes, there was no relationship between implicit ambivalence and attitudes toward Black elected officials ($b = -0.044, SE = 0.055, p = .421$). The same interaction pattern is evident for perceptions of the influence of Blacks in U.S. politics, attitudes toward Black elected officials “in terms of intelligence,” and marginally regarding support for the federal government ensuring fair job treatment for Blacks.

As hypothesized, implicit racial ambivalence showed evidence of moderating the roles of principles in predicting all political outcomes. However, there was substantial variation in how implicit ambivalence moderated principles across dependent variables. There were five instances of implicit ambivalence moderating the role of principles in the expected direction – that is, principles were less predictive of political outcomes at high levels of implicit racial ambivalence. For example, as shown in Figure 4.2, ideology exhibited a negative relationship with support for affirmative action at mean levels of implicit ambivalence ($OR = 0.739, CI = 0.674-0.810, p < .001$), but this relationship was significantly stronger at low levels of implicit ambivalence ($OR =$

0.648, $CI = 0.568-0.736, p < .001$) and significantly weaker at high levels of implicit ambivalence ($OR = 0.843, CI = 0.748-0.950, p < .01$).



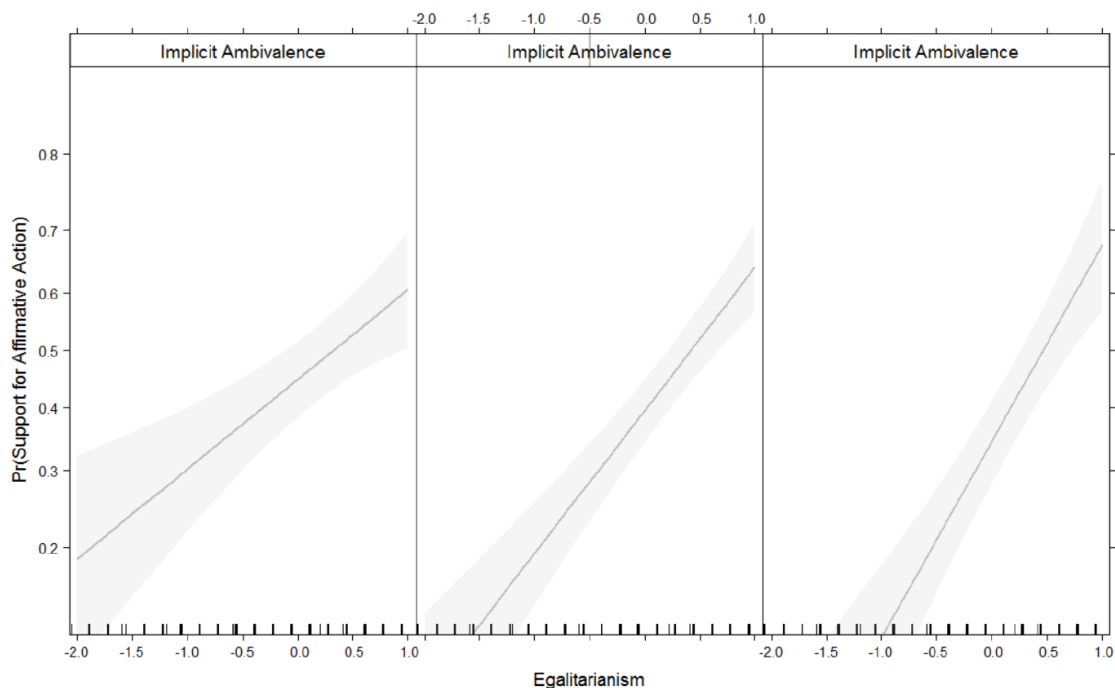
Stacked probabilities of each level of support are indicated by variations in shading. Smaller proportions of the y-axis shaded for each level of support indicate smaller probabilities of that level. Left column indicates -1 SD Implicit Ambivalence, middle column indicates mean Implicit Ambivalence, and right column indicates +1 SD Implicit Ambivalence. Rug plot shown to illustrate distribution of cases across the x-axis.

Figure 4.2: Implicit Racial Ambivalence x Ideology Predicting Support for Affirmative Action

This same pattern was evident predicting the probability of voting for Obama, symbolic racism, perceptions of the influence of Blacks in U.S. politics, and attitudes toward Black elected officials “in terms of intelligence”, but moderation of ideology was dependent on the direction of ambivalence for all of these outcomes except for the probability of voting for Obama. Regarding symbolic racism and attitudes toward Black elected officials “in terms of intelligence,” ideology was only moderated by implicit ambivalence among those whose explicit attitudes were more

(relatively) biased than their implicit attitudes. Regarding perceptions of the influence of Blacks in U.S. politics, ideology was only moderated by implicit ambivalence among those whose implicit attitudes were more (relatively) biased than their explicit attitudes. This variation in the importance of direction of ambivalence is somewhat erratic, but all of these analyses nonetheless provide evidence for the hypothesis that the influence of political ideology on political outcomes is dependent on the degree to which individuals' implicit and explicit preferences conflict with one another.

Critically, however, the effects of egalitarianism were never moderated by implicit ambivalence in the expected direction. In fact, there are four instances of the effects of egalitarianism being amplified at high levels of implicit racial ambivalence, and one instance of political ideology being amplified at high levels of implicit racial ambivalence. Figure 4.3 depicts the amplification of egalitarianism by implicit ambivalence when predicting the probability of having voted for Obama in 2008.



Y-axis indicates probability of reporting having voted for Barack Obama in the 2008 general election. Left column indicates -1 SD Implicit Ambivalence, middle column indicates mean Implicit Ambivalence, and right column indicates +1 SD Implicit Ambivalence. Rug plot shown to illustrate distribution of cases across the x-axis.

Figure 4.3: Implicit Racial Ambivalence x Egalitarianism Predicting the Probability of Voting for Obama in 2008

At mean levels of implicit ambivalence, egalitarianism was significantly associated with a higher likelihood of reporting having voted for Obama ($b = 1.005$, $SE = 0.130$, $p < .001$). Yet at low levels of implicit ambivalence, this relationship was significantly weaker ($b = 0.632$, $SE = 0.167$, $p < .001$), and at high levels of implicit ambivalence, this relationship was significantly stronger ($b = 1.379$, $SE = 0.200$, $p < .001$). This same pattern of amplification of the effects of egalitarianism was evident with regard to symbolic racism, attitudes toward Black elected officials, and support for the federal government ensuring fair job treatment for Blacks, although this amplification effect was marginal and dependent on direction of ambivalence for symbolic racism and support for fair job treatment for Blacks. In both cases where direction of

ambivalence seemed to matter, it was mainly among individuals whose implicit attitudes were more (relatively) biased than their explicit attitudes that effects of egalitarianism were amplified.

What about Ingroup Favoritism?

The results presented thus far shed light on the construct of implicit racial ambivalence, its covariates, and its ability to moderate the role of higher-order principles in predicting race-related political outcomes. However, a key component of the theoretical framework of this dissertation that I have not been able to examine with these data is how the role of ingroup favoritism might depend on levels of implicit ambivalence. I have argued that the controlled processing that comes along with resolving implicit-explicit discrepancies in racial attitudes yields deference to principles when evaluating race-targeted policies, but I have also suggested in previous chapters that controlled processing diminishes the influence of sheer ingroup favoritism. Accordingly, it should be expected that individuals who have resolved discrepancies between their implicit and explicit attitudes should be the least likely to support or oppose a race-targeted policy simply because of ingroup favoritism.

In these data, I do not have a measure of ingroup favoritism, and the racial attitude variables I have are used in my measures of implicit ambivalence. However, by briefly turning back to the data used in Chapter 3, I can construct a measure of implicit ambivalence using the implicit and explicit minimal groups attitude measures from that study and see how it relates to ingroup favoritism in terms of resource allocation – i.e. the behavioral task. Therefore, before discussing the main findings of this chapter, I do exactly that.

An implicit ambivalence variable was constructed the same exact way as the implicit racial ambivalence variable that was created using the ANES data. The only difference was that

the groups were the minimal ingroup and outgroup from the MGP rather than Blacks and Whites ($M = 1.03$, $SD = 1.02$). In these data, roughly 61% of the sample had more (relatively) biased implicit than explicit attitudes. Given high skewness, this variable was also subjected to a square root transformation before being included in statistical models. Across behavioral MG bias measures, ingroup favoritism was, as expected, significantly positively related to implicit ambivalence (ingroup favoritism: $b = 4.397$, $SE = 1.392$, $p < .01$; maximum differentiation: $b = 4.353$, $SE = 1.378$, $p < .01$; parity: $b = -1.944$, $SE = 0.741$, $p < .01$).¹⁴ In other words, individuals with more resolved attitudes – *even in the minimal groups setting* – were less likely to favor their ingroup. The relationship between implicit ambivalence and the behavioral variable representing ingroup favoritism over maximum joint profit is illustrated in Figure 4.4.

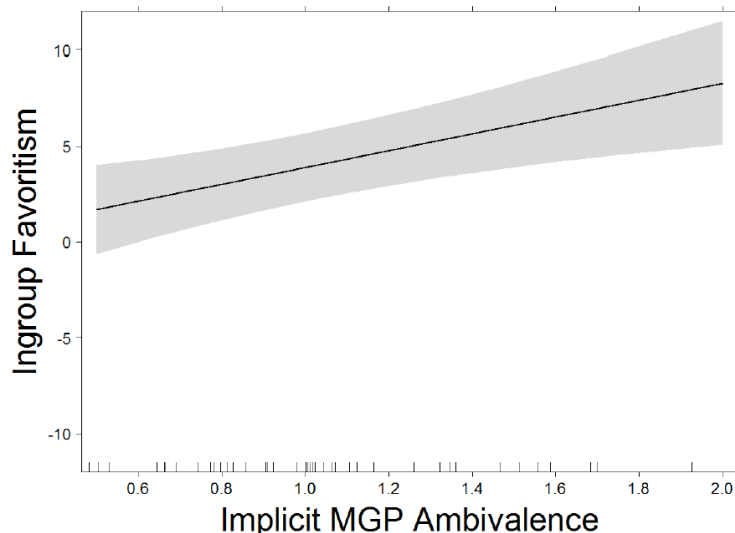


Figure 4.4: Implicit MGP Ambivalence and Ingroup Favoritism

¹⁴ Due to substantial skewness in the parity variable, a new variable dichotomizing parity into whether or not individuals chose one of the top three most equitable parity options was used as the dependent variable. All models controlled for direction of ambivalence and the control variables used in the models in Chapter 3. There were no cases in which the effect of implicit ambivalence was dependent on direction of ambivalence.

Translating from Automatic to Controlled

The theoretical foundations of this dissertation emanate largely from work in psychology showing that attitudes exist on a spectrum from automatic, gut-level and implicit to conscious, deliberate, and explicit. This work has shown that implicit attitudes are useful for understanding a variety of political attitudes and behaviors (e.g. Greenwald et al. 2009; Pasek et al. 2009; Pèrez 2010; Pèrez 2016) but are also often mediated by explicit self-reports (e.g. Ditonto, Lau, and Sears 2013; Payne et al. 2010). I have argued that what is missing from this literature is work that goes beyond simply estimating main effects of implicit and explicit attitude measures and acknowledges the interactions between the automatic and controlled processes that underlay these measures. Therefore, part of recognizing the spectrum of attitudes from implicit to explicit involves looking at within-person discrepancies between implicitly and explicitly held preferences. This is not only relevant to race-targeted policy attitudes, but also is an extension beyond how we understand the range of attitudes and considerations people bring to the table when evaluating political phenomena. In this chapter I have proposed that instead of comparing the utility of implicit and explicit measures to find the “best” measures for predicting political outcomes, it is more fruitful to embrace individual differences in the correspondence between implicit and explicit attitudes as a useful and meaningful tool for understanding race and politics as well as public opinion more broadly.

That is not to say measures of implicit ambivalence should in any way replace separate measures of implicit and explicit preferences; indeed, the effect sizes of the implicit and explicit attitude variables were often larger than that of the ambivalence variables. Instead, I suggest implicit ambivalence should be acknowledged alongside the direct roles of implicit and explicit

prejudice as a critical factor in determining how much “room” there is for other considerations when people evaluate political targets. Implicit ambivalence helps us to understand the translation process from automatic to controlled processes, which I have suggested is key to understanding how people evaluate race-targeted policies.

The results presented in this chapter largely support my proposed overarching framework, albeit in somewhat mixed fashion. First, it is worth emphasizing that according to the 2008 ANES data, Americans are roughly split in terms of where they fall on implicit and explicit racial attitude measures relative to the mean of each variable. A common argument built on research showing the tremendous breadth of implicit racial biases is that for most people, implicit attitudes are more prejudiced and people tailor their explicit self-reports to appear unprejudiced. However, the simple descriptive statistics on implicit racial ambivalence in these data suggest potentially that controlled processes can pull people’s explicit self-reports in either direction. It is worth reiterating that because the implicit and explicit attitude variables in the implicit ambivalence measure are both standardized, positive/negative implicit ambivalence scores cannot be said to directly indicate that one’s implicit/explicit racial attitudes are more “prejudiced” than their explicit/implicit racial attitudes. However, the relative differences between implicit and explicit scores are nonetheless telling, and it is perhaps possible to imagine individuals who are less biased at an automatic level but experience social pressure to exhibit prejudiced racial attitudes. This suggests controlled processes may not be universally used to inhibit prejudice. Instead, for some people, controlled processes may be more complex.

Implicit ambivalence was found to be associated with a broad swatch of politically relevant constructs, consistently results in more prejudiced race-related political attitudes and

behaviors in the aggregate, and, critically, often moderates the roles of competing considerations that political evaluations might be based on such as ideology and egalitarianism. In other words, holding explicit racial attitudes constant, the degree to which one's implicit attitudes differ from one's explicit attitudes holds substantial implications for how people evaluate race-related political outcomes, including levels of symbolic racism, attitudes to the role of Blacks in politics, race-targeted government assistance, and the probability of voting for Black candidates.

I did not expect implicit ambivalence to have a straightforward main effect on race-related political attitudes and behavior. Specifically, greater implicit racial ambivalence was associated with more prejudiced race-related political attitudes. This may be because as implicit racial ambivalence increases, the fact that the issue deals with a racial outgroup becomes more of a factor in people's evaluations of the target. This may seem somewhat counter-intuitive. If an individual is high in implicit racial ambivalence, it means that some aspect of their racial attitudes are substantially prejudiced, and so it makes intuitive sense that their political evaluations will be more prejudiced than those of someone who holds nonprejudiced explicit and implicit attitudes. However, this would also suggest that individuals high in implicit racial ambivalence exhibited more prejudiced political opinions than individuals with prejudiced implicit and explicit attitudes, which is unexpected.

Yet we may be able to explain this effect by drawing from work on cognitive dissonance. According to the action-based model of dissonance, conflicting thoughts and cognitions yield negative feelings, which individuals strive to resolve by adjusting their cognition in a way that involves the least resistance (Harmon-Jones 2008; Harmon-Jones, Amodio, and Harmon-Jones 2009). Critically, this cognitive adjustment aimed at resolving dissonance can lead individuals to

“double down” on the cognition that is less easily manipulated. Given that the suppression of prejudice has been shown to entail substantial cognitive effort and resources (e.g. Cunningham et al. 2004; Devine 1989; Devine, Plant, Amodio, Harmon-Jones, and Vance 2002; Plant and Devine 1998; Stanley, Phelps, and Banaji 2008), the “easier” cognition on which to double down might be the prejudiced one. Yet that would also suggest the automatic preference is not necessarily the “easier” one, which would not fit with any existing theory.

The moderation of the role of principles by implicit ambivalence in this chapter is crucial because it speaks to the “principles or prejudice” literature as well as whether “principles” versus “prejudice” are primarily responsible for constructs such as symbolic racism. The results presented in this chapter suggest the degree to which individuals are conflicted or ambivalent over their racial preferences consistently influences the degree to which “principles” or “prejudice” are relevant in the first place.

There are several caveats and unanswered questions pertaining to the findings of this chapter. First, the direction of ambivalence was sometimes found to matter and sometimes found not to matter, and no clear pattern emerged regarding why direction mattered in some cases and not others. It was sometimes the case that when explicit attitudes differed from implicit attitudes, explicit attitudes tended to “win out,” which fits in line with work showing explicit attitudes can mediate implicit attitudes (e.g. Dito, Lau, and Sears 2013). However, the more implicit attitudes deviated from explicit attitudes, the *more* they “won out.” Given prior research, implicit ambivalence should yield greater scrutiny and deliberation aimed at the target of ambivalence – in this case, race – regardless of direction (Petty et al. 2006), but the sporadic influence of

direction in this study suggests the ramifications of increased racial scrutiny via implicit ambivalence vary depending on the target being evaluated.

Second, implicit racial ambivalence had a mostly consistent tempering effect on the role of political ideology in predicting race-related political outcomes, as expected. However, implicit ambivalence actually seemed mostly to *amplify* the effects of egalitarianism (and in one instance the effects of political ideology) on several political outcomes, suggesting in some cases, values may be deferred to in the face of racial scrutiny. Such an effect would be consistent with some work suggesting anti-egalitarian values can be utilized as a socially acceptable tool for expressing prejudiced attitudes (e.g. Federico and Sidanius 2002; Kinder and Mendelberg 2000). This could also indicate that egalitarianism is more directly related to group-based attitudes than political ideology, which arguably encompasses a wide range of “principles” beyond how people feel about equality in society. Anti-egalitarianism may thus offer a closer structural fit for expressing prejudice than political ideology.

Ultimately, the nuances in this chapter’s findings leave several questions unanswered. However, the purpose of this chapter was to examine the implications of conflict between automatic and controlled processes. As expected, individuals who were more resolved in their racial preferences, which presumably came with effortful processing over time, were more likely to rely on ideology as a guideline for evaluating race-related political objects. More work is needed to obtain a more fine-grained understanding of the implications of when individuals’ implicit and explicit preferences diverge, but the translation process between automatic and controlled processes seems to, as expected, largely determine the relative role of higher-order principles.

CHAPTER 5

Inhibition or Ideology? The Neural Mechanisms of Principles and Prejudice

Bigotry is the disease of ignorance, of morbid minds; enthusiasm of the free and buoyant. Education and free discussion are the antidotes of both.
- Thomas Jefferson¹⁵

Despite having owned slaves, Thomas Jefferson has been quoted as condemning bigotry as the product of foolishness. But is racial prejudice due to ignorance, or does it reflect genuine moral conviction? A predominant view, with deep roots in eighteenth-century enlightenment thinking, is that bigotry and prejudice stem from deficiencies. Whether the deficiencies involve education, reason, awareness, or even cognitive ability, the common assumption is that prejudice is due to a lack of complex, enlightened thought. Accordingly, this view suggests prejudice reflects a default, simplistic way of thinking and can be overcome through effortful thought and learning about the logical fallacies of prejudiced beliefs. Although this viewpoint is primarily evident in lay rhetoric, it is corroborated by work suggesting that people are most able to inhibit racial biases when they engage in controlled processing – i.e. when they “think about” race (e.g. Cunningham et al. 2004; Mendelberg 2001; Valentino, Hutchings, and White 2002). However, although it may be the case that controlled processing is associated with lower levels of racial bias, this does not necessarily mean that inhibition of prejudice is the universal outcome of controlled processing. It could be the case that for some, controlled processing allows for prejudice inhibition, but for others, controlled processing actually exacerbates racial biases.

¹⁵ Jefferson, Thomas, 1735-1826. (1971). *The Adams-Jefferson Letters: The complete correspondence between Thomas Jefferson and Abigail and John Adams*. L. J. Cappon (Ed.). New York: Simon & Schuster.

I have argued that the function of controlled processing is not simply to diminish racial biases, nor is it to produce more reasoned thought in general for that matter. Instead, in line with work in social neuroscience (e.g. Bunge and Zelazo 2006; Cunningham and Zelazo 2007), my model suggests controlled processing is part of an iterative routine whereby simple evaluations become increasingly complex, and basic behavioral rules are transformed into broad principles that can be applied to mass-scale society. It follows that controlled processing will not always produce egalitarianism and inhibition of prejudice, and may often yield anti-egalitarian beliefs or exacerbation of prejudice as long as individuals are then able to apply these attitudes to mass-scale society broadly. Put simply, I propose that controlled processing yields deference to ideological principles rather than universal inhibition of prejudice.

The results presented throughout this dissertation thus far have largely supported my proposed model, but without direct measures of the neural mechanisms assumed to be at work, it remains difficult to understand the role controlled processing plays. In this chapter, I utilize functional Magnetic Resonance Imaging (fMRI) to directly observe the neural processes occurring when people evaluate race in the context of politics, as well as what happens when automatic evaluations are translated into higher-level principles.

The contributions of this chapter are twofold. First of all, this chapter provides a significant methodological advancement to not only this dissertation but also existing literature by allowing me to directly observe the neural mechanisms assumed to underlay race-related evaluations. Some work in social neuroscience has examined the neural correlates of automatic racial biases as well as inhibitory processes that diminish racial biases (e.g. Cunningham et al. 2004; Shkurko 2012), but for the most part (and especially in political science), these neural

processes are assumed, ignored, or measured indirectly (e.g. Huber and Lapinski 2006; Hurwitz and Peffley 2005; Mendelberg 2001; Valentino, Hutchings, and White 2002). This is in no small part due to the tremendous cost of using direct measures such as fMRI, but it is also due to the relative youth of fMRI technology and research necessary to formulate hypotheses about specific neural processes. In this chapter, I build on the surfeit of social neuroscience research on race-related evaluations that has accumulated over the past two decades to develop specific hypotheses about which neural processes should underlay controlled versus automatic processing of race-related political stimuli.

The second contribution of this chapter is to test the predictions of my model regarding the role of controlled processing. Specifically, as mentioned above, whereas existing work largely suggests controlled processing is used to inhibit racial biases, I suggest controlled processing allows people to defer to broad ideological principles, which may diminish or exacerbate racial biases depending on one's ideological preferences. Without direct measures of automatic versus controlled processes, it is impossible to know, for example, whether conservatives tend to exhibit more anti-Black political views because of a lack of controlled processing, or because the consequences of controlled processing are different for conservatives than they are for liberals. Even in Chapter 4, the finding that ideology played a stronger role in predicting race-related political attitudes at low levels of implicit ambivalence could be due to controlled processing leading to different outcomes for liberals and conservatives, or it could be due to liberals simply utilizing controlled processing to a greater degree than conservatives when resolving conflicts in their racial preferences. Put differently, without directly observing

controlled processes, we cannot determine whether encouraging people to “think about race” leads to greater racial tolerance across the board or ideological polarization.

To be clear, the hypotheses derived from my model are not necessarily incompatible with findings from existing literature (that controlled processing is associated with greater racial tolerance). It may very well be the case that individuals who express lower levels of racial bias exhibit greater levels of controlled processing than individuals who express high levels of racial bias. However, the central test of my model is not whether controlled processing is greater among racially tolerant individuals than among racially intolerant individuals. The central test is whether controlled processing results in greater racial tolerance for everyone, or in greater racial tolerance for those embracing liberal, hierarchy-attenuating ideologies and greater racial intolerance for those embracing conservative, hierarchy-enhancing ideologies.

The Consequences of “Thinking about Race”

The idea that encouraging effortful thought can reduce prejudice emanates largely from work on when racial cues in political messages are most likely to trigger racial attitudes. This work focuses on political attitudes broadly rather than race-targeted policy attitudes. Mendelberg (2001) developed a framework describing the conflict that most White Americans face when they evaluate a racialized political message. According to this framework, the racial appeals most likely to be influential are those that only implicitly prime race because when race is primed explicitly, people are cued to reject prejudiced instincts and abide by egalitarian norms encouraged by society. However, when race is primed implicitly, the argument is that prejudiced instincts can influence political attitudes uninhibited. This model is supported by an array of empirical evidence showing that a wide range of implicit racial cues can trigger race-based

evaluations of political messages, and that the mechanism implicit cues operate through is by making race cognitively accessible (Mendelberg 2001; Valentino, Hutchings, and White 2002). The assumptions of this model are therefore that implicit racial cues trigger uncontrollable racial biases, that explicit racial cues trigger higher-level cognitive processes, and that higher-level cognitive processes triggered by explicit racial cues necessarily yield inhibition of prejudice. The implication is that when race is explicit (as it always is with race-targeted policies), prejudice influences race-targeted policies because some individuals simply do not engage in controlled inhibition. As explained by Valentino, Hutchings, and White, “For most people, the argument goes, negative racial attitudes affect political thinking automatically or not at all” (2002, 77).

Despite evidence supporting this framework, some findings call its assumptions into question. The effects of implicit racial cues are diminished when the cue presents a racial group in a counter-stereotypic way (Valentino, Hutchings, and White 2002) – a finding explained as due to greater controlled processing of counter-stereotypic information (Brewer, Dull, and Lui 1981; Hastie 1981). Also, the greater ability of implicit racial cues rather than explicit racial cues to trigger racial attitudes when evaluating a political message seems concentrated only among the highly educated (Huber and Lapinski 2006). These findings suggest that although the framework constructed by Mendelberg (2001) may be somewhat robust, there are significant limits to its generalizability. It seems that implicit cues do not *necessarily* lead to uncontrollable racial biases. Further, the finding that the model holds mainly among the highly educated suggests that the effects of making race explicit depend on the person: among highly educated individuals it seems likely that controlled processing will reduce the effects of prejudice, but less educated individuals may hold egalitarian values a bit more loosely, and so controlled processing

may often yield an equal or greater role for prejudice. Indeed, work on “motivation to respond without prejudice” suggests that individuals vary a great deal in the degree to which they are internally or externally motivated to use controlled processes to inhibit prejudice, even when race is explicit (e.g. Plant and Devine 1998). Not everyone holds the egalitarian views assumed to yield controlled inhibition when racial appeals are explicit. Nonetheless, this assumption remains largely accepted: that higher-level, controlled processing universally yields prejudice inhibition. The best way to reduce the role of prejudice then is to make race explicit and encourage greater controlled processing.

In this study, I explore the possibility that equal levels of controlled processing can nonetheless yield variation in the degree to which prejudice influences policy opinions. For some, according to this hypothesis, controlled processing leads to inhibition of prejudice because of higher-level principles that call for hierarchy attenuation. For others, controlled processing simply reinforces and even exacerbates the role of prejudice because of higher-level motivations to enhance hierarchy and *reject* egalitarian norms.

A Neural Framework of the Automatic and Controlled Components of Prejudice

A variety of methodologies have been used to estimate the influence of racial prejudice on political attitudes, but these methods rely heavily on assumptions regarding the mechanisms by which prejudice influences policy attitudes. To understand *how* prejudice influences such attitudes and thus how its role can be diminished, the most promising avenue of research is to directly examine the hypothesized mechanisms at work in the brain. Advances in medical technology have made it possible to obtain quantitative measurements of brain activity by tracking the increases in oxygenated blood that accompany increases in activity in particular

parts of the brain. Functional Magnetic Resonance Imaging (fMRI) uses a magnetic field in conjunction with pulses of radio frequencies to track the iron that comes with movement in oxygenated blood (Friston 2009).

There is one simple, overarching finding of research using fMRI methodology – the human brain is extremely complex, and to suggest that any attitude or behavior can be traced solely to a single region of the brain, or conversely that any single region of the brain has just one behaviorally relevant function, would be tremendously misguided. Nonetheless, research has advanced our understanding of neural activity substantially, and we can look at trends in brain activity in response to specific stimuli to infer which parts of the brain are *generally* active in particular contexts, including contexts related to politics (see Schreiber 2011). Through replication and convergence, we can develop frameworks of specific neural processes so brain activity can be used to predict “real-world,” ecologically valid outcomes (Berkman and Falk 2013). Neuroscience research on prejudice has provided an excellent starting point for understanding the mechanisms by which prejudice might influence policy attitudes. Findings from this work have essentially left us with a roadmap for how the brain functions when prejudice occurs automatically versus when controlled processes inhibit racial biases.

Group-based categorization seems hardwired into the brain’s processing of social information. Some work has shown that even when group membership is meaningless (i.e. created within the confines of an experimental task), the brain reacts to fellow ingroup members by activating certain regions associated with awareness of self (Volz et al. 2009; Shkurko, 2012; Molenberghs et al. 2013; Van Bavel, Packer, and Cunningham 2008). With regard to reactions to outgroups, the most commonly activated brain regions include the amygdala and insula,

suggesting that people immediately categorize outgroup members as emotionally significant (e.g. Cunningham et al. 2004; but also see Wheeler and Fiske 2005) and associate them with uncertainty or danger, which must be avoided (Rilling et al. 2008; Risheson et al. 2003; see Shkurko 2012 for meta-analysis). This reaction – that is, the activation of these brain regions – occurs automatically and even when race is only primed implicitly (Cunningham et al. 2004; Cunningham et al. 2007).

Something interesting happens, however, when race is consciously recognized and there is a conflict between people's implicit adverse reactions to outgroups and their explicit goals of being or appearing unprejudiced (e.g. "I feel threatened by the Black man walking toward me but know I should not"): the anterior cingulate cortex (ACC) detects the goal conflict, the amygdala response is suppressed, and certain parts of the prefrontal cortex (PFC) kick in that are associated with higher-level cognitive functioning (see Stanley, Phelps, and Banaji 2008). Controlled processes are thus triggered when race makes the jump from covert to overt, and these processes can allow people to inhibit prejudiced reactions. This basically means the precise mechanisms assumed to be occurring under the framework of this dissertation are detectable when observing brain data. Overall, we are left with relatively distinct patterns of neural activity when someone reacts adversely to a racial outgroup member, detects a conflict between this adverse reaction and one's conscious goals regarding race, and subsequently attempts to supersede adverse reactions through controlled processing. Note that conflict detection and controlled processing to inhibit adverse reactions should only be evident when race is explicit enough so as to be consciously recognized; otherwise, automatic reactions should occur uninhibited (Cunningham et al. 2004).

Critically, the research on the aforementioned neural processes does not suggest that controlled processing in the PFC will *necessarily* yield inhibition of prejudiced reactions – just that it is necessary for such inhibition to be possible. An alternative possibility, which has not been explored thoroughly in the neuroscience literature let alone in political science, is that controlled processes ultimately lead individuals to reinforce their pre-existing biases as a form of rationalization. In terms of the brain, PFC activation may be necessary for inhibition to occur, but in some situations, it may act as a catalyst for “doubling down” on prejudiced instincts. The implication for reducing the role of prejudice is glaring: encouraging people to think more about the racial component of race-targeted policies will not diminish the influence of prejudice and may even bolster its influence because people will turn to ideological principles that may either enhance or attenuate group hierarchy.

Expectations for the Role of Controlled Processing

The central test of this chapter has to do with the expected role of controlled processing in people’s race-related political evaluations. I expect that broadly, controlled processing will be associated with greater deference to ideological principles that either enhance or attenuate hierarchy, and so individuals who hold most strongly to hierarchy-enhancing or hierarchy-attenuating ideologies should experience the greatest amount of controlled processing when evaluating race-related political stimuli. I test this broad hypothesis using data from an fMRI experiment during which participants make decisions about whether or not individuals should receive monetary aid from the government. The race of the individuals applying for aid varies between Black and White, and race is primed either implicitly or explicitly. By priming race

either implicitly or explicitly, I am able to simulate the conditions under which we should expect automatic versus controlled processing related to race.

In line with the research described earlier in this chapter, it should be expected that when race is primed implicitly, individuals will exhibit greater automatic processing when evaluating Black applicants for government assistance compared to White applicants. When race is primed explicitly, individuals should exhibit greater controlled processing when evaluating Black applicants compared to White applicants. This has been established in prior research (e.g. Cunningham et al. 2004; Shkurko 2012). The critical test in relation to this dissertation has to do with the correlates of controlled processing when race is explicit. When race is explicit, I expect anti-Black racial biases in support for government assistance to be diminished among those who endorse hierarchy-attenuating ideologies and exacerbated among those who endorse hierarchy-enhancing ideologies. Further, because my model proposes that these effects of priming race explicitly are due to controlled processing among both “hierarchy-attenuators” and “hierarchy-enhancers,” controlled processing when race is explicit should be greater for those at the tail-ends of the ideological spectrum (i.e. those who are more ideologically extreme) compared to those who are more ideologically moderate. The critical tests of this study in terms of relevance to this dissertation are therefore constrained to involving activation in the cortical regions of the brain such as the ACC and PFC (i.e. the brain regions associated with controlled processing), and so only results pertaining to these brain regions are reported.¹⁶

¹⁶ Nonetheless, results regarding brain regions associated with automatic processing (e.g. amygdala and insula) are available upon request.

Data and Methods

Twenty-Three White adults were recruited to participate in the study through flyers placed around the community surrounding a large Midwestern university, as well as a volunteer participant registry created by the facility housing the MRI. Participants were each compensated \$25. All participants were right-handed and had normal or corrected-to-normal vision. Participants were safety screened to ensure eligibility for MRI and provided informed consent in accord with study approval by the Institutional Review Board. The sample was 60.9% female, 65.2% Christian, and 100% U.S.-born with English as a first language. The median age was 22 years old (range: 19 to 57) and 39.1% of the sample had a 4-year college degree or higher. Some screening was also done during recruitment to ensure the sample was adequately diverse in terms of political ideology and partisan affiliation. Nonetheless, the sample exhibited a liberal/Democratic skew (16 participants identified as liberal or strongly liberal, 1 identified as moderate, and 6 identified as conservative or strongly conservative; 13 participants identified as Democrats, 1 identified as Independent, 7 identified as Republican, and 2 identified as “something else”).

Experimental Design

The experimental task used a 2 (*race of applicant*: White vs. Black) x 2 (*conscious awareness*: non-conscious/implicit vs. conscious/explicit) rapid event-related within-subjects design consisting of 4 blocks. *Race of applicant* varied at the trial-level (i.e. within blocks) and *conscious awareness* varied at the block-level (i.e. between blocks). There were 40 trials within each block for a total of 160 trials. Each block lasted approximately 7 minutes. Experimental stimuli were presented in the scanner using PsychoPy2 v1.83.01 (Peirce, 2007; 2009). Both trial

order and interstimulus interval (ISI) duration were predetermined using Optseq2 (<https://surfer.nmr.mgh.harvard.edu/optseq>), a software package that maximizes efficiency in modeling the hemodynamic response in rapid event-related designs (see Burock, Buckner, Woldorff, Rosen, & Dale, 1998; Dale, Greve, & Burock, 1999).

Each trial presented participants with a random piece of information about the applicant and asked them to decide whether they strongly oppose, oppose, support, or strongly support monetary aid to that individual using response pads (one response pad for each hand, with buttons under each of participants' pointer and middle fingers). Response options were shown at the top of the screen throughout the task, and their positions on the screen remained constant across trials. Descriptions were constructed in such a way that they would not explicitly indicate any sort of deservingness or merit on the part of the applicant (e.g. "has an outie belly button" or "hums when eating"; see Appendix for complete list of statements). All applicants were males in order to avoid any effects of applicant sex. The phrase "A man that:" was shown on the screen just below the response options at all times.

Prior the description of each individual, a Black or White face was presented for an amount of time that either allowed for conscious recognition of the face (*Explicit* blocks) or not (*Implicit* blocks). The order of *Conscious* and *Nonconscious* blocks was randomized for each participant. All faces were taken from the Chicago Face Database (Ma, Correll, and Wittenbrink, 2015), a database of high-resolution standardized photographs of 158 Black and White individuals between the ages of 18-40. Each image has been coded for 41 physical characteristics (e.g. nose width, face width to height ratio, eye closeness) and rated along 22 perceptual dimensions (e.g. threatening, attractive, masculine) by an independent sample of raters, and was matched as

closely as possible across these characteristics and dimensions. Only 80 images were used (40 Black faces and 40 White faces), and so each image was shown twice (once during an *Explicit* block and once during an *Implicit* block). Images were randomly paired with applicant descriptions for each participant.

Each trial began with a fixation cross shown for 500 milliseconds (ms). In *Explicit* blocks, this was followed by an image of a Black or White face for 1 second, and then an applicant description for 3.5 seconds. In *Implicit* blocks, the fixation cross was followed by a noise mask (a picture of random black and white pixels) for 934 ms, a Black or White face for 30 ms, a second noise mask for 36 ms, and then an applicant description for 3.5 seconds. As such, *Explicit* and *Implicit* trials were identical except that whereas *Explicit* trials showed a Black or White face for a full second before each applicant description, *Implicit* trials showed noise masks during that second with a Black or White face shown rapidly in between masks. Thus, participants were expected not to be consciously aware of the influence of the Black or White face in the *Implicit* trials. Prior research has explored similar protocols and shown that a short stimulus presentation time such as the one being used here for the implicit face primes is sufficient to yield only nonconscious recognition of a stimulus (see Rohr, Degner, and Wentura 2015). The exact timings used for this experiment were pre-tested on a small sample of undergraduate students as well as during pilot runs in the fMRI scanner.

Survey Measures

Participants were in the scanner for approximately 45 minutes to 1 hour. Before being removed from the scanner, participants were asked a series of questions about whether they were able to see faces during the *Implicit* blocks (8 participants reported being able to see faces, and

11 said they were able to make out the race of faces seen between noise masks). Afterward, they completed a 15-minute long survey, which included, among other things, the 3 race-based policy support items used in chapters 2 and 3 ($\alpha = 0.89$; $M = 5.36$, $SD = 1.44$), SDO (SDO-D: $\alpha = 0.82$, $M = 2.19$, $SD = 0.98$; SDO-E: $\alpha = 0.90$, $M = 2.27$, $SD = 1.11$), and demographic items. SDO-D and SDO-E are used as the primary measures of hierarchy-attenuating/hierarchy-enhancing ideological principles. Both SDO scales were folded at their midpoint so that higher values indicate greater ideological extremity in either direction (SDO-D: $M = 1.14$, $SD = 0.54$; SDO-E: $M = 1.21$, $SD = 0.51$).

MRI Data Acquisition

A Siemens Skyra 3.0 Tesla MRI with a 32-channel head coil was used to collect brain activation data. Prior to functional imaging, a high-resolution T1-weighted 3D anatomical image (MPRAGE; *field of view (FoV) read* = 256 millimeters, *slice thickness* = 1.0 x 1.0 x 1.0 mm, *repetition time (TR)* = 2400 ms, *echo time (TE)* = 3.37 ms, *inversion time (TI)* = 991 ms, *prescan normalize on*, *PAT mode GRAPPA*) was obtained for spatial normalization. Functional MRI data was acquired with acquisition parallel to the AC-PC line to maximize whole-brain coverage (42 slices, *FoV read* = 220 mm, *slice thickness* = 3.0 x 3.0 x 3.0 mm, *TR* = 2500 ms, *TE* = 30 ms, *flip angle* = 80 degrees, *prescan normalize off*). The first 12.5 seconds of each block showed a blank grey screen so that the first five volumes of each block could be discarded to avoid variability due to pre-steady state functional data.

MRI Data Preprocessing and Analysis

MRI data were preprocessed using fMRI Expert Analysis Tool (FEAT) in FMRIB Software Library (FSL; Jenkinson, Beckmann, Behrens, Woolrich, & Smith, 2012; Smith et al.,

2004) on Mac OS X. The MPRAGE image was skull stripped using FSL's Brain Extraction Function (BET; Smith, 2002). Functional data were subjected to normalization, registration to both MPRAGE and standard space (MNI152), spatial smoothing at FWHM of 5mm, slice timing correction (to correct for interleaved data acquisition), and motion correction using MCFLIRT (Jenkinson, Bannister, Brady, & Smith, 2002).

General linear models (GLMs) were used as implemented in FSL. Time-series data was modeled at the first level (the trial level) using FMRIB's Improved Linear Model (FILM), and then higher-level analysis (across blocks first, and then across participants) was done using FMRIB's Local Analysis of Mixed Effects (FLAME; see Smith et al., 2004). The blood oxygen level-dependent (BOLD) signal was first modeled at the trial level for each run as a function of *race* (Black/White). At the block-level, BOLD signal differences for Black compared to White trials were modeled as a function of *conscious awareness* (Explicit/Implicit). The subject-level analyses were then combined into group-level region of interest (ROI) analyses using FSL FLAME1. ROI analyses on left amygdala, right amygdala, bilateral insula, anterior cingulate cortex (ACC), and frontal orbital cortex (OFC) were masked prior to analysis (using anatomical masks from the Harvard-Oxford Cortical/Subcortical Atlases provided with FSL). Due to the small sample size, these analyses remained uncorrected for multiple comparisons.¹⁷ Analyses of the *race*conscious awareness* contrasts (before taking into account between-subject variables) were used to examine whether controlled processing was, as expected, more likely for Black compared to White trials in *Explicit* compared to *Implicit* blocks. To examine how brain activation mapped onto individual differences in SDO extremity, BOLD signal differences for

¹⁷ As such, the results presented in this chapter are preliminary and should not be cited without consulting the author.

the *race*conscious awareness* contrasts were modeled as a function of each of the SDO extremity variables (analyses for SDO-D and SDO-E were carried out separately).

Results

Behavioral Task Results

Before turning to the fMRI results, analyses of participants' responses to the task were done to determine whether support for government assistance was influenced by race, conscious awareness, and SDO in the expected directions. Multilevel models were run with trial as the unit of analysis nested within blocks, nested within participant. An empty model was run estimating random effects for participant, block, trial number, and face image (because each image was shown twice). Only the random effects for participant explained a substantial amount of variation in support for government assistance (approximately 10% of the variance), and so all subsequent models include only random effects for participant. All models also controlled for fixed effects of participant's sex, education, religion (Christian or non-Christian), and whether participants reported being able to decipher the race of faces shown in between noise masks in the Explicit blocks.

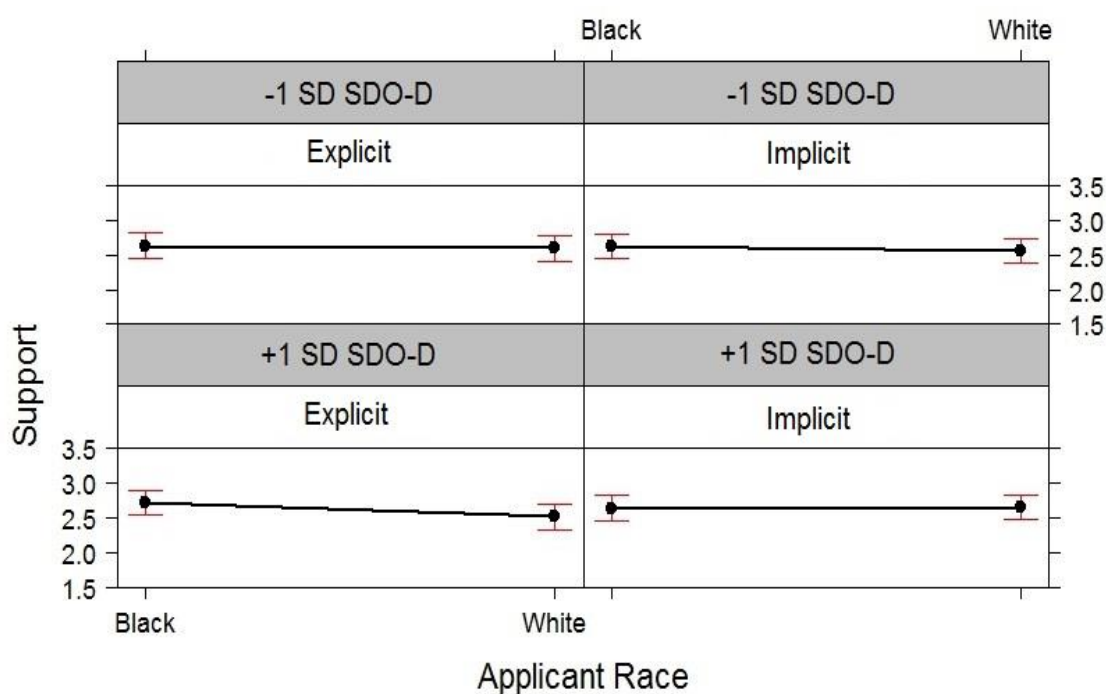
First, I examined the relationship between racial biases in support for government assistance during the task and support for race-based policies as measured during the post-scan survey in order to determine the degree to which the task reflected evaluations of "real-world" policies. Support for government assistance during the task was regressed on a dummy variable for *conscious awareness* condition (0 = Explicit, 1 = Implicit), a dummy variable for race of the applicant (0 = Black, 1 = White), race-based policy support (mean-centered), and the interaction between race of applicant and race-based policy support. Indeed, there was a significant

interaction between race of applicant and race-based policy support ($b = -0.052, SE = 0.022, p < .05$). Once decomposed, it was evident that at low levels of support for race-based policies (-1 SD from the mean), there were no significant racial biases in support for government assistance during the task ($b = 0.011, SE = 0.048, p = .811$), but at high levels of support for race-based policies (+1 SD from the mean), there was a significant pro-Black bias ($b = -0.135, SE = 0.044, p < .01$). This suggests that although opposition to “real-world” race-based policies was not associated with anti-Black bias during the task, support for race-based policies was indeed associated with pro-Black bias during the task.

The primary behavioral analyses concern the interaction between race of applicant and conscious awareness. A main effects model indicated no significant main effect of conscious awareness, but there was a significant main effect of applicant race such that on average, participants were less likely to support White applicants than Black applicants ($b = -0.069, SE = 0.033, p < .05$). Next, an interaction was calculated between applicant race and conscious awareness. The interaction was marginally significant ($b = 0.106, SE = 0.057, p = .062$), and showed that although a pro-White racial bias existed in Explicit blocks ($b = -0.122, SE = 0.044, p < .01$), no significant racial bias existed in Implicit blocks ($b = -0.017, SE = 0.043, p = 0.696$). Thus, although the lack of racial bias in Implicit blocks was unexpected, more racially tolerant evaluations were observed in Explicit blocks as hypothesized.

The final, critical, behavioral analysis relating to the hypotheses of this chapter involved a three-way interaction between applicant race, conscious awareness, and SDO. Separate models were run testing SDO-D and SDO-E. With regard to SDO-D, a significant three-way interaction was evident. At mean levels of SDO-D, there is a significant pro-Black bias ($b = -0.120, SE =$

0.040, $p < .01$), and this bias is marginally stronger in Explicit blocks than in Implicit blocks. At low levels of SDO-D, there is a stronger pro-Black bias ($b = -0.399$, $SE = 0.143$, $p < .01$), but this bias is *only* significant in the Explicit blocks ($b = 0.516$, $SE = 0.200$, $p < .01$). Finally, among those high in SDO-D, there is no significant racial bias in either Implicit or Explicit blocks. Therefore, although individuals high in SDO-D unexpectedly showed no evidence of racial biases regardless of conscious awareness of race, individuals low in SDO-D exhibited the expected pattern of greater racial tolerance in Explicit compared to Implicit blocks. Figure 5.1 illustrates this interaction.



The y-axis indicates support for government assistance, and applicant race is plotted on the x-axis. Left panels indicate Explicit blocks and right panels indicate Implicit blocks. Top panels indicated -1 SD SDO-D, and bottom panels indicate +1 SD SDO-D.

Figure 4.1: Interaction between Applicant Race, Conscious Awareness, and SDO-D

The same three-way regarding SDO-E was not significant. However, both two-way interactions comprising the three-way were significant. An applicant race by SDO-E interaction

showed that as should be expected, greater levels of SDO-E were associated with lower levels of pro-Black bias ($b = 0.066$, $SE = 0.026$, $p < .05$). Unexpectedly, a conscious awareness by SDO-E interaction showed that greater levels of SDO-E were also associated with greater support for government assistance in the Explicit blocks compared to the Implicit blocks, regardless of race ($b = -0.067$, $SE = 0.026$, $p < .05$).

Taken together, these behavioral results suggest that contrary to expectations, the predominant effect of applicant race was pro-Black/anti-White. Even among individuals high in SDO, there were no racial biases rather than pro-White racial biases. However, in line with expectations, and more importantly with regard to the framework of this dissertation, conscious awareness of race (i.e. explicit racial primes) was associated with more pro-Black evaluations, but only among individuals low in SDO-D (although this three-way interaction was not evident for SDO-E).

fMRI Results

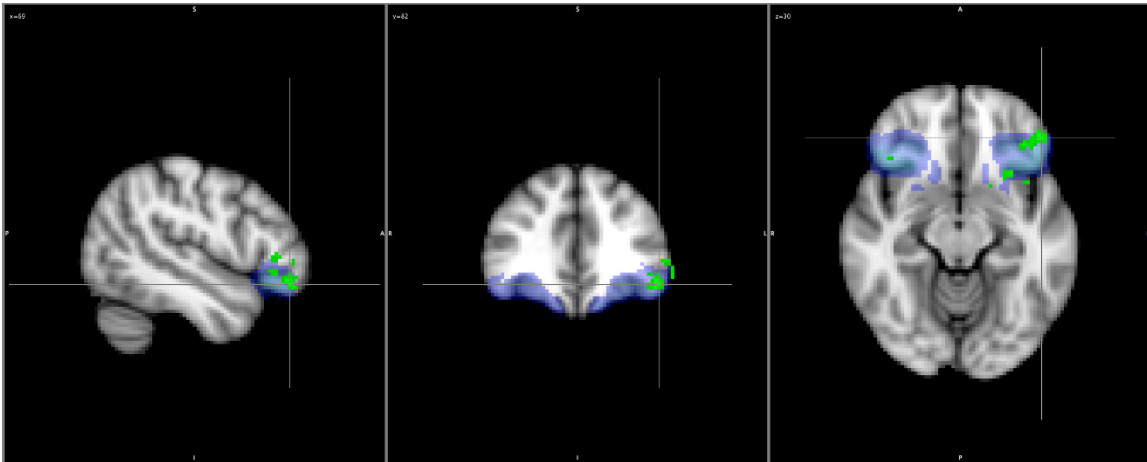
Two sets of analyses were conducted. First BOLD signal was modeled as a function of the interaction between applicant race and conscious awareness to see which brain regions exhibited significant activation for Black (compared to White) faces in Explicit blocks (compared to Implicit blocks) as well as Implicit (compared to Explicit) blocks. Then, the SDO extremity variable was added to the model to see which brain regions in the above contrast exhibited significant activation that was associated with SDO extremity. ROI analyses were done for each contrast for ACC and OFC in order to limited analyses to only the brain regions implicated in the critical tests of this chapter. Further, SDO analyses were carried out separately for SDO-D and SDO-E.

In the aggregate, there was a small but significant cluster of activation in ACC for Implicit compared to Explicit trials, and contrary to the expectation that ACC would be more active in Explicit trials, significant activation was also found in ACC for Explicit compared to Implicit trials but it was only a single voxel. Thus, ACC activation was evident in both types of trials for Black compared to White faces, but this activation was stronger in Implicit trials. Regarding OFC, significant clusters of activation were evident for both contrasts (i.e. in both Explicit and Implicit trials), but as expected, activation was stronger for the Explicit > Implicit contrast than for the Implicit > Explicit contrast. Therefore, although ACC activation did not behave as expected, OFC activation did. These results are shown in Table 5.1, and the significant clusters of OFC associated with Explicit compared to Implicit trials are depicted in Figure 5.2.

Table 5.1: Significant Clusters of Bold Activation in ROIs for Applicant Race*Conscious Awareness

Contrast	Anatomical Label(s)	Cluster Size (# Voxels)	Peak Activation (Z-Score)	p-value	X	Y	Z
<i>ACC</i>							
Implicit > Explicit	52% ACC, 19% Juxtapositional Lobule Cortex, 2% Cingulate Gyrus, posterior division, 1% Precentral Gyrus	12	2.153	0.031	48	59	57
Explicit > Implicit	76% Paracingulate Gyrus, 13% ACC, 1% Superior Frontal Gyrus	1	1.697	0.090	43	80	52
<i>OFC</i>							
Implicit > Explicit	24% Frontal Pole, 13% OFC	17	2.305	0.021	37	81	26
Explicit > Implicit	57% Frontal Pole, 25% OFC, 1% Inferior Frontal Gyrus, pars triangularis	170	2.537	0.011	69	82	30

All contrasts are interactions with applicant race and thus represent activation in Black face compared to White face trials. X, Y, and Z indicate the coordinates of activation in MNI152 space.



Images were created by overlaying the thresholded Z -statistic image on a standard space template (MNI152). Images are centered on the peak voxel for each cluster from the ROI analyses. Areas highlighted in blue indicate the ROI for the analysis, and areas highlighted in green indicate regions of significant activation.

Figure 5.2: BOLD Activation in OFC in Response to Explicit > Implicit Blocks for Black > White Trials

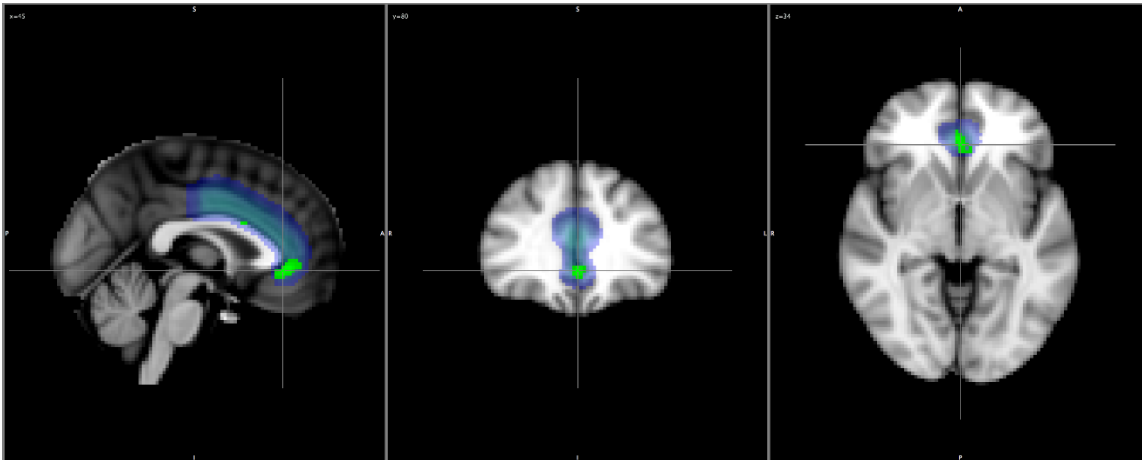
If the theoretical framework of this dissertation is correct, besides an interaction between applicant race and conscious awareness, it should also be evident that activation in ACC and OFC in Explicit (compared to Implicit) blocks for Black face (compared to White face) trials is associated with greater extremity in one's ideological principles – in this case, SDO. Analyses were first conducted for SDO-D extremity. Small yet significant clusters of activation in ACC were found for both contrasts, suggesting some degree of ACC activity occurred in both Implicit and Explicit blocks was associated with greater SDO-D extremity. However, contrary to expectations, the cluster for the Implicit > Explicit contrast predicting SDO-D extremity was somewhat stronger. Conversely, patterns of activation in OFC supported my hypotheses. That is, although significant activation in OFC was found to be associated with greater SDO-D extremity for both contrasts, this activation was somewhat stronger for the Explicit > Implicit contrast (despite involving fewer voxels).

Regarding SDO-E, there were significant clusters of activation associated with SDO-E extremity across all contrasts for both ACC and OFC, but the relative strengths of activation varied entirely as expected. For ACC, activation associated with high SDO-E extremity was stronger for Explicit (compared to Implicit) blocks than it was for Implicit (compared to Explicit) blocks. Further, this same relative difference in strength of activation was evident for OFC. These associations between brain activation patterns and SDO-E extremity strongly support the hypothesis that greater controlled processing when race is explicitly (rather than implicitly) primed is associated with greater adherence to ideological principles regardless of the direction of those principles. These results are presented in Table 5.2, and the clusters of ACC and OFC for Explicit (compared to Implicit) blocks that are significantly associated with SDO-E extremity are illustrated in Figures 5.3 and 5.4, respectively.

Table 5.2: Significant Clusters of Bold Activation in ROIs for Applicant Race*Conscious Awareness*SDO

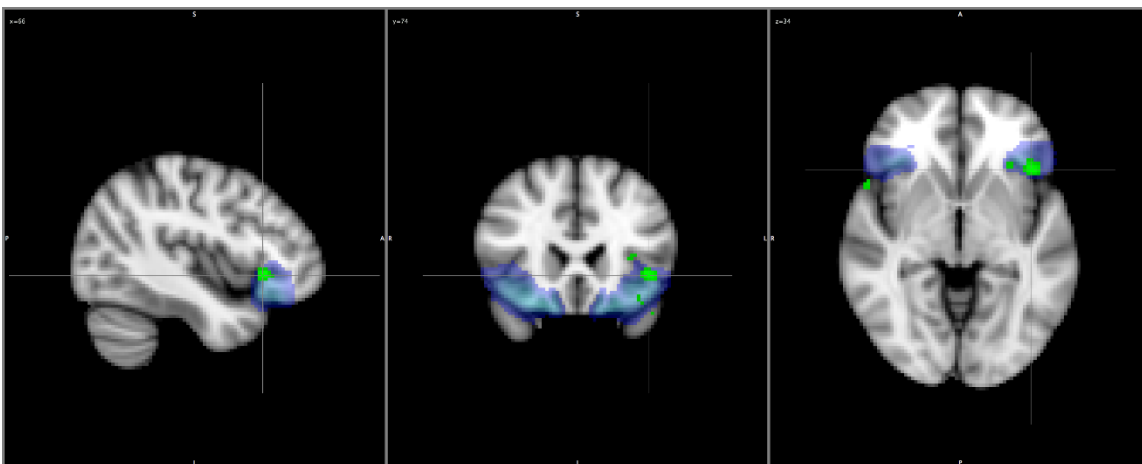
Contrast	Anatomical Label(s)	Cluster Size (# Voxels)	Peak Activation (Z-Score)	p-value	X	Y	Z
SDO-D							
<i>ACC</i>							
Implicit > Explicit	30% Paracingulate Gyrus, 23% ACC	28	2.046	0.041	39	72	53
Explicit > Implicit	19% ACC, 3% Paracingulate Gyrus	3	1.822	0.069	38	83	39
<i>OFC</i>							
Implicit > Explicit	25% OFC, 16% Frontal Pole	26	2.038	0.042	21	79	27
Explicit > Implicit	11% OFC, 3% Subcallosal Cortex	11	2.338	0.019	53	72	30
SDO-E							
<i>ACC</i>							
Implicit > Explicit	31% Paracingulate Gyrus, 16% ACC	13	1.973	0.049	39	73	53
Explicit > Implicit	66% ACC, 9% Paracingulate Gyrus, 9% Subcallosal Cortex	114	2.451	0.014	45	80	34
<i>OFC</i>							
Implicit > Explicit	11% OFC, 4% Frontal Pole	11	2.160	0.031	21	79	26
Explicit > Implicit	42% OFC, 28% Frontal Operculum Cortex, 4% Insular Cortex, 1% Inferior Frontal Gyrus, pars triangularis	159	2.966	0.003	66	74	34

All contrasts are interactions with applicant race and thus represent activation in Black face compared to White face trials. X, Y, and Z indicate the coordinates of activation in MNI152 space.



Images were created by overlaying the thresholded Z -statistic image on a standard space template (MNI152). Images are centered on the peak voxel for each cluster from the ROI analyses. Areas highlighted in blue indicate the ROI for the analysis, and areas highlighted in green indicate regions of significant activation.

**Figure 5.3: BOLD Activation Associated with SDO-E Extremity
In ACC in Response to Explicit > Implicit Blocks for Black > White Trials**



Images were created by overlaying the thresholded Z -statistic image on a standard space template (MNI152). Images are centered on the peak voxel for each cluster from the ROI analyses. Areas highlighted in blue indicate the ROI for the analysis, and areas highlighted in green indicate regions of significant activation.

**Figure 5.4: BOLD Activation Associated with SDO-E Extremity
In OFC in Response to Explicit > Implicit Blocks for Black > White Trials**

Inside the Black Box of “Principles and Prejudice”

The results presented in this chapter are mostly consistent with the model laid out in this dissertation. The findings presented in chapters 3 and 4 have largely corroborated a theoretical

framework in which race-related political objects trigger automatic processes associated with ingroup favoritism but controlled processes allow individuals to defer to ideological principles. However, up to this point, the neural mechanisms thought to be responsible for deference to ideological principles have yet to be directly observed. The use of fMRI allowed for such observation, and an experiment was developed to distinguish between neural processes occurring during explicitly race-related political evaluations and neural processes occurring during political evaluations that are only implicitly related to race. Prior research in social neuroscience has established that whereas implicit racial cues are associated with automatic processing (e.g. amygdala and insula), explicit racial cues are associated with more controlled processing (e.g. ACC and PFC), which can lead to decreases in racial bias (e.g. Cunningham et al. 2004). However, it remained unknown whether controlled processing universally leads to inhibition of racial biases or to deference to ideological principles, as my model suggests.

The behavioral task results suggested that individuals were, as expected, less likely to express bias against a racial outgroup when racial primes were explicit than when they were implicit. However, the precise pattern of results was somewhat unexpected. Contrary to the hypothesis that anti-Black racial biases would be strongest during Implicit blocks, no anti-Black racial biases in support for government assistance were observed anywhere in these data. This could have been due to multiple reasons. First, although brain activation patterns indicated that participants did respond differently to Black than White faces in Implicit blocks, it could be the case that the primes were shown too fast for any behavioral consequences to manifest. This would be somewhat surprising given the pilot testing as well as the fact that quicker primes have been used in prior research to produce racial biases (e.g. Rohr, Degner, and Wentura 2015).

Another possibility is simply limited sample size. Although the within-subjects design allowed for a large sample in terms of trials, the limited between-subject variation may have made it difficult to detect anti-Black biases. Nonetheless, pro-Black biases were evident only in Explicit blocks, which suggests that on average, explicit racial primes led to more pro-Black evaluations.

The behavioral task results involving SDO told a similar story. Unexpectedly, individuals high in SDO-D or SDO-E exhibited no racial bias. Although this was contrary to expectations, it makes some sense for SDO-E because the ideology promoted by high SDO-E is thought to involve endorsement of existing group inequality rather than a preference to enhance or exacerbate inequality (Jost and Thompson 2000; Kugler, Cooper, and Nosek 2010). In that sense, a pro-Black bias among those low in SDO may indicate an acknowledgement of existing racial inequality and efforts to address it rather than simply a lack of racial prejudice. However, SDO-D is thought to involve active degradation of groups perceived to be low in the societal hierarchy, and so the lack of racial bias among individuals high in SDO-D remains unexplained. This lack of bias may, as with the lack of anti-Black bias in the aggregate, be due to sample size issues. Nonetheless, SDO-D interacted with applicant race and conscious awareness in the expected direction, as pro-Black biases in support for government assistance among individuals low in SDO-D were primarily evident during Explicit blocks rather than Implicit blocks.

The primary contribution of this chapter centers on the fMRI results. It is first worth noting that across analyses, significant brain activation in both the ACC and OFC was evident for both Implicit and Explicit blocks, suggesting that in line with the IR model (Cunningham, Haas, and Jahn 2011; Cunningham and Zelazo 2007; Cunningham, Zelazo, Packer, and Van Bavel 2007), controlled processes seem to be active to some degree at all stages of race-related

evaluation (i.e. when race is implicitly primed as well as when it is explicitly primed). The main tests of the hypotheses of this chapter therefore involve relative strength of activation between Explicit > Implicit contrasts and Implicit > Explicit contrasts.

In the aggregate, brain activation patterns in the OFC followed expectations, as activation for Black versus White face trials was substantially stronger in Explicit (compared to Implicit) blocks than in Implicit (compared to Explicit) blocks. However, only low levels of activation in ACC for Black versus White face trials were observed to differ across blocks, and activation was somewhat stronger for the Implicit > Explicit contrast. This result for ACC was unexpected, but may fit with understandings of the functions of the ACC. The findings described earlier in this chapter suggest the ACC plays sort of a mediating a role in racial evaluations by detecting conflict between automatic reactions and conscious beliefs. If this is the case, it would not be surprising if conflict detection was occurring during Implicit blocks, and by the time evaluations were made during Explicit blocks, conflicts had already been detected and the OFC kicked in to regulate one's response. However, this finding would still be contrary to some existing work (e.g. Cunningham et al. 2004). It could also be the case that the ACC was responding simply to the "strangeness" of the Implicit trials, because a series of flashing black and white noise masks were shown prior to every applicant description.

The fMRI analyses regarding SDO were done using extremity variables because my model suggests controlled processing should be evident when race is primed explicitly for individuals at either end of the ideological spectrum. My hypotheses for SDO were supported more strongly for SDO-E than for SDO-D, which is not surprising given the fact that SDO-E is thought to be associated with redistributive policies, as described in Chapter 2. Both ACC and

OFC showed stronger activation associated with SDO-E extremity in the Explicit > Implicit contrasts than in the Implicit > Explicit contrasts. This same pattern was evident for SDO-D regarding activation in OFC but not ACC. However, even with regard to the OFC results, the activation associated with SDO-D extremity in the Explicit > Implicit contrast was stronger than the activation in the Implicit > Explicit contrast but was observed in fewer voxels (i.e. less anatomical space). ACC activation was more strongly associated with SDO-D extremity in the Implicit > Explicit contrast than in the Explicit > Implicit contrast, and in all analyses activation was relatively weak. These results regarding SDO-D are not terribly concerning given that prior research has suggested SDO-D is more directly related to social identity motives than SDO-E (Jost and Thompson 2000; Kugler, Cooper, and Nosek 2010), and SDO-E has been the primary workhorse of group-based principles throughout this dissertation, but further research should nonetheless investigate these differences in the neural underpinnings of SDO-E versus SDO-D.

CHAPTER 6

A Neural Toolkit for Change, Anchored in the Past

“Nihilists! Fuck me. I mean, say what you want about the tenets of National Socialism, Dude, at least it’s an ethos.”

- Walter Sobchak (John Goodman) in *The Big Lebowski*¹⁸

The question posed at the beginning of this dissertation was: how much control do people have over their racial attitudes and behaviors? A review of the literature, in conjunction with the results presented throughout this dissertation, suggests the answer is (in true academic fashion): it depends. It depends largely on the degree to which automatic versus controlled processes have influenced one’s attitudes or behavior. From work in social neuroscience, we know that attitudes exist on a spectrum from “more influenced by automatic processes” to “more influenced by controlled process,” and that where a particular attitude falls on this spectrum is largely a function of the degree to which over time, controlled processes are able to interact with automatic processes and thus influence evaluations (Cunningham, Haas, and Jahn 2011; Cunningham and Zelazo 2007; Cunningham, Zelazo, Packer, and Van Bavel 2007). Therefore, when it comes to the topic of this dissertation – race-related political attitudes – we should expect controlled, conscious processes to be quite powerful, but also for people’s attitudes to be grounded substantially in innate reflexes.

I have argued that group-based principles such as SDO are the most dominant predictors of race-related political attitudes because they reflect the translation process between people’s automatic group biases and the controlled processes that allow them to interact with mass-scale

¹⁸ Coen, J. and Coen, E. (Writers and Directors). (1998). *The Big Lebowski* [Motion picture]. United States: Polygram Film Entertainment and Working Title Films.

society. More specifically, as people think about race in the domain of politics and how societal resources should be allocated across racial groups, automatic ingroup favoritism is translated through controlled processes into higher-order principles about how groups should be organized in society. Accordingly, people's attitudes toward policies such as affirmative action and race-based education spending are not simply the result of whether or not the policy benefits their group, nor are they the result solely of their beliefs about deservingness and the proper role of government. Instead, race-targeted policy attitudes seem primarily driven by individuals' preferences regarding how groups in society should be organized hierarchically. The results of the studies presented in this dissertation have largely supported this theoretical framework.

In the remainder of this chapter, I elaborate on the implications of the theory and results of this dissertation. I start by summarizing the findings of each chapter. Then, I expand on the implications of this dissertation for the "principles or prejudice" literature, and how an acknowledgement that "principles" and "prejudice" operate together can benefit this literature. Next, I dive deeper into the distinct consequences of group-based principles as opposed to sheer ingroup favoritism. Turning to the "principled" component of group-based principles, I talk about what this dissertation suggests principles really are, and how this conceptualization of principles differs from the conceptualizations of prior work. I then describe how the findings of this dissertation call for more cross-fertilization between the NMSoT and CCSoT. Ultimately, I suggest this dissertation holds tremendous implications for how we might go about reducing the role of prejudice in politics, and so I conclude by discussing ways in which public policy may be informed by the theoretical model of this dissertation.

Summary of Results

In Chapter 2, I surveyed the landscape by pitting survey measures of various constructs against one another to see which factors were most strongly associated with race-based policy attitudes. Using data from a moderately sized online sample, I looked at three broad categories of factors: social identity motives, individualism, and group-based principles. I hypothesized that group-based principles, measured as SDO-E and SDO-D, would outperform the social identity motives (i.e. the role of ingroup favoritism) and individualism in predicting attitudes toward race-based policies. This hypothesis was strongly supported. SDO-E, in particular, was the foremost predictor of race-based policy attitudes. Social identity motives played a minor role, and the only independent effects of social identity motives were positive relationships between ingroup importance and commitment and support for race-based government assistance, which was unexpected. As for the role of individualism, work ethic and self-reliance both maintained weak but significant independent relationships with race-based policy attitudes. However, an experimental manipulation of how hardworking Blacks versus Whites were perceived to be had no effect on policy attitudes. Further, contrary to some existing research (Federico and Sidanius 2002; Sidanius, Pratto, and Bobo 1996), the hypothesis that the role of individualism is explained by group-based preferences was not supported.

These results suggest first and foremost that group-based principles rather than ingroup favoritism or individualistic values underlay race-targeted policy opinions. This finding, in particular, is central to the theoretical framework outlined in this dissertation. These findings also suggest that “principles,” as traditionally conceptualized in the “principles or prejudice”

literature, are not simply socially acceptable vehicles for the expression of group biases but are also not necessarily “pure” indicators of deservingness-based evaluations either.

In Chapter 3, I utilized the Minimal Groups Paradigm (MGP) to strip away any context, history, and stereotypes that might be attached to “real-world” groups and thus obtain purer estimates of the relationships between ingroup favoritism, perceptions of deservingness, group-based principles, and race-based policy attitudes. I started by replicating the correlational results of Chapter 2 in a smaller, student sample (except for the unexpected results regarding ingroup importance and commitment, which did not replicate). SDO-E was the strongest predictor of race-based policy attitudes even when the measures of social identity motives were specific to the minimal groups in the experiment. In order to manipulate the perceived deservingness of the minimal groups, and thus gauge the degree to which individuals evaluated groups based on merit rather than sheer ingroup favoritism, I experimentally manipulated whether the participant was told their ingroup was relatively good or bad at a task they completed during the study. In the aggregate, participants showed significant ingroup bias and no sensitivity to the merit manipulations. However, most importantly for the purposes of this dissertation, the only component of the minimal groups task that significantly predicted race-based policy attitudes was *implicit* ingroup favoritism. Further, SDO was associated with implicit ingroup favoritism *as well as* some instances of explicit ingroup favoritism and sensitivity to the merit manipulations in a way indicative of merit-based evaluation.

These results are important for two reasons. First, they suggest ingroup favoritism is not entirely unrelated to race-based policy attitudes, but instead its role is implicit and based in

automatic processes. Second, group-based principles seem to represent some combination of automatic (and to some degree controlled) group biases and higher-order principles, as expected.

Although Chapter 3 demonstrated that automatic group biases are associated with race-based policy opposition, and group-based principles seem to embody some combination of ingroup bias and deservingness-based values, I was nonetheless unable to test the hypothesis that individuals become more “principled” in their evaluations as controlled processes interact with automatic group biases. Therefore, in Chapter 4, I used nationally representative data from the 2008 ANES to gauge implicit racial ambivalence – or the discrepancy between one’s implicit and explicit racial attitudes – as it exists in the United States and its consequences for how people evaluate race-related political issues. Psychology work has shown that implicit ambivalence is associated with greater controlled processing aimed at resolving dissonance associated with the target of ambivalence, and so I use implicit ambivalence as a measure of the degree to which individuals’ controlled and automatic preferences have interacted. I hypothesized that individuals low in implicit racial ambivalence – i.e. those who are the most “resolved” in their racial attitudes – would be the most likely to rely on higher-order principles such as ideology and egalitarianism to evaluate race-related political objects.

In support of this hypothesis, this was the case with regard to the role of ideology across a range of race-related political outcome variables. The role of egalitarianism was also moderated by implicit racial ambivalence but counter to expectations, such that the effects of egalitarianism were exacerbated at high levels of implicit ambivalence. I suggest this is because the egalitarianism items are inherently group-based and so this interaction actually depicts greater reliance on group biases at high implicit ambivalence.

These results suggest that the degree to which individuals defer to higher-order principles when evaluating race-related political objects is, as hypothesized, largely moderated by the degree to which they have used controlled processing to “adjust” their automatic biases. Further, the result of the interaction between controlled processing and automatic biases is not always inhibition of racial biases or more liberal policy attitudes. In half of the sample, explicit preferences were more biased relative to the mean of the explicit bias measure than implicit preferences, and having resolved racial preferences was associated with ideological polarization rather than simply more liberal policy positions. Returning to the data from the experiment in Chapter 3, I also was able to construct an implicit ambivalence variable using the implicit and explicit attitudes measures in the MGP and show that, as expected, individuals low in implicit ambivalence toward the minimal groups (i.e. those with more resolved attitudes) were the least likely to express ingroup favoritism behaviorally in the MGP.

In Chapter 5, I directly examined the neural mechanisms thought to be at play throughout this dissertation by using fMRI. Specifically, I examined brain activation patterns in cortical areas of the brain – the ACC and OFC, in particular – when people evaluated applicants for government assistance who varied in race as well as whether their race was primed explicitly or implicitly. I found that in terms of decisions to support or oppose government assistance, contrary to expectations, no racial biases were apparent when race was primed implicitly, but more importantly for the purposes of this dissertation, pro-Black evaluations were greater when race was primed explicitly than implicitly, and this pattern was exacerbated among individuals low in SDO-D (but not SDO-E). In terms of fMRI results, I found that contrary to expectations, activation in ACC was greater when race was primed implicitly. However, in line with

expectations, OFC activation was greater when race was primed explicitly, and this pattern was exacerbated at high levels of SDO extremity (i.e. both ends of the SDO-E and SDO-D spectrums). This same pattern was evident for SDO-D for OFC, but not for ACC.

Principles, Prejudice, or Principled Prejudice?

The questions addressed in this dissertation stemmed largely from the “principles or prejudice” literature. The central question of the “principles or prejudice” literature is: are people’s attitudes toward race-targeted policies like affirmative action due more to racial prejudice or individualistic values? Research aimed at answering this question has spanned decades, and ultimately the result has been a vast array of mixed findings.

On the one hand, proponents of constructs like *symbolic racism* (Kinder and Sears 1981; Sears and Henry 2003) and *modern racism* (McConahay 1986) have argued that racial prejudice has become inherently fused with conservative, individualistic values, and that “principled” arguments against policies like affirmative action are in fact socially acceptable means of expressing racial prejudice (Bobo and Smith 1998; Kinder and Mendelberg 2000; Kinder and Sander 1996; Sears, van Laar, Carrillo, and Kosterman 1997). An array of empirical findings support these claims (e.g. Gilens 1999; Henry and Sears 2003; Rabinowitz et al. 2009; Sears et al. 1997). On the other hand, other research argues that constructs like symbolic racism confound racial prejudice with nonracial conservative values, and that when accounted for separately, individualism explains the bulk of variation in race-targeted policy attitudes (Sniderman and Carmines 1997; Sniderman et al. 1996).

I have proposed that the mixed findings that beset the “principles or prejudice” literature are largely due to limited, unidimensional conceptualizations of what “principles” and

“prejudice” *are*. Scholars have primarily been interested in whether racial biases underlay opposition to race-targeted policies or not, and so a simple dichotomy between racial prejudice and individualism has been functional. However, the limited conceptualizations of principles and prejudice (as well as the mechanisms of each) have prevented the literature from advancing along with work in political psychology and social neuroscience. By decomposing the multiple dimensions comprising common conceptualizations of prejudice and individualism, I have derived critical tests that may advance the “principles or prejudice” literature.

One critical test that a limited conceptualization of principles and prejudice has obscured is whether racial biases in support for government assistance (i.e. effects of prejudice) are due to social identity motives – motivations to favor ingroups and derogate outgroups – or broad ideologies about how groups in society should be organized hierarchically. The implications of this test are straightforward. If race-targeted policy attitudes are primarily due to ingroup favoritism, it suggests a sort of self-interest at play whereby individuals are mainly out for their own groups’ (and thus their own selves’) best interest and so will support policies that benefit their group and oppose policies that benefit other groups. Members of disadvantaged groups are thus, as Jost, Banaji, and Nosek put it, “revolutionaries-in-waiting” (2004, p. 883), who are in constant conflict with members of advantaged groups who cling onto their disproportionate share of resources. In terms of race-based policy attitudes, this suggests variation in support for policies that assist Blacks is a function of variation in race as well as the degree to which people identify with their racial ingroup. However, the findings in this dissertation suggest the role of prejudice is explained mainly by group-based principles, and so opposition to race-based policies is less about any kind of self-interest and more about efforts to instill order on society (Chapters

2 and 3). What is important is the system rather than the ingroup, and individuals who believe in maintaining group-based hierarchy will oppose government assistance to *any* disadvantaged group, regardless of whether or not they are a member of that group.

Another critical test allowed by a more comprehensive view of “principles” and “prejudice” is of what sorts of principles are most important with regard to race-based policy opinions. In the existing literature, the term “principles” has been used primarily to reflect the influence of any values not indicative of racial animus. Some work has deferred primarily to Weber’s ideas about the Protestant Work Ethic, or a high valuation of hard work and the belief that hard work is the key to success; other work has defined individualism more in terms of Tocqueville’s or Emerson’s conceptualizations of self-reliance and isolation from mass-scale society; still other work explicitly talks about both versions of individualism and then measures only one dimension of it. The implications of which version of individualism actually underlays people’s race-based policy opinions are substantial because it essentially determines which sort of principles a government assistance policy might be framed around in order to manipulate support for that policy.

Results from this dissertation suggest that the most dominant form of principles driving race-based policy attitudes is group-based principles. Work ethic and self-reliance maintain independent but weak relationships with race-targeted policies when controlling for group-based principles (Chapters 2 and 3), and the only evidence of deservingness-based evaluations found in this dissertation was among individuals high in SDO (Chapter 3). Existing work thus seems to have missed the mark to some degree in postulating that the principles of individualism are the

primary values at work when people evaluate race-based policies. Instead, the principles people seem to be basing their opinions on have to do with how groups should be arranged in society.

A final critical test made possible by broadening our view of “principles or prejudice” is of the mechanisms that drive the effects of ingroup favoritism, principles, and group-based principles on race-based policy attitudes. By acknowledging that principles and prejudice are multidimensional, interrelated constructs, we can ask the question: *when* are people more likely to evaluate government assistance policies based on group biases versus ideological principles? I have argued that based on research in social neuroscience as well as evolutionary psychology, we should expect the role of ingroup favoritism to be driven primarily by automatic processes and the role of higher-order principles to be driven primarily by controlled processes.

Indeed, I found that only implicit measures of ingroup favoritism predicted race-based policy opposition (Chapter 3), and individuals who had utilized controlled processing to resolve discrepancies between their implicit and explicit racial attitudes (i.e. those low in implicit racial ambivalence) were the most likely to rely on ideology to evaluate a range of race-related political outcomes (Chapter 4). Further, group-based principles – the dominant predictor of race-based policy opinions – seem to consist of a combination of these forces (Chapter 3). Finally, by studying the neural processes thought to underlay automatic ingroup biases versus ideological thinking directly using fMRI, it was found that racial biases aligned with one’s ideology (in terms of SDO) most when race was primed explicitly, and that individuals at either extreme of SDO were most likely to exhibit brain activation in OFC (and to some degree ACC) when race was primed explicitly, which suggests ideological thinking is exacerbated by controlled processing (Chapter 5). Understanding the neural mechanisms of race-targeted policy attitudes

allows us to understand the potential for change in these attitudes, as well as the sorts of interventions most likely to create attitude change. In other words, instead of simply looking at the degree to which race-targeted policy opinions are primarily principled or prejudiced, we can start to investigate how to reduce the role of prejudice that exists.

Other Group-Based Principles, Symbolic Racism, and System Justification

In this dissertation, the primary operationalization of group-based principles was SDO (more specifically, the two subdimensions of SDO-D and SDO-E). However, it is reasonable to wonder what other sorts of group-based principles exist and the role they would play in influencing race-targeted policy attitudes. Although SDO is likely one of the most influential constructs representing group-based principles, I would not suggest it is the end-all-be-all of group-based principles, or that the term “group-based principles” is simply a synonym for SDO. Instead, I use the term group-based principles to represent conscious beliefs about how groups in society should be organized, and I argue that they stem from interactions between innate group biases and higher-order values and beliefs.

SDO is one manifestation of group-based principles and reflects the degree to which people think groups should be arranged hierarchically and high-status groups should dominate subordinate groups. However, it is also possible to imagine other group-based principles such as beliefs about how rigid group boundaries should be, whether diversity should be tolerated within groups, how conflicts between groups should be addressed (e.g. through hostility or diplomacy), whether groups should be integrated or stay separate from one another, or whether groups should even be defined in the first place. Only SDO is tested in this dissertation, and its dominant role in predicting race-based attitudes shows that broadly, beliefs about how groups *should* be organized

matter more than positive feelings toward ingroups. Yet if group-based principles comprise the interaction between automatic group biases and higher-order principles, it should be expected that this interaction results in other sorts of group-based principles as well, and so the roles of other group-based principles should be tested in future research. In fact, different types of group-based policies may be driven by different group-based principles. Race-targeted government assistance such as affirmative action and race-based education spending has most directly to do with the distribution of resources across groups in a country, but other policies may be more aimed at integration between groups, equal protection under and enforcement of the law, or conflicts between warring nations of relatively equal status in terms of global hierarchy.

Symbolic Racism: Not So “New”

Given the theoretical framework developed in this dissertation, I would also suggest that constructs such as symbolic racism in some ways tap the effects of group-based principles. As explained earlier, proponents of symbolic racism argue that it reflects a fusion between racial prejudice against Blacks and individualistic values. The theoretical model developed in this dissertation accounts for such a fusion. Specifically, the model suggests symbolic racism is a form of group-based principles specific to Blacks in the contemporary United States. Americans – especially White Americans – have internalized negative feelings toward Blacks, but these negative feelings run counter to social pressures to be unprejudiced, and so as Kinder and Mendelberg (2000) suggest, individualism becomes a socially acceptable vehicle for the expression of anti-Black prejudice. Indeed, implicit racial biases as well as explicit racial biases and ideology have been shown to contribute to symbolic racism (e.g. Ditonto, Lau, and Sears 2013).

The model I propose interprets the origins of symbolic racism (and thus the implications of its effects) differently in three major ways. First, my model suggests that despite the fact that group biases and higher-order principles often interact, they are nonetheless separable components of group-based principles. Indeed, although non-group-based principles such as work ethic and self-reliance did not play a substantial role throughout my studies, they did at times exhibit independent relationships with race-based policy attitudes.

Second, my model suggests the construct of symbolic racism is unnecessarily constrained to attitudes toward Blacks, because although the unique history of Blacks in the United States has undoubtedly influenced the particular stereotypes that are applied to them as a group, the interaction between principles and prejudice is not unique to attitudes toward Blacks and in fact is perfectly explained by evolutionary processes. Said differently, the processes underlying symbolic racism have likely been around for millennia, and so the term “new racism,” as symbolic racism has often been referred to, is in some ways misleading. My model proposes that since humans have interacted with one another in small-scale groups, automatic group-based adaptations have interacted with controlled processes that respond to the nuances of the immediate social context.

Second, my model suggests the degree to which constructs like symbolic racism (and group-based principles more broadly) are driven by group biases versus higher-order principles like individualism depends on the amount of controlled processing applied to the attitude object (in this case, Blacks). As demonstrated in Chapter 4, the degree to which political ideology predicted constructs including symbolic racism was moderated by how resolved their implicit and explicit racial attitudes were.

SDO or System Justification?

Another construct that should be addressed in relation to group-based principles is system justification. System Justification Theory (SJT) posits that individuals vary in the degree to which they strive to defend and rationalize the existing social, economic, and political systems as well as the status quo more broadly (Jost and Banaji 1994; Jost, Banaji, and Nosek 2004; Jost, Kay, and Thorisdottir 2009;). SJT has been incredibly influential and supported across a range of empirical tests in a variety of domains (e.g. Balabanis and Diamantopoulos 2016; Blasi and Jost 2006; Feygina, Jost, and Goldsmith 2010; Jost et al. 2012; Jost and Kay 2005; Kay and Jost 2003; Montieth, Burns, Rupp, and Mihalec-Adkins 2016).

A critical difference between SJT and group-based theories revolving around ingroup favoritism or ethnocentrism is that according to SJT, members of disadvantaged groups will often favor the status quo even when it works against them because the structure and order provided by the system alleviates the distress associated with thinking the world is unfair (Jost and Banaji 1994; Jost and Burgess 2000). This tenet of SJT is nearly identical to what differentiates SDO from social identity motives. Although some work has recognized this similarity between SJT and SDO (Jost and Thompson 2000; Kugler, Cooper, and Nosek 2010), a thorough examination of the differences between the two constructs remains to be done and the most notable difference is simply that SDO has explicitly to do with groups as the unit of analysis whereas SJT does not necessarily use groups as the unit of analysis. Future research is needed to parse out the distinct implications of SJT compared to SDO, as SJT was not measured in this dissertation.

In fact, the model in this dissertation may be applied to SJT as a potential explanation of the mechanisms underlying system justification. Are system justification motives driven by the interaction between innate group biases and higher-order principles? Some SJT work suggests system justification motives can manifest implicitly (e.g. Jost, Banaji, and Nosek 2004; Jost, Pelham, and Carvalho 2002), suggesting the automatic components of system justification are not necessarily indicative of sheer ingroup favoritism. However, a comprehensive analysis of the degree to which system justification is driven by automatic or controlled processes has yet to be done. From an evolutionary standpoint, it makes sense that “the system” the brain is responding to automatically is not mass-scale society, but instead the reflection of some system that could exist in small-scale groups. My model suggests the brain might utilize controlled processes to adjust these automatic responses to address the social, political, and economic “systems” of contemporary mass-scale society.

Rules for Radicals, and Reactionaries Too

Principles, values, ethos, belief systems, philosophies, dogmas, ideologies, moral codes; regardless of what you call them, higher-order belief structures seem important to understanding attitudes toward government assistance. Through this dissertation, I have attempted to unify, to some degree, the divergent conceptualizations of “principles” that would be posited at the extreme ends of the NMSoT and CCSoT. The implications of my proposed model for how we should conceptualize principles flies in the face of the extreme forms of both schools of thought.

I propose that reliance on principles to evaluate political objects stems from evolutionary adaptations aimed at adjusting behavioral rules to accommodate the needs of a complex

immediate environment. This conceptualization of principles is based on work in social neuroscience on the functions of the prefrontal cortex (PFC) regarding complex behavioral rules used by humans (Bunge and Zelazo 2006; Zelazo, Frye, and Rapus 1996; Zelazo, Muller, Frye, and Marcovitch 2003), and on work in evolutionary psychology on how political cognition largely emanates from adaptations suited for small-scale group interactions (Petersen 2014; Petersen and Aarøe 2012; 2013; Petersen et al. 2013). Essentially, my model's conceptualization of principles comes from acknowledging that controlled processes are driven by millennia-old neural hardware that primarily serve the function of allowing humans to respond to a complex and rapidly changing environment in a nuanced way. Small-scale group interactions did not call for behavioral guidelines that extended beyond a handful of tribes, yet the contemporary political world often necessitates belief systems that apply universally and span the globe. Ultimately, I therefore suggest controlled processes do not necessarily simply lead people to express more egalitarian views. Instead, controlled processes lead people to adopt and adhere to a set of principles, whether those principles are progressive, conservative, radical, reactionary, capitalist, socialist, hierarchy-attenuating, hierarch-enhancing, or ideological in some other way.

Acknowledging Factors Formerly “Beyond Our Scope”

This conceptualization of what principles represent is somewhat counter to the tenets of the CCSoT because it suggests that when individuals utilize controlled processing, it does not necessarily entail the deliberate weighing of costs and benefits or active choosing of one value system over another, and it certainly does not indicate processing that is detached from innate predispositions. As mentioned throughout this dissertation, most scholars – even those working mainly within the CCSoT – have acknowledged that predispositions and evolutionary forces

likely play a role in shaping political opinions but neglected to study such forces directly due to the lack of tools available for studying them (e.g. Zaller 1992). However, the tools for studying automatic processes, innate predispositions, and nonconscious motivations are no longer beyond the scope of political research, and so it seems overdue that a model accounting for these factors is applied to the study of race-targeted policy opinions.

Social Desirability “Influence,” Not “Bias”

The conceptualization of principles offered in this dissertation also runs counter to some models that have focused on NMSoT factors. As referenced in the epigraph of the first chapter in this dissertation, Stephen Pinker (2008) has suggested that just because science has been uncovering the mysteries associated with what he refers to as the “temptation systems” in the brain, it does not mean we need to disregard the “inhibition systems” in the brain. In other words, the research by scholars of the NMSoT has advanced our understandings of automatic processes tremendously, but this does not mean we should neglect the role of controlled processes. Most scholars primarily focused on studying NMSoT factors have recognized this (e.g. Smith et al. 2011). As mentioned in the caveats outlined in Chapter 1, my use of the term “controlled processing” should not be taken to indicate control in a metaphysical sense or “free will.” However, I do suggest (in line with work in social neuroscience) that controlled processes are more likely to reach conscious awareness. Individuals are sometimes, though not always, able and even willing to self-report their attitudes (mostly) “accurately.” Especially with regard to politics, individuals take enough time before reporting their attitudes that controlled processing are able to lay substantial claim to evaluations of a political object, whether it be a policy, a candidate, a party, or an issue.

Some work has explicitly acknowledged the role of controlled processing but in a way that suggests it is less valuable than the role of automatic processing. One thing some researchers are particularly wary of with regard to surveys, and an example often held up as representing the limitations of self-report measures, is social desirability “bias.” The argument regarding social desirability “bias” is often that social desirability leads individuals to report responses that are somehow disingenuous. Social desirability, the argument goes, leads people to obfuscate or skew their self-reported attitudes away from their “true” attitudes, which are deep-seated, in order to appease conformity pressures or abide by perceived social norms (e.g. Greenwald et al. 2009). Or, alternatively, it is sometimes argued that people are simply unaware of their “true” attitudes. However, the model proposed in this dissertation suggests controlled processes are an inherent component of attitudes. Further, the concept of a single, coherent “true” attitude does not fit well with the model I propose, because the model suggests that due to the constant iterative process between automatic and controlled processes, any particular attitude will manifest differently depending on when and how it is expressed. In fact, evaluations that are heavily influenced by controlled processes are expected to differ substantially from evaluations made primarily at the automatic level, because over time controlled processes should adjust evaluations of an attitude object to fit the needs of the immediate environment (e.g. Cunningham and Zelazo 2007). Even in cases when individuals are consciously aware of how they feel yet report something different to exude a socially acceptable image, I suggest that this is nonetheless an example of controlled processing adjusting one’s evaluations (albeit evaluations that have already been somewhat influenced by controlled processing). Ultimately, this suggests social desirability “bias” would be better worded as social desirability “influence.”

The model in this dissertation also holds substantial implications for what effects we should expect from controlled processes and social desirability influence. A common assumption of NMSoT work is that when people utilize controlled processing or tailor self-reports in light of social pressures, they tailor their attitude toward being more egalitarian and unprejudiced. A common assumption of scholars using implicit measures is that they are useful because at the conscious level, people realize it is socially unacceptable to be racially prejudiced so adjust their self-reports accordingly (e.g. Greenwald et al. 1998; Greenwald et al. 2009; Pasek et al. 2010; Payne et al. 2005; Payne et al. 2010; see Fazio and Olson 2003). According to Tali Mendelberg (2001), implicit racial cues in political campaigns are the most effective at evoking prejudice because when racial cues are explicit, controlled processes allow individuals to inhibit racial biases and abide by egalitarian norms (see also Valentino, Hutchings, and White 2002). Even work in social neuroscience suggests that controlled neural processes are associated with individuals inhibiting automatic racial biases (Cunningham et al. 2004). However, if controlled processes are simply utilized to adapt automatic processes to the social pressures of the immediate environment, then should it be expected that controlled processes always lead to less prejudiced, egalitarian, liberal views? I suggest not.

Instead, I argue controlled processing can lead people to mimic culturally learned stereotypes, support group hierarchy, or defer to anti-egalitarian principles instead of egalitarian ones. Indeed, the results of Chapter 4 support this proposition. Although ingroup favoritism was diminished at low levels of implicit racial ambivalence (i.e. among those who had used controlled processing to resolve their racial attitudes), controlled processing seemed to lead some individuals (approximately half the sample in those data) to self-report racial attitudes that were

more (relatively) biased than their implicit racial attitudes. Further, the controlled processing associated with low implicit racial ambivalence was also related to a stronger role of political ideology. In other words, individuals who presumably underwent the most controlled processing to resolve their racial attitudes were not the most likely to report liberal race-based policy attitudes, but instead were the most likely to report race-based policy attitudes in line with their ideology. The result of having resolved racial attitudes was ideological polarization rather than universal inhibition of prejudice.

Given this broader understanding of controlled processing, I argue that the impact of social desirability is not greater egalitarianism and inhibition of group biases but instead the application of some broader set of principles to one's automatic group biases. For some, this means appearing unprejudiced, and for others, it means "telling it like it is" and rejecting "political correctness." For some, this means expressing pro-social attitudes, and for others, it means abiding by the "bootstraps" model of individualism. For some, it means striving to attenuate hierarchy, and for others, it means striving to enhance it. As advocated by Walter Sobchak in *The Big Lebowski* in the quote at the beginning of this chapter, what really seems to be universally socially desirable is just being "principled" in some way.

Nudging Away Prejudice?

One of the ultimate goals of the model developed in this dissertation is to inform efforts to reduce the role of prejudice in politics. Cass Sunstein and Richard Thaler, in their book *Nudge: Improving Decisions about Health, Wealth, and Happiness*, suggest that through psychologically informed public policy, relatively minor adjustments in how policies are developed and implemented can allow government to take a proactive role in increasing the well-

being of the public while also refraining from becoming invasive (2008; see also Sunstein 2013). If decreasing the role of racial prejudice in politics is a desirable outcome (at least for many people), then an understanding of the mechanisms of the role of prejudice could be quite useful to policymakers. The model developed in this dissertation offers several tangible tools for nudging people's attitudes toward race-targeted policies.

This dissertation advances our understandings of the mechanisms of how prejudice influences political attitudes – when people are most likely to abide by higher-order ideological principles when evaluating race-related political objects rather than resorting to innate, primal ingroup biases. Indeed, according to the model, humans are not helplessly bound by millennia-old motivations to favor ingroups over outgroups. There is room for change. However, humans are also not detached from these millennia-old forces. Controlled processes allow individuals to adjust automatic, reflexive group biases in accordance with higher-order principles. The major role controlled processes are argued to play offers some room for optimism among those interested in reducing the role prejudice plays in driving policy opinions.

Structuring (Not Simply Encouraging) Racial Thought

Two of the major heuristics Sunstein and Thaler suggest are useful for policymakers are framing, which structures the way in which individuals think about a topic, and priming, which influences what issues come to mind when people evaluate a topic (see e.g., Iyengar and Kinder 1987; Kahneman and Tversky 1984). In this dissertation, it is argued that prejudice and ingroup favoritism are least likely to drive people's attitudes toward race-targeted policies when they have time to think about it, when controlled processes have a chance to “reiterate” automatic evaluations, and when individuals have had the opportunity to use controlled processes to resolve

disparities that exist between their gut-level preferences and their consciously held values. In terms of public policy, then, it seems efforts to frame policies in a manner that encourages effortful thought or prime relatively complex aspects of an issue may be useful in discouraging the role of prejudice.

However, I suggest this conclusion would be problematic. As described above, it should not be assumed that controlled processing yields egalitarian, liberal political views. Instead, it may very well be the case that controlled processing can result in ideological polarization or even exacerbate racially prejudiced views. Getting people to “think more” about race or priming people to think about race explicitly has been upheld in much research as a tool for diminishing racial prejudice (e.g. Cunningham et al. 2004; Hurwitz and Peffley 2005; Mendelberg 2001; Valentino, Hutchings, and White 2002). Yet as mentioned in Chapter 5, some of these findings likely exist in the aggregate but are limited by a range of factors linked to whether one embraces egalitarian views, such as education (Huber and Lapinski 2006).

The results presented in this dissertation and the model more broadly suggest that although some controlled processing allows some individuals to abide by egalitarian norms, it also allows other individuals to abide by anti-egalitarian norms. The group-based principles people abide by can entail a range of beliefs, some of which explicitly exacerbate group-based inequality. As such, getting people simply to “think about race” is likely not enough to shift opinions in a conservative or liberal direction, or even in an unprejudiced direction. If the goal is simply to remove the influence of ingroup favoritism, the model suggests sheer encouragement of controlled processing will be effective, but racial biases may manifest through more “principled” means. If the goal is to diminish racial biases in support for government assistance,

or to encourage support for group-based equality more broadly, more direct restructuring of controlled processes through policy framing and priming is likely necessary.

Further, careful consideration should be taken not to prime individuals to think in terms that are easily translatable into small-scale, tribal interactions. The automatic processes underlying people's evaluations get "first dibs" when it comes to shaping policy attitudes, and so the easier it is for automatic group biases to take root, the less need there will be for controlled processing to step in. Indeed, Nelson and Kinder (1996) showed that group centrism was significantly more likely to play a role in people's policy opinions when issues were framed in a way that drew attention to the policy's beneficiaries. When it is easy for people to revert to thinking in terms of the ingroup-outgroup distinctions of millennia past, they likely will.

Leveraging Social Desirability Influence

Another psychological phenomenon Sunstein and Thaler suggest might be useful to policymakers is the spotlight effect, whereby individuals tend to overestimate the degree to which others notice and assign importance to their actions (Gilovich, Medvec, and Savitsky, 2000; Gilovich and Savitsky 1999). Since classic conformity studies like those conducted by Solomon Asch, Stanley Milgram, and Philip Zimbardo, social psychology research has shown that social pressure is a very powerful force (Asch 1956; Milgram 1964; Haney, Banks, and Zimbardo 1973), and can be leveraged even for political ends such as encouraging voter turnout (e.g. Gerber and Green 2008; Panagopoulos 2010). The model in this dissertation suggests individuals use controlled processing to translate their automatic group preferences into higher-order, group-based principles that address the demands of their complex social environment. As such, if the social environment demands egalitarianism then we should expect controlled

processing to push people toward egalitarianism. In this way, social desirability can be thought of as a tool for persuasion rather than a “bias” to be avoided. “Public shaming” has in many ways become a problem in social media (e.g. Whittaker and Kowalski 2015), but at the same time it is largely through social influence that norms and principles are created (e.g. Benford and Snow 2000). When individuals are “put on the spot” – when they perceive their attitudes and behaviors to have social consequences – they may be more likely to utilize controlled processing to adjust their deep-seated biases in a more egalitarian direction, *as long as* the social environment encourages egalitarianism.

BIBLIOGRAPHY

- Aldrich, J. H. (1993). Rational choice and turnout. *American Journal of Political Science*, 37(1), 246-278.
- Aldrich, J. H. (1995). *Why parties?: The origin and transformation of political parties in America*. Chicago, IL: University of Chicago Press.
- Alvarez, R. M., & Brehm, J. (1997). "Are Americans ambivalent towards racial policies?." *American Journal of Political Science*, 41(2), 345-374.
- Arkes, H. R., & Tetlock, P. E. (2004). "Attributions of implicit prejudice, or "would Jesse Jackson 'fail' the Implicit Association Test?"". *Psychological Inquiry*, 15(4), 257-278.
- Arrow, K. J. (1963). *Social choice and individual values*. New Haven, CT: Yale University Press.
- Asch, S. E. (1956). Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychological monographs: General and Applied*, 70(9), 1.
- Balabanis, G., & Diamantopoulos, A. (2016). Consumer Xenocentrism as Determinant of Foreign Product Preference: A System Justification Perspective. *Journal of International Marketing*, 24(3), 58-77.
- Baron, A. S., & Banaji, M. R. (2006). "The development of implicit attitudes evidence of race evaluations from ages 6 and 10 and adulthood." *Psychological Science*, 17(1), 53-58.
- Benford, R. D., & Snow, D. A. (2000). Framing processes and social movements: An overview and assessment. *Annual Review of Sociology*, 26(1), 611-639.
- Berinsky, A. J., Huber, G. A., and Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon. com's mechanical turk. *Political Analysis*, 20(3), 351-368.
- Berkman, E. T., & Falk, E. B. (2013). Beyond brain mapping using neural measures to predict real-world outcomes. *Current Directions in Psychological Science*, 22(1), 45-50.
- Bettencourt, B., Charlton, K., Dorr, N., & Hume, D. L. (2001). Status differences and in-group bias: a meta-analytic examination of the effects of status stability, status legitimacy, and group permeability. *Psychological Bulletin*, 127(4), 520.
- Blanton, H., Jaccard, J., Klick, J., Mellers, B., Mitchell, G., & Tetlock, P. E. (2009). "Strong claims and weak evidence: reassessing the predictive validity of the IAT." *Journal of Applied Psychology*, 94(3), 567.
- Blasi, G., & Jost, J. T. (2006). System justification theory and research: Implications for law, legal advocacy, and social justice. *California Law Review*, 94(4), 1119-1168.
- Bornstein, G., Crum, L., Wittenbraker, J., Harring, K., Insko, C. A., & Thibaut, J. (1983a). On the measurement of social orientations in the minimal group paradigm. *European Journal of Social Psychology*, 13(4), 321-350.
- Bornstein, G., Crum, L., Wittenbraker, J., Harring, K., Insko, C. A., & Thibaut, J. (1983b). Reply to Turner's comments. *European Journal of Social Psychology*, 13(4), 369-381.
- Brewer, M. B. (1999). The psychology of prejudice: Ingroup love and outgroup hate?. *Journal of Social Issues*, 55(3), 429-444.
- Brewer, M. B. (2001). Ingroup identification and intergroup conflict: When does ingroup love become outgroup hate?. In R. D. Ashmore, L. Jussim, and D. Wilder (Eds.), *Social*

- identity, intergroup conflict, and conflict reduction* (pp. 17-41). New York: Oxford University Press.
- Bobo, L. (1998). Race, interests, and beliefs about affirmative action unanswered questions and new directions. *American Behavioral Scientist*, *41*(7), 985-1003.
- Bobo, L., & Licari, F. C. (1989). Education and Political Tolerance: Testing the Effects of Cognitive Sophistication and Target Group Affect. *Public Opinion Quarterly*, *52*, 285–308.
- Bourhis, R. Y., Sachdev, I., & Gagnon, A. (1994). Intergroup research with the Tajfel matrices: Methodological notes. In *The psychology of prejudice: The Ontario symposium* (Vol. 7, pp. 209-232). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Brewer, M. B., Dull, V., & Lui, L. (1981). Perceptions of the elderly: Stereotypes as prototypes. *Journal of Personality and Social Psychology*, *41*(4), 656.
- Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk A New Source of Inexpensive, Yet High-Quality, Data?. *Perspectives on Psychological Science*, *6*(1), 3-5.
- Burock, M. A., Buckner, R. L., Woldorff, M. G., Rosen, B. R., & Dale, A. M. (1998). Randomized event-related experimental designs allow for extremely rapid presentation rates using functional MRI. *NeuroReport*, *9*(16), 3735-3739.
- Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, *42*(1), 116-131.
- Clifford, S., Jewell, R. M., & Waggoner, P. D. (2015). Are samples drawn from Mechanical Turk valid for research on political ideology?. *Research and Politics*, *2*(4).
- Commins, B., & Lockwood, J. (1979). The effects of status differences, favoured treatment and equity on intergroup comparisons. *European Journal of Social Psychology*, *9*(3), 281-289.
- Correll, J., Urland, G. R., & Ito, T. A. (2006). "Event-related potentials and the decision to shoot: The role of threat perception and cognitive control." *Journal of Experimental Social Psychology*, *42*(1), 120-128.
- Cunningham, W. A., Johnson, M. K., Raye, C. L., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2004). "Separable neural components in the processing of black and white faces." *Psychological Science*, *15*(12), 806-813.
- Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). "Implicit attitude measures: Consistency, stability, and convergent validity." *Psychological Science*, *12*(2), 163-170.
- Dale, A. M., Greve, D. N., & Burock, M. A. (1999). Optimal stimulus sequences for event-related fMRI. *NeuroImage*, *9*, S33-S33.
- Devine, P. G. (1989). "Stereotypes and prejudice: Their automatic and controlled components." *Journal of Personality and Social Psychology*, *56*(1), 5.
- Devine, P. G., Plant, E. A., Amodio, D. M., Harmon-Jones, E., & Vance, S. L. (2002). "The regulation of explicit and implicit race bias: the role of motivations to respond without prejudice." *Journal of Personality and Social Psychology*, *82*(5), 835.
- Ditonto, T. M., Lau, R. R., and Sears, D. O. (2013). AMPing racial attitudes: Comparing the power of explicit and implicit racism measures in 2008. *Political Psychology*, *34*(4), 487-510.

- Dodd, M. D., Balzer, A., Jacobs, C. M., Gruszczynski, M. W., Smith, K. B., & Hibbing, J. R. (2012). The political left rolls with the good and the political right confronts the bad: connecting physiology and cognition to preferences. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1589), 640-649.
- Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). "On the nature of prejudice: Automatic and controlled processes." *Journal of Experimental Social Psychology*, 33(5), 510-540.
- Dovidio, J. F., Mann, J., & Gaertner, S. L. (1989). Resistance to affirmative action: The implications of aversive racism. In F. A. Blanchard & F. J. Crosby (Eds.), *Affirmative Action in Perspective* (pp. 83-102). Springer New York.
- Downs, A. (1957). An economic theory of political action in a democracy. *Journal of Political Economy*, 65(2), 135-150.
- Downs, G. W., & Rocke, D. M. (1994). Conflict, agency, and gambling for resurrection: The principal-agent problem goes to war. *American Journal of Political Science*, 362-380.
- Ellemers, N., Wilke, H., & Van Knippenberg, A. (1993). Effects of the legitimacy of low group or individual status on individual and collective status-enhancement strategies. *Journal of Personality and Social Psychology*, 64(5), 766.
- Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). "Variability in automatic activation as an unobtrusive measure of racial attitudes: a bona fide pipeline?." *Journal of Personality and Social Psychology*, 69(6), 1013.
- Fazio, R. H., & Olson, M. A. (2003). "Implicit measures in social cognition research: Their meaning and use." *Annual Review of Psychology*, 54(1), 297-327.
- Fearon, J. D. (1994). Domestic political audiences and the escalation of international disputes. *American Political Science Review*, 88(03), 577-592.
- Feddersen, T., Gailmard, S., & Sandroni, A. (2009). Moral bias in large elections: theory and experimental evidence. *American Political Science Review*, 103(02), 175-192.
- Federico, C. M. (2004). "When do welfare attitudes become racialized? The paradoxical effects of education." *American Journal of Political Science*, 48(2), 374-391.
- Federico, C. M. & Sidanius, J. (2002). Racism, ideology, and affirmative action revisited: the antecedents and consequences of "principled objections" to affirmative action. *Journal of Personality and Social Psychology*, 82, 488-502.
- Feldman, S. (1988). Structure and consistency in public opinion: The role of core beliefs and values. *American Journal of Political Science*, 416-440.
- Feldman, S., & Huddy, L. (2005). Racial Resentment and White Opposition to Race-Conscious Programs: Principles or Prejudice?. *American Journal of Political Science*, 49(1), 168-183.
- Feldman, S., & Zaller, J. (1992). "The political culture of ambivalence: Ideological responses to the welfare state." *American Journal of Political Science*, 268-307.
- Feygina, I., Jost, J. T., & Goldsmith, R. E. (2010). System justification, the denial of global warming, and the possibility of "system-sanctioned change". *Personality and Social Psychology Bulletin*, 36(3), 326-338.
- Forgas, J. P., & Fiedler, K. (1996). Us and them: Mood effects on intergroup discrimination. *Journal of Personality and Social Psychology*, 70(1), 28.

- French, J. A., Smith, K. B., Alford, J. R., Guck, A., Birnie, A. K., & Hibbing, J. R. (2014). Cortisol and politics: variance in voting behavior is predicted by baseline cortisol levels. *Physiology & Behavior*, *133*, 61-67.
- Friston, K. J. (2009). Modalities, modes, and models in functional neuroimaging. *Science*, *326*(5951), 399-403.
- Furnham, A. (1990). *The Protestant work ethic: The psychology of work-related beliefs and behaviours*. New York: Routledge.
- Gaertner, S. L., and Dovidio, J. F. (1986). *The aversive form of racism*. Academic Press.
- Gawronski, B., & Bodenhausen, G. V. (2006). "Associative and propositional processes in evaluation: an integrative review of implicit and explicit attitude change." *Psychological Bulletin*, *132*(5), 692.
- Gerber, A. S., Green, D. P., & Larimer, C. W. (2008). Social pressure and voter turnout: Evidence from a large-scale field experiment. *American Political Science Review*, *102*(01), 33-48.
- Gilovich, T., Medvec, V. H., & Savitsky, K. (2000). The spotlight effect in social judgment: an egocentric bias in estimates of the salience of one's own actions and appearance. *Journal of Personality and Social Psychology*, *78*(2), 211.
- Gilovich, T., & Savitsky, K. (1999). The spotlight effect and the illusion of transparency: Egocentric assessments of how we are seen by others. *Current Directions in Psychological Science*, *8*(6), 165-168.
- Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, *96*(5), 1029.
- Green, D., & Shapiro, I. (1996). *Pathologies of rational choice theory: A critique of applications in political science*. New Haven, CT: Yale University Press.
- Greenwald, A. G., & Banaji, M. R. (1995). "Implicit social cognition: attitudes, self-esteem, and stereotypes." *Psychological Review*, *102*(1), 4.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). "Measuring individual differences in implicit cognition: the implicit association test." *Journal of Personality and Social Psychology*, *74*(6), 1464.
- Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). "Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity." *Journal of Personality and Social Psychology*, *97*(1), 17.
- Haney, C., Banks, W. C., & Zimbardo, P. G. (1973). A study of prisoners and guards in a simulated prison. *Naval Research Review*, *30*, 4-17.
- Harmon-Jones, E., & Harmon-Jones, C. (2008). "Action-Based Model of Dissonance: A Review of Behavioral, Anterior Cingulate, and Prefrontal Cortical Mechanisms." *Social and Personality Psychology Compass*, *2*(3), 1518-1538.
- Harmon-Jones, E., Amodio, D. M., & Harmon-Jones, C. (2009). "Action-based model of dissonance: A review, integration, and expansion of conceptions of cognitive conflict." *Advances in Experimental Social Psychology*, *41*, 119-166.
- Hastie, R. (1981). Schematic Principles in Human Memory. In E. Tory Higgins, C. Peter Herman, and Mark P. Zanna (Eds.), *Social Cognition: The Ontario Symposium*, Vol. 1 (pp. 39-88). Hillsdale, NJ: Erlbaum.

- Henry, P. J. & Sears, D. O. (2002). The symbolic racism 2000 scale. *Political Psychology*, 23, 253-283.
- Ho, A. K., Sidanius, J., Kteily, N., Sheehy-Skeffington, J., Pratto, F., Henkel, K. E., Foels, R., & Stewart, A. L. (2015). The Nature of Social Dominance Orientation: Theorizing and Measuring Preferences for Intergroup Inequality Using the New SDO7 Scale. *Journal of Personality and Social Psychology*, 109(6), 1003-1028.
- Ho, A. K., Sidanius, J., Pratto, F., Levin, S., Thomsen, L., Kteily, N., & Sheehy-Skeffington, J. (2012). Social Dominance Orientation Revisiting the Structure and Function of a Variable Predicting Social and Political Attitudes. *Personality and Social Psychology Bulletin*, 38(5), 583-606.
- Hofmann, W., & Baumert, A. (2010). Immediate affect as a basis for intuitive moral judgement: An adaptation of the affect misattribution procedure. *Cognition and Emotion*, 24(3), 522-535.
- Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2020). "A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures." *Personality and Social Psychology Bulletin*, 31(10), 1369-1385.
- Huber, G. A., & Lapinski, J. S. (2006). The "race card" revisited: Assessing racial priming in policy contests. *American Journal of Political Science*, 50(2), 421-440.
- Hui, C. H. (1988). Measurement of individualism-collectivism. *Journal of Research in Personality*, 22(1), 17-36.
- Hurwitz, J., & Peffley, M. (1997). Public perceptions of race and crime: The role of racial stereotypes. *American Journal of Political Science*, 375-401.
- Hurwitz, J., & Peffley, M. (2005). Playing the race card in the post-Willie Horton Era the impact of racialized code words on support for punitive crime policy. *Public Opinion Quarterly*, 69(1), 99-112.
- Hutchings, V. L., & Valentino, N. A. (2004). "The centrality of race in American politics." *Annual Review of Political Science*, 7, 383-408.
- Imhoff, R., Schmidt, A. F., Bernhardt, J., Dierksmeier, A., & Banse, R. (2011). An inkblot for sexual preference: A semantic variant of the Affect Misattribution Procedure. *Cognition and Emotion*, 25(4), 676-690.
- Iyengar, S., & Kinder, D. R. (2010). *News that matters: Television and American opinion*. Chicago, IL: University of Chicago Press.
- Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimisation for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, 17, 825-841.
- Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W., & Smith, S. M. (2012). FSL. *NeuroImage*, 62, 782-790.
- Jennings, M. K., Stoker, L., & Bowers, J. (2009). Politics across generations: Family transmission reexamined. *The Journal of Politics*, 71(3), 782-799.
- Jonas, K., Diehl, M., & Brömer, P. (1997). "Effects of attitudinal ambivalence on information processing and attitude-intention consistency." *Journal of Experimental Social Psychology*, 33(2), 190-210.

- Jost, J. T., Banaji, M. R., & Nosek, B. A. (2004). A decade of system justification theory: Accumulated evidence of conscious and unconscious bolstering of the status quo. *Political Psychology*, 25(6), 881-919.
- Jost, J. T., & Burgess, D. (2000). Attitudinal ambivalence and the conflict between group and system justification motives in low status groups. *Personality and Social Psychology Bulletin*, 26(3), 293-305.
- Jost, J. T., Chaikalis-Petrtsis, V., Abrams, D., Sidanius, J., Van Der Toorn, J., & Bratt, C. (2012). Why men (and women) do and don't rebel: Effects of system justification on willingness to protest. *Personality and Social Psychology Bulletin*, 38(2), 197-208.
- Jost, J. T., Glaser, J., Kruglanski, A. W., and Sulloway, F. J. (2003). Political conservatism as motivated social cognition. *Psychological Bulletin*, 129(3), 339.
- Jost, J. T., & Kay, A. C. (2005). Exposure to benevolent sexism and complementary gender stereotypes: consequences for specific and diffuse forms of system justification. *Journal of Personality and Social Psychology*, 88(3), 498.
- Jost, J. T., Kay, A. C., & Thorisdottir, H. (2009). *Social and psychological bases of ideology and system justification*. Oxford University Press.
- Jost, J. T., Pelham, B. W., & Carvallo, M. R. (2002). Non-conscious forms of system justification: Implicit and behavioral preferences for higher status groups. *Journal of Experimental Social Psychology*, 38(6), 586-602.
- Jost, J. T., and Thompson, E. P. (2000). Group-based dominance and opposition to equality as independent predictors of self-esteem, ethnocentrism, and social policy attitudes among African Americans and European Americans. *Journal of Experimental Social Psychology*, 36(3), 209-232.
- Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, 39(4), 341.
- Kaplan, K. J. (1972). "On the ambivalence-indifference problem in attitude theory and measurement: A suggested modification of the semantic differential technique." *Psychological Bulletin*, 77(5), 361.
- Kay, A. C., & Jost, J. T. (2003). Complementary justice: effects of "poor but happy" and "poor but honest" stereotype exemplars on system justification and implicit activation of the justice motive. *Journal of Personality and Social Psychology*, 85(5), 823.
- Kinder, D. R., & D. Kam. (2009). *Us Against Them: The Ethnocentric Foundations of American Opinion*. Chicago, IL: University of Chicago Press.
- Kinder, D. R. & Mendelberg, T. (2000). Individualism reconsidered: Principles and prejudice in contemporary American opinion. In J. S. David O. Sears and L. Bobo (Eds.), *Racialized politics: The debate about racism in America* (pp. 44-74). Chicago, IL: University of Chicago Press.
- Kinder, D. R., & Sanders, L. M. (1996). *Divided by color: Racial politics and democratic ideals*. University of Chicago Press.
- Kinder, D. R. & Sears, D. O. (1981). Prejudice and politics: Symbolic racism versus racial threats to the good life. *Journal of Personality and Social Psychology*, 40, 414-431.
- Knowles, E. D., Lowery, B. S., & Schaumberg, R. L. (2010). "Racial prejudice predicts opposition to Obama and his health care reform plan." *Journal of Experimental Social Psychology*, 46(2), 420-423.

- Kugler, M. B., Cooper, J., & Nosek, B. A. (2010). Group-based dominance and opposition to equality correspond to different psychological motives. *Social Justice Research, 23*(2), 117-155.
- Kuklinski, J. H., Sniderman, P. M., Knight, K., Piazza, T., Tetlock, P. E., Lawrence, G. R., & Mellers, B. (1997). Racial prejudice and attitudes toward affirmative action. *American Journal of Political Science, 40*, 402-419.
- Lavine, H. (2001). "The electoral consequences of ambivalence toward presidential candidates." *American Journal of Political Science, 9*, 915-929.
- Levine, D. K., & Palfrey, T. R. (2007). The paradox of voter participation? A laboratory study. *American Political Science Review, 101*(01), 143-158.
- Lippmann, W. (1946). *Public opinion* (Vol. 1). New Brunswick, NJ: Transaction Publishers.
- Lodge, M., & Taber, C. S. (2005). The automaticity of affect for political leaders, groups, and issues: An experimental test of the hot cognition hypothesis. *Political Psychology, 26*(3), 455-482.
- Lodge, M., & Taber, C. S. (2013). *The rationalizing voter*. New York: Cambridge University Press.
- Lowery, B. S., Unzueta, M. M., Knowles, E. D., & Goff, P. A. (2006). Concern for the in-group and opposition to affirmative action. *Journal of Personality and Social Psychology, 90*(6), 961.
- Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods, 47*(4), 1122-1135.
- Mackie, D. M., Devos, T., & Smith, E. R. (2000). Intergroup emotions: explaining offensive action tendencies in an intergroup context. *Journal of personality and social psychology, 79*(4), 602.
- McClosky, H., & Zaller, J. (1984). *The American Ethos*. Cambridge, MA: Harvard University Press.
- McConahay, J. B. (1983). "Modern Racism and Modern Discrimination The Effects of Race, Racial Attitudes, and Context on Simulated Hiring Decisions." *Personality and Social Psychology Bulletin, 9*(4), 551-558.
- McConahay, J. B. (1986). Modern racism, ambivalence, and the Modern Racism Scale. In J. F. Dovidio and S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 91-125). San Diego, CA: Academic Press.
- Mendelberg, T. (2001). *The race card: Campaign strategy, implicit messages, and the norm of equality*. Princeton University Press.
- Miller, M. J., Woehr, D. J., & Hudspeth, N. (2002). The meaning and measurement of work ethic: Construction and initial validation of a multidimensional inventory. *Journal of Vocational Behavior, 60*(3), 451-489.
- Milgram, S. (1964). Group pressure and action against a person. *The Journal of Abnormal and Social Psychology, 69*(2), 137.
- Mondak, J. J. (2010). *Personality and the foundations of political behavior*. New York, NY: Cambridge University Press.
- Monteith, M. J., Burns, M. D., Rupp, D. E., & Mihalec-Adkins, B. P. (2016). Out of work and out of luck? Layoffs, system justification, and hiring decisions for people who have been laid off. *Social Psychological and Personality Science, 7*(1), 77-84.

- Muller, E., & Satterthwaite, M. A. (1977). The equivalence of strong positive association and strategy-proofness. *Journal of Economic Theory*, *14*(2), 412-418.
- Mutz, D. C., & Reeves, B. (2005). The new videomalaise: Effects of televised incivility on political trust. *American Political Science Review*, *99*(01), 1-15.
- Neiman, J., Giuseffi, K., Smith, K., French, J., Waismel-Manor, I., & Hibbing, J. (2015). Voting at Home Is Associated with Lower Cortisol than Voting at the Polls. *PloS One*, *10*(9), e0135289.
- Nosek, B. A. (2005). "Moderators of the relationship between implicit and explicit evaluation." *Journal of Experimental Psychology: General*, *134*(4), 565.
- Nosek, B. A. (2007). "Implicit–explicit relations." *Current Directions in Psychological Science*, *16*(2), 65-69.
- Olson, M. A., & Fazio, R. H. (2009). Implicit and explicit measures of attitudes: The perspective of the MODE model. In Richard E. Petty, Russell H. Fazio, and Pablo Briñol (Eds.), *Attitudes: Insights from the new implicit measures* (pp. 119-164). New York, NY: Psychology Press.
- Oswald, F. L., Mitchell, G., Blanton, B., Jaccard, J., & Tetlock, P. E. (2013). "Predicting ethnic and racial discrimination: A meta-analysis of IAT criterion studies." *Journal of Personality and Social Psychology*, *105*(2), 171.
- Oxley, D. R., Smith, K. B., Alford, J. R., Hibbing, M. V., Miller, J. L., Scalora, M., Hatemi, P. K., & Hibbing, J. R. (2008). Political attitudes vary with physiological traits. *Science*, *321*(5896), 1667-1670.
- Panagopoulos, C. (2010). Affect, social pressure and prosocial motivation: Field experimental evidence of the mobilizing effects of pride, shame and publicizing voting behavior. *Political Behavior*, *32*(3), 369-386.
- Pastor-Bernier, A., Plott, C. R., & Schultz, W. (2017). Monkeys choose as if maximizing utility compatible with basic principles of revealed preference theory. *Proceedings of the National Academy of Sciences*, 201612010.
- Payne, B. K., Burkley, M. A., & Stokes, M. B. (2008). "Why do implicit and explicit attitude tests diverge? The role of structural fit." *Journal of Personality and Social Psychology*, *94*(1), 16.
- Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology*, *89*, 277-293.
- Payne, B. K., Krosnick, J. A., Pasek, J., Leikes, Y., Akhtar, O., & Tompson, T. (2010). Implicit and explicit prejudice in the 2008 American Presidential Election. *Journal of Experimental Social Psychology*, *46*, 367-374.
- Peffley, M., Hurwitz, J., & Sniderman, P. M. (1997). Racial stereotypes and whites' political views of blacks in the context of welfare and crime. *American Journal of Political Science*, *41*(1), 30-60.
- Peirce, J. W. (2007) PsychoPy - Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1-2), 8-13.
- Peirce, J. W. (2009) Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, *2*:10.

- Pérez, E. O. (2010). "Explicit evidence on the import of implicit attitudes: The IAT and immigration policy judgments." *Political Behavior*, 32(4), 517-545.
- Pérez, E. O. (2016). *Unspoken Politics: Implicit Attitudes and Political Thinking*. New York: Cambridge University Press.
- Petty, R. E., Briñol, P., & Johnson, I. (2012). Implicit ambivalence. In Bertram Gawronski, & Fritz Strack (Eds.), *Cognitive consistency: A fundamental principle in social cognition*, pp. 178-201. New York: Guilford Press.
- Petty, R. E., Fazio, R. H., & Briñol, P. (Eds.). (2012). *Attitudes: Insights from the new implicit measures*. New York: Psychology Press.
- Petty, R. E., Tormala, Z. L., Briñol, P., & Jarvis, W. B. G. (2006). "Implicit ambivalence from attitude change: an exploration of the PAST model." *Journal of Personality and Social Psychology* 90(1), 21.
- Pinker, S. (2008). A biological understanding of human nature. In Brockman, J. (Ed.), *Science at the Edge: Conversations with Leading Scientific Thinkers of Today* (pp. 37-56). New York, NY: Sterling Publishing Company.
- Plant, E. A., & Devine, P. G. (1998). "Internal and external motivation to respond without prejudice." *Journal of Personality and Social Psychology*, 75(3), 811.
- Pratto, F., Sidanius, J., & Levin, S. (2006) Social dominance theory and the dynamics of intergroup relations: Taking stock and looking forward, *European Review of Social Psychology*, 17(1), 271-320.
- Pratto, F., Sidanius, J., Stallworth, L. M., & Malle, B. F. (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology*, 67, 741-763.
- Ratner, K. G., Kaul, C., & Van Bavel, J. J. (2013). Is race erased? Decoding race from patterns of neural activity when skin color is not diagnostic of group boundaries. *Social Cognitive and Affective Neuroscience*, 8(7), 750-755.
- Redlawsk, D. P. (2002). Hot cognition or cool consideration? Testing the effects of motivated reasoning on political decision making. *Journal of Politics*, 64(4), 1021-1044.
- Reyna, C., Henry, P. J., Korfmacher, W., & Tucker, A. (2005). Examining the Principles in Principled Conservatism: The Role of Responsibility Stereotypes as Cues for Deservingness in Racial Policy Decisions. *Journal of Personality and Social Psychology*, 90(1), 109-128.
- Rhemtulla, M., Brosseau-Liard, P. É., & Savalei, V. (2012). When Can Categorical Variables Be Treated as Continuous? A Comparison of Robust Continuous and Categorical SEM Estimation Methods Under Suboptimal Conditions. *Psychological Methods*, 17(3), 354-373.
- Richeson, J. A., Baird, A. A., Gordon, H. L., Heatherton, T. F., Wyland, C. L., Trawalter, S., & Shelton, J. N. (2003). An fMRI investigation of the impact of interracial contact on executive function. *Nature Neuroscience*, 6(12), 1323-1328.
- Rilling, J. K., Dagenais, J. E., Goldsmith, D. R., Glenn, A. L., & Pagnoni, G. (2008). Social cognitive neural networks during in-group and out-group interactions. *Neuroimage*, 41(4), 1447-1461.

- Roccas, S., Sagiv, L., Schwartz, S., Halevy, N., & Eidelson, R. (2008). Toward a unifying model of identification with groups: Integrating theoretical perspectives. *Personality and Social Psychology Review*, 12(3), 280-306.
- Rohr, M., Degner, J., & Wentura, D. (2015). The “emotion misattribution” procedure: Processing beyond good and bad under masked and unmasked presentation conditions. *Cognition and Emotion*, 29(2), 196-219.
- Rudolph, T. J. (2011). “The dynamics of ambivalence.” *American Journal of Political Science*, 55(3), 561-573.
- Rudolph, T. J., & Popp, E. (2007). “An information processing theory of ambivalence.” *Political Psychology*, (5), 563-585.
- Schreiber, D. (2011). From SCAN to neuropolitics. In P. Hatemi and McDermott, R. (Eds.), *Man is by Nature a Political Animal: Evolution, Biology and Politics* (pp. 273-299). Chicago and London: The University of Chicago Press.
- Sears, D. O. & Henry, P. J. (2003). The origins of symbolic racism. *Journal of Personality and Social Psychology*, 85, 259-275.
- Sears, D. O., Van Laar, C., Carrillo, M., & Kosterman, R. (1997). Is it really racism? The origins of white Americans' opposition to race-targeted policies. *The Public Opinion Quarterly*, 61, 16-53.
- Shkurko, A. V. (2013). Is social categorization based on relational ingroup/outgroup opposition? A meta-analysis. *Social Cognitive and Affective Neuroscience*, 8(8), 870-877.
- Sidanius, J., & Pratto, F. (1993). “The inevitability of oppression and the dynamics of social dominance.” In P. Sniderman, P. E. Tetlock, and E. G. Carmines (Eds.), *Prejudice, politics, and the American dilemma* (pp. 173-211). Stanford, CA: Stanford University Press.
- Sidanius, J., & Pratto, F. (1999). *Social dominance: An intergroup theory of social hierarchy and oppression*. New York: Cambridge University Press.
- Sidanius, J., Pratto, F., & Bobo, L. (1996). Racism, conservatism, affirmative action, and intellectual sophistication: A matter of principled conservatism or group dominance? *Journal of Personality and Social Psychology*, 70, 476-490.
- Simon, H. A. (1946). The proverbs of administration. *Public Administration Review*, 6(1), 53-67.
- Simon, H. A. (1982). *Models of bounded rationality: Empirically grounded economic reason* (Vol. 3). Cambridge, MA: MIT press.
- Smith, K. B., Oxley, D. R., Hibbing, M. V., Alford, J. R., & Hibbing, J. R. (2011). Linking genetics and political attitudes: Reconceptualizing political ideology. *Political Psychology*, 32(3), 369-397.
- Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, 17, 143-155.
- Smith, S. M., Jenkinson, M., Woolrich, M. w., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobnjak, I, Flitney, D. E., Niazy, R. K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, M., & Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, 23, S208-S219.
- Sniderman, P. M. & Carmines, E.G. (1997). *Reaching Beyond Race*. Cambridge, MA.: Harvard University Press.

- Sniderman, P. M., Crosby, G. C., & Howell, W. G. (2000). The politics of race. In D. O. Sears, J. Sidanius, & L. Bobo (Eds.), *Racialized politics: The debate about racism in America* (pp. 236-279). Chicago, IL: University of Chicago Press.
- Sniderman, P. M., & Piazza, T. (1993). *The Scar of Race*. Cambridge, MA: Harvard University Press.
- Sniderman, P. M., Piazza, T., Tetlock, P. E., & Kendrick, A. (1991). The new racism. *American Journal of Political Science*, 423-447.
- Sniderman, P. M., Carmines, E. G., Laymen, G. C., & Carter, M. (1996). Beyond race: Social justice as a race neutral ideal. *American Journal of Political Science*, 40, 33-55.
- Sniderman, P. M. & Tetlock, P. E. (1986). Symbolic racism: Problems of motive attribution in political analysis. *Journal of Social Issues*, 42, 129-150.
- Sobel M. E. (1988). Direct and indirect effects in linear structural equation models. In J. S. Long (Ed.), *Common Problems/Proper Solutions* (pp. 46–64). Beverly Hills, CA: Sage.
- Stanley, D., Phelps, E., & Banaji, M. R. (2008). “The neural basis of implicit attitudes.” *Current Directions in Psychological Science*, 17(2), 164-170.
- Sumner, W. G. (1906). *Folkways: A study of the sociological importance of usages, manners, customs, mores, and morals*. Boston, MA: The Athenaeum Press.
- Sunstein, C. R. (2013). *Simpler: The future of government*. New York, NY: Simon and Schuster.
- Taber, C. S., & Lodge, M. (2006). “Motivated skepticism in the evaluation of political beliefs.” *American Journal of Political Science*, 50(3), 755-769.
- Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behavior. *European Journal of Social Psychology*, 1(2), 149-178.
- Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. *The Social Psychology of Intergroup Relations*, 33, 47.
- Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving Decisions about Health, Wealth, and Happiness*. New York, NY: Penguin Books.
- Thompson, M. M., & Zanna, M. P. (1995). “The Conflicted Individual: Personality-Based and Domain Specific Antecedents of Ambivalent Social Attitudes.” *Journal of Personality*, 63(3), 259-288.
- Thompson, M. M., & Zanna, M. P., & Griffin, D. W. (1995). “Let’s not be indifferent about (attitudinal) ambivalence.” In R. E. Petty & J. A. Krosnick (Eds.) *Attitude strength: Antecedents and consequences* (pp. 361-386). New York: Lawrence Erlbaum Associates.
- Triandis, H. C., Bontempo, R., Villareal, M. J., Asai, M., & Lucca, N. (1988). Individualism and collectivism: Cross-cultural perspectives on self-ingroup relationships. *Journal of Personality and Social Psychology*, 54(2), 323.
- Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Oxford: Blackwell.
- Valentino, N. A., Hutchings, V. L., & White, I. K. (2002). Cues that matter: How political ads prime racial attitudes during campaigns. *American Political Science Review*, 96(01), 75-90.
- Van Bavel, J. J., Packer, D. J., & Cunningham, W. A. (2008). The neural substrates of in-group bias a functional magnetic resonance imaging investigation. *Psychological Science*, 19(11), 1131-1139.

- Van Bavel, J. J., Packer, D. J., & Cunningham, W. A. (2011). Modulation of the fusiform face area following minimal exposure to motivationally relevant faces: evidence of in-group enhancement (not out-group disregard). *Journal of Cognitive Neuroscience*, *23*(11), 3343-3354.
- Volz, K. G., Kessler, T., & von Cramon, D. Y. (2009). In-group as part of the self: in-group favoritism is mediated by medial prefrontal cortex activation. *Social Neuroscience*, *4*(3), 244-260.
- Weber, M. (1905). *The Protestant ethic and the spirit of capitalism*. New York: Scribners.
- Wheeler, M. E., & Fiske, S. T. (2005). Controlling racial prejudice social-cognitive goals affect amygdala and stereotype activation. *Psychological Science*, *16*(1), 56-63.
- Whittaker, E., & Kowalski, R. M. (2015). Cyberbullying via social media. *Journal of School Violence*, *14*(1), 11-29.
- Wilson, D. C., Moore, D. W., McKay, P. F., & Avery, D. R. (2008). Affirmative Action Programs for Women and Minorities Expressed Support Affected by Question Order. *Public Opinion Quarterly*, *72*, 514-522.

APPENDIX

Chapter 2

Group Identification:

1. I feel strongly affiliated with most groups I belong to.
2. Other groups can learn a lot from the groups I belong to.
3. Belonging to groups is an important part of my identity.
4. In times of trouble, the only way to know what to do is to rely on the group leaders.
5. I am glad to contribute to the groups I belong to.
6. Compared to other groups of the same kind, the groups I belong to are particularly good.
7. It is important to me that I view myself as a member of groups I belong to.
8. All group members should respect the customs, the institutions, and the leaders of the group.
9. I am strongly committed to the groups I belong to.
10. Relative to other groups, the groups I belong to are very moral groups.
11. It is important to me that others see me as a member of groups I belong to.
12. It is disloyal to criticize the group.
13. I like to help the groups I belong to.
14. The groups I belong to are better than other groups in all respects.
15. When I talk about the group members, I usually say “we” instead of “they.”
16. There is usually a good reason for every rule and regulation that the group leaders propose.

1 (Strongly Disagree)

2 (Disagree)

3 (Neither Agree nor Disagree)

4 (Agree)

5 (Strongly Agree)

Importance – 3, 7, 11, 15; Commitment – 1, 5, 9, 13; Deference – 4, 8, 12, 16; Superiority – 2, 6, 10, 14.

Error covariance estimated between items 5 and 13.

SDO:

1. Some groups of people are just more worthy than others.
2. It would be good if all groups could be equal.
3. In getting what your group wants, it is sometimes necessary to use force against other groups.
4. Group equality should be our ideal.
5. Superior groups should dominate inferior groups.
6. All groups should be given an equal chance in life.
7. To get ahead in life, it is sometimes necessary to step on other groups.
8. We should do what we can to equalize conditions for different groups.

9. If certain groups of people stayed in their place, we would have fewer problems.
10. Increased social equality.
11. It's probably a good thing that certain groups are at the top and other groups are at the bottom.
12. We would have fewer problems if we treated different groups more equally.
13. Inferior groups should stay in their place.
14. We should strive to make incomes more equal.
15. Sometimes other groups must be kept in their place.
16. No one group should dominate in society.

1 (Strongly Disagree)

2 (Disagree)

3 (Neither Agree nor Disagree)

4 (Agree)

5 (Strongly Agree)

SDO-D – 1, 3, 5, 7, 9, 11, 13, 15; SDO-E – 2, 4, 6, 8, 10, 12, 14, 16.

Error covariances estimated between items 3 and 7, items 2 and 4, and items 8 and 14.

Individualism:

1. To be truly successful, a person should be self-reliant.
2. Nothing is impossible if you work hard enough.
3. Self-reliance is the key to being successful.
4. If one works hard enough, one is likely to make a good life for oneself.
5. Hard work makes one a better person.
6. People would be better off if they depended on themselves.
7. One should live one's own life independent of others as much as possible.
8. I do not like having to depend on other people.
9. Working hard is the key to being successful.
10. By working hard a person can overcome every obstacle that life presents.
11. One must avoid dependence on other persons whenever possible.
12. Any problem can be overcome with hard work.
13. I strive to be self-reliant.
14. If you work hard you will succeed.
15. Anyone who is able and willing to work hard has a good chance of succeeding.
16. Having a great deal of independence from others is very important to me.
17. A person should always do the best job possible.
18. Only those who depend on themselves get ahead in life.
19. It is important to control one's destiny by not being dependent on others.
20. By simply working hard enough, one can achieve one's goals.

1 (Strongly Disagree)

2 (Disagree)

3 (Neither Agree nor Disagree)

4 (Agree)

5 (Strongly Agree)

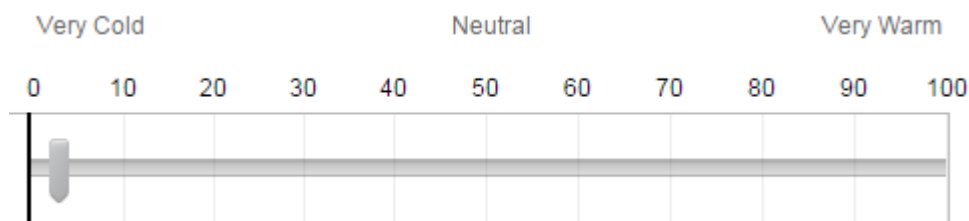
Work Ethic – 2, 4, 5, 9, 10, 12, 14, 15, 17, 20; Self-Reliance – 1, 3, 6, 7, 8, 11, 13, 16, 18, 19.
Error covariances estimated between items 8 and 16, items 7 and 13, items 11 and 13, items 10 and 12, and items 2 and 12.

Chapter 3

Affective Evaluations:

**Asked with regard to the minimal ingroup and minimal outgroup.*

1. Please rate how "warm" or "cold" you feel toward each of the following groups on the below scales from 0-100. Ratings between 50 degrees and 100 degrees mean that you feel favorable and warm toward the group. Ratings between 0 degrees and 50 degrees mean that you don't feel favorable toward the group and that you don't care too much for that group. You would rate the group at the 50 degree mark if you don't feel particularly warm or cold toward the group.



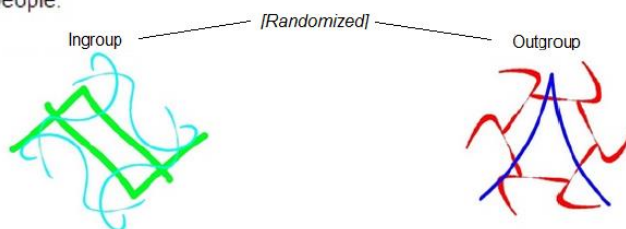
Attitudes toward race-targeted policies:

1. Affirmative action for Blacks in hiring for jobs.
2. Increased federal education spending in neighborhoods predominantly populated by Blacks.
3. Federally funded college scholarships for Blacks.
4. Policies that make it illegal for law enforcement to profile Blacks (use the fact that an individual is Black to decide whether or not to engage in enforcement).
 - 1 (Strongly Oppose)
 - 2 (Moderately Oppose)
 - 3 (Slightly Oppose)
 - 4 (Neither Oppose nor Support)
 - 5 (Slightly Support)
 - 6 (Moderately Support)
 - 7 (Strongly Support)

Minimal Groups Logos and Description

Research has shown that people who more accurately estimate the number of dots tend to have a particular set of psychological traits that are significantly different from people who less accurately estimate the number of dots. We call these different "numerical estimation styles." Specifically, some research has shown that various traits are correlated with numerical estimation style. These traits that are correlated with numerical estimation style include personality, problem-solving approaches, what aspects of the world individuals consider beautiful, and even preferences for mate choice.

These findings have been shown in numerous studies (e.g., Feldhouse, Grimm, and Banks, 2004; Maxwell and Tunjii, 1998; Nill, 2000), and have even been backed up by neurological examinations of brain activity (Bellows, Coleman, and Kale, 2010). It seems, from the findings, that people's numerical estimation style can actually predict stable traits in people.



Minimal Groups Resource Allocation Task

**For detailed instructions used, contact author.*

As a reminder, on each matrix you are to award points (which are valuable in this experiment) to two other participants in this experiment. The top row of numbers within the boxes are the points awarded to a fellow member of your group, the [Ingroup], and the bottom row are points to be given to a member of the other group, the [Outgroup]. After looking at each box of the matrix, you must choose only one box that represents your choice of how you wish to award the points.

Remember that you are distributing resources between members of your group and the other group, but no yourself.

Points to member of the [Ingroup]	19	18	17	16	15	14	13	12	11	10	9	8	7
Points to member of the [Outgroup]	1	3	5	7	9	11	13	15	17	19	21	23	25

Points to member of the [Ingroup]	25	23	21	19	17	15	13	11	9	7	5	3	1
Points to member of the [Outgroup]	1	8	9	10	11	12	13	14	15	16	17	18	19

Points to member of the [Ingroup]	19	18	17	16	15	14	13	12	11	10	9	8	7
Points to member of the [Outgroup]	25	23	21	19	17	15	13	11	9	7	5	3	1

Points to member of the [Ingroup]	1	3	5	7	9	11	13	15	17	19	21	23	25
Points to member of the [Outgroup]	7	8	9	10	11	12	13	14	15	16	17	18	19

Points to member of the [Ingroup]	16	17	18	19	20	21	22	23	24	25	26	27	28
Points to member of the [Outgroup]	16	15	14	13	12	11	10	9	8	7	6	5	4

Points to member of the [Ingroup]	4	5	6	7	8	9	10	11	12	13	14	15	16
Points to member of the [Outgroup]	28	27	26	25	24	23	22	21	20	19	18	17	16

Chapter 4

Symbolic Racism

1. Irish, Italians, Jewish and many other minorities overcame prejudice and worked their way up. Blacks should do the same without any special favors.
2. Generations of slavery and discrimination have created conditions that make it difficult for blacks to work their way out of the lower class.
3. Over the past few years, blacks have gotten less than they deserve.
4. It's really a matter of some people not trying hard enough; if blacks would only try harder they could be just as well off as whites.

- 1 (Strongly Disagree)
- 2 (Disagree)
- 3 (Neither Agree nor Disagree)
- 4 (Agree)
- 5 (Strongly Agree)

Influence of Blacks in U.S. Politics

1. Would you say that blacks have TOO MUCH INFLUENCE in American politics, JUST ABOUT THE RIGHT AMOUNT of influence in American politics, or TOO LITTLE influence in American politics?

Black Elected Officials

1. Do you think that most white candidates who run for political office are better suited to be an elected official than are most black candidates, that most black candidates are better suited to be an elected official than are most white candidates, or do you think white and black candidates are equally suited to be an elected official?

Respondents then asked how strongly they hold that position.

Black Elected Officials - Intelligence

1. Do you think that most white candidates who run for political office are better suited in terms of their intelligence to serve as an elected official than are most black candidates, that most black candidates are better suited in terms of their intelligence to serve as an elected official than are most white candidates, or do you think white and black candidates are equally suited in terms of their intelligence to serve as an elected official?

Respondents then asked how strongly they hold that position.

Affirmative Action

1. Some people say that because of past discrimination, Blacks should be given preference in hiring and promotion. Others say that such preference in hiring and promotion of Blacks is wrong because it gives Blacks advantages they haven't earned. What about your opinion - are you for or against preferential hiring and promotion of Blacks?

Fair Job Treatment

1. Some people feel that if black people are not getting fair treatment in jobs, the government in Washington ought to see to it that they do. Others feel that this is not the federal government's business. Have you had enough interest in this question to favor one side over the other?

How do you feel? Should the government in Washington see to it that black people get fair treatment in jobs OR is this not the federal government's business?

Respondents then asked how strongly they hold that position.

Vote Obama

1. How about the election for President? Did you vote for a candidate for PRESIDENT? Who did you vote for?

Feeling Thermometers

1. Ratings between 50 degrees and 100 degrees mean that you feel favorable and warm toward the person. Ratings between 0 degrees and 50 degrees mean that you don't feel favorable toward the person and that you don't care too much for that person. You would rate the person at the 50 degree mark if you don't feel particularly warm or cold toward the person.

Need for Cognition

1. Some people like to have responsibility for handling situations that require a lot of thinking, and other people don't like to have responsibility for situations like that. What about you? Do you like having responsibility for handling situations that require a lot of thinking, do you dislike it, or do you neither like it nor dislike it? Do you like it a lot or just somewhat?
2. Some people prefer to solve simple problems instead of complex ones, whereas other people prefer to solve more complex problems. Which type of problem do you prefer to solve: simple or complex?

Education

1. What is the highest grade of school or year of college you have completed?
2. Did you get a high school diploma or pass a high school equivalency test?
3. What is the highest degree that you have earned?

Political Interest

The political interest and political discussion variables were part of an experimental manipulation in the 2008 ANES, in which respondents were randomly assigned to receive either old or new versions of questions. For political interest, the old version asked only about how often respondents follow government and public affairs whereas the new version asked one item about how closely respondents follow government and politics and a second item about how often. Thus, for respondents in the new condition, political interest is measured as the mean of the two items. The old version for political discussion first asked whether or not respondents discuss politics with family and friends and then asked how often they did in the past week, whereas the new version just asked how often in a typical week.

“old”

1. Some people seem to follow what's going on in government and public affairs most of the time, whether there's an election going on or not. Others aren't that interested. Would you say you follow what's going on in government and public affairs MOST OF THE TIME, SOME OF THE TIME, ONLY NOW AND THEN, or HARDLY AT ALL?

“new”

1. How closely do you pay attention to information about what's going on in government and politics? EXTREMELY CLOSELY, VERY CLOSELY, MODERATELY CLOSELY, SLIGHTLY CLOSELY, or NOT CLOSELY AT ALL?
2. How often do you pay attention to what's going on in government and politics? ALL THE TIME, MOST OF THE TIME, ABOUT HALF THE TIME, ONCE IN A WHILE, or NEVER?

Ideology

Where would you place YOURSELF on this scale, or haven't you thought much about this?
Extremely liberal, Liberal, Slightly liberal, Moderate; middle of the road, Slightly conservative, Conservative, Extremely conservative

Egalitarianism

1. Our society should do whatever is necessary to make sure that everyone has an equal opportunity to succeed.

2. We have gone too far in pushing equal rights in this country.
3. One of the big problems in this country is that we don't give everyone an equal chance.
4. This country would be better off if we worried less about how equal people are.
5. It is not really that big a problem if some people have more of a chance in life than others.
6. If people were treated more equally in this country we would have many fewer problems.

Table A4.1: Full Model Predicting Symbolic Racism

	Main Effects	Direction Int.	Ideology Int.	Egalitarianism Int.
Implicit Ambivalence	0.163* (0.064)	0.247** (0.090)	0.277** (0.090)	0.246** (0.090)
Ambivalence Direction	0.091‡ (0.052)	0.092‡ (0.052)	0.086 (0.054)	0.093‡ (0.052)
Ideology	0.121*** (0.020)	0.121*** (0.020)	0.114*** (0.026)	0.120*** (0.020)
Egalitarianism	-0.449*** (0.038)	-0.447*** (0.038)	-0.440*** (0.038)	-0.464*** (0.050)
NFC	0.029 (0.083)	0.027 (0.083)	0.026 (0.083)	0.028 (0.083)
Education	-0.094*** (0.018)	-0.095*** (0.018)	-0.092*** (0.018)	-0.096*** (0.018)
Political Interest	-0.032 (0.028)	-0.030 (0.028)	-0.036 (0.028)	-0.029 (0.028)
Age	-0.0003 (0.002)	-0.0004 (0.002)	-0.00001 (0.002)	-0.001 (0.002)
Male	-0.046 (0.054)	-0.049 (0.054)	-0.054 (0.054)	-0.055 (0.054)
Imp Amb:Amb Direction	-	-0.168 (0.126)	-0.211 (0.130)	-0.171 (0.126)
Imp Amb*Ideology	-	-	-0.157** (0.058)	-
Imp Amb*Egalitarianism	-	-	-	0.045 (0.107)
Ideology*Amb Direction	-	-	0.016 (0.036)	-
Egalitarianism*Amb Direction	-	-	0.037 (0.068)	-
Imp Amb*Ideology*Amb Direction	-	-	0.178* (0.088)	-
Imp Amb*Egalitarianism*Amb Direction	-	-	-	-0.304‡ (0.165)
Constant	4.902*** (0.230)	5.033*** (0.224)	5.485*** (0.184)	3.532*** (0.148)
N	970	970	970	970
R ²	0.279	0.280	0.286	0.284

Coefficients are linear regression coefficients with standard errors in parentheses. Thresholds shown with standard errors in parentheses. ***p < .001; **p < .01; *p < .05; ‡p < .10.

Table A4.2: Full Model Predicting Perceived Influence of Blacks in Politics

	Main Effects	Direction Int.	Ideology Int.	Egalitarianism Int.
Implicit Ambivalence	0.919 (0.699,1.208)	0.612* (0.419,0.895)	0.622* (0.425,0.910)	0.920 (0.699,1.209)
Ambivalence Direction	0.949 (0.762,1.183)	0.945 (0.758,1.177)	1.020 (0.812,1.283)	0.948 (0.761,1.182)
Ideology	0.774*** (0.711,0.843)	0.770*** (0.707,0.838)	0.842** (0.754,0.940)	0.774*** (0.711,0.842)
Egalitarianism	1.993*** (1.688,2.357)	1.985*** (1.681,2.348)	2.014*** (1.703,2.384)	1.995*** (1.689,2.360)
NFC	1.230 (0.869,1.741)	1.258 (0.889,1.782)	1.258 (0.887,1.784)	1.227 (0.867,1.736)
Education	1.000 (0.929,1.077)	1.005 (0.933,1.082)	1.006 (0.934,1.083)	1.001 (0.930,1.078)
Political Interest	0.942 (0.836,1.062)	0.936 (0.830,1.055)	0.932 (0.826,1.051)	0.943 (0.836,1.062)
Age	0.999 (0.992,1.006)	0.999 (0.992,1.006)	0.999 (0.992,1.006)	0.999 (0.992,1.006)
Male	0.855 (0.681,1.074)	0.869 (0.691,1.091)	0.847 (0.673,1.065)	0.857 (0.682,1.077)
Imp Amb*Amb Direction	-	2.274** (1.331,3.892)	-	1.966* (1.126,3.436)
Imp Amb*Ideology	-	-	0.896 (0.699,1.151)	-
Ideology*Amb Direction	-	-	0.815** (0.698,0.949)	-
Imp Amb*Egalitarianism	-	-	-	1.075 (0.764,1.520)
Imp Amb*Ideology*Amb Direction	-	-	1.481* (1.012,2.169)	-
Too Much About Right	-1.266* (0.496)	-1.195* (0.484)	-0.121 (0.403)	-3.533*** (0.336)
About Right Not Enough	1.858*** (0.496)	1.948*** (0.485)	3.047*** (0.411)	-0.409 (0.317)
AIC	2297.435	2290.407	2285.624	2299.264
N	1319	1319	1319	1319

Coefficients are odds ratios with 95% confidence intervals in parentheses. Thresholds shown with standard errors in parentheses. ***p < .001; **p < .01; *p < .05; †p < .10.

Table A4.3: Full Model Predicting Attitudes toward Black Elected Officials

	Main Effects	Direction Int.	Ideology Int.	Egalitarianism Int.
Implicit Ambivalence	-0.290*** (0.040)	-0.535*** (0.055)	-0.274*** (0.041)	-0.290*** (0.040)
Ambivalence Direction	0.094** (0.032)	0.090** (0.032)	0.095** (0.032)	0.092** (0.032)
Ideology	-0.012 (0.013)	-0.015 (0.012)	-0.011 (0.013)	-0.013 (0.013)
Egalitarianism	0.057* (0.024)	0.051* (0.024)	0.060* (0.024)	0.058* (0.024)
NFC	0.025 (0.052)	0.031 (0.051)	0.029 (0.052)	0.020 (0.052)
Education	0.002 (0.011)	0.005 (0.011)	0.003 (0.011)	0.003 (0.011)
Political Interest	0.016 (0.018)	0.010 (0.017)	0.014 (0.018)	0.016 (0.018)
Age	0.0002 (0.001)	0.001 (0.001)	0.0003 (0.001)	0.0003 (0.001)
Male	0.001 (0.034)	0.012 (0.033)	0.002 (0.034)	0.005 (0.034)
Imp Amb:Amb Direction	-	0.491*** (0.077)	-	-
Imp Amb*Ideology	-	-	-0.046 (0.027)	-
Imp Amb*Egalitarianism	-	-	-	0.102* (0.051)
Constant	3.868*** (0.145)	3.633*** (0.139)	3.564*** (0.116)	3.823*** (0.093)
N	959	959	959	959
R ²	0.073	0.111	0.076	0.077

Coefficients are linear regression coefficients with standard errors in parentheses. Thresholds shown with standard errors in parentheses. ***p < .001; **p < .01; *p < .05; †p < .10.

Table A4.4: Full Model Attitudes toward Black Elected Officials – Intelligence

	Main Effects	Direction Int.	Ideology Int.	Egalitarianism Int.
Implicit Ambivalence	-0.223*** (0.041)	-0.485*** (0.057)	-0.456*** (0.057)	-0.223*** (0.041)
Ambivalence Direction	0.123*** (0.033)	0.118*** (0.033)	0.107** (0.034)	0.122*** (0.033)
Ideology	-0.021 (0.013)	-0.023‡ (0.013)	-0.035* (0.016)	-0.022‡ (0.013)
Egalitarianism	0.057* (0.025)	0.053* (0.024)	0.058* (0.024)	0.058* (0.025)
NFC	0.016 (0.053)	0.021 (0.052)	0.019 (0.052)	0.013 (0.053)
Education	0.008 (0.011)	0.011 (0.011)	0.013 (0.011)	0.008 (0.011)
Political Interest	0.022 (0.018)	0.015 (0.018)	0.010 (0.018)	0.022 (0.018)
Age	-0.0003 (0.001)	0.0001 (0.001)	0.0004 (0.001)	-0.0002 (0.001)
Male	0.040 (0.034)	0.053 (0.034)	0.049 (0.034)	0.042 (0.035)
Imp Amb:Amb Direction	-	0.524*** (0.079)	0.489*** (0.082)	-
Imp Amb*Ideology	-	-	-0.138*** (0.037)	-
Ideology*Amb Direction	-	-	0.025 (0.022)	-
Imp Amb*Egalitarianism	-	-	-	0.062 (0.055)
Imp Amb*Ideology*Amb Direction	-	-	0.148** (0.055)	-
Constant	3.801*** (0.149)	3.616*** (0.142)	3.499*** (0.116)	3.813*** (0.096)
N	960	960	960	960
R ²	0.062	0.104	0.118	0.064

Coefficients are linear regression coefficients with standard errors in parentheses. Thresholds shown with standard errors in parentheses. ***p < .001; **p < .01; *p < .05; ‡p < .10.

Table A4.5: Full Model Predicting Support for Affirmative Action

	Main Effects	Direction Int.	Ideology Int.	Egalitarianism Int.
Implicit Ambivalence	0.947 (0.703,1.272)	0.925 (0.615,1.386)	0.915 (0.679,1.231)	0.984 (0.728,1.326)
Ambivalence Direction	0.915 (0.719,1.164)	0.915 (0.719,1.164)	0.916 (0.719,1.166)	0.924 (0.726,1.176)
Ideology	0.746*** (0.681,0.817)	0.746*** (0.681,0.817)	0.739*** (0.674,0.810)	0.749*** (0.683,0.820)
Egalitarianism	1.860*** (1.545,2.248)	1.859*** (1.544,2.247)	1.839*** (1.524,2.227)	1.871*** (1.552,2.263)
NFC	0.649* (0.440,0.956)	0.650* (0.441,0.957)	0.627* (0.424,0.925)	0.657* (0.446,0.968)
Education	0.915* (0.841,0.994)	0.915* (0.841,0.995)	0.906* (0.832,0.985)	0.913* (0.839,0.992)
Political Interest	1.072 (0.939,1.226)	1.072 (0.938,1.225)	1.086 (0.949,1.242)	1.071 (0.937,1.224)
Age	0.991* (0.983,0.999)	0.991* (0.983,0.999)	0.990* (0.982,0.998)	0.990* (0.983,0.998)
Male	1.116 (0.869,1.433)	1.117 (0.870,1.435)	1.109 (0.863,1.425)	1.098 (0.855,1.412)
Imp Amb:Amb Direction	-	1.050 (0.588,1.875)	-	-
Imp Amb*Ideology	-	-	1.380** (1.126,1.693)	-
Imp Amb*Egalitarianism	-	-	-	0.736 (0.494,1.096)
Strongly Against Not Strongly Against	0.762 (0.550)	0.806 (0.530)	1.865*** (0.444)	-1.303*** (0.347)
Not Strongly Against Not Strongly For	2.211*** (0.554)	2.256*** (0.534)	3.324*** (0.453)	0.149 (0.348)
Not Strongly For Strongly For	3.082*** (0.562)	3.126*** (0.542)	4.200*** (0.464)	1.020** (0.358)
AIC	2240.069	2242.042	2232.393	2239.790
N	1344	1344	1344	1344

Coefficients are odds ratios with 95% confidence intervals in parentheses. Thresholds shown with standard errors in parentheses. ***p < .001; **p < .01; *p < .05; †p < .10.

Table A4.6: Full Model Predicting Support for Fair Job Treatment for Blacks

	Main Effects	Direction Int.	Ideology Int.	Egalitarianism Int.
Implicit Ambivalence	0.682* (0.475,0.978)	0.486** (0.294,0.799)	0.642* (0.442,0.932)	0.507** (0.304,0.836)
Ambivalence Direction	0.741* (0.551,0.995)	0.726* (0.540,0.977)	0.745‡ (0.554,1.001)	0.729* (0.542,0.980)
Ideology	0.918 (0.822,1.026)	0.920 (0.823,1.028)	0.916 (0.819,1.023)	0.923 (0.825,1.032)
Egalitarianism	3.595*** (2.867,4.547)	3.582*** (2.856,4.531)	3.566*** (2.843,4.513)	3.884*** (2.883,5.316)
NFC	0.836 (0.513,1.358)	0.837 (0.513,1.362)	0.827 (0.507,1.344)	0.853 (0.521,1.393)
Education	0.993 (0.898,1.099)	0.995 (0.900,1.101)	0.992 (0.897,1.097)	0.994 (0.898,1.100)
Political Interest	1.402*** (1.202,1.639)	1.390*** (1.191,1.625)	1.407*** (1.207,1.645)	1.392*** (1.193,1.628)
Age	0.999 (0.989,1.008)	0.999 (0.990,1.009)	0.998 (0.989,1.008)	0.999 (0.990,1.009)
Male	1.104 (0.815,1.498)	1.123 (0.828,1.525)	1.098 (0.810,1.490)	1.135 (0.837,1.542)
Imp Amb:Amb Direction	–	2.017‡ (0.991,4.109)	–	1.864‡ (0.909,3.836)
Imp Amb*Ideology	–	–	1.165 (0.905,1.503)	–
Egalitarianism*Amb Direction	–	–	–	0.836 (0.559,1.252)
Imp Amb*Egalitarianism	–	–	–	0.759 (0.418,1.411)
Imp Amb*Egalitarianism*Amb Direction	–	–	–	2.484‡ (0.928,6.666)
Strongly Against Not Strongly Against	3.901*** (0.695)	4.214*** (0.687)	4.499*** (0.582)	–0.079 (0.462)
Not Strongly Against Not Strongly For	4.529*** (0.700)	4.843*** (0.692)	5.128*** (0.588)	0.553 (0.463)
Not Strongly For Strongly For	5.085*** (0.706)	5.402*** (0.697)	5.684*** (0.596)	1.114* (0.465)
AIC	1599.754	1598.013	1600.352	1600.133
N	749	749	749	749

Coefficients are odds ratios with 95% confidence intervals in parentheses. Thresholds shown with standard errors in parentheses. ***p < .001; **p < .01; *p < .05; ‡p < .10.

Table A4.7: Full Model Predicting Self-Reported Vote for Obama

	Main Effects	Direction Int.	Ideology Int.	Egalitarianism Int.
Implicit Ambivalence	-0.534* (0.475,0.978)	-0.505 [‡] (0.294,0.799)	-0.652 (0.442,0.932)	-0.535** (0.304,0.836)
Ambivalence Direction	-0.094* (0.551,0.995)	-0.095 (0.540,0.977)	-0.095 (0.554,1.001)	-0.134 (0.542,0.980)
Ideology	-0.965 (0.822,1.026)	-0.964*** (0.823,1.028)	-0.980 (0.819,1.023)	-0.980*** (0.825,1.032)
Egalitarianism	0.976*** (2.867,4.547)	0.977*** (2.856,4.531)	0.974*** (2.843,4.513)	1.005*** (2.883,5.316)
NFC	-0.232 (0.513,1.358)	-0.233 (0.513,1.362)	-0.264 (0.507,1.344)	-0.279 (0.521,1.393)
Education	-0.274 (0.898,1.099)	-0.274*** (0.900,1.101)	-0.281*** (0.897,1.097)	-0.270*** (0.898,1.100)
Political Interest	0.092*** (1.202,1.639)	0.093 (1.191,1.625)	0.124 (1.207,1.645)	0.094 (1.193,1.628)
Age	0.002 (0.989,1.008)	0.002 (0.990,1.009)	0.001 (0.989,1.008)	0.003 (0.990,1.009)
Male	0.160 (0.815,1.498)	0.159 (0.828,1.525)	0.150 (0.810,1.490)	0.186 (0.837,1.542)
Imp Amb:Amb Direction	-	-0.057 (0.991,4.109)	-	-
Imp Amb*Ideology	-	-	0.425 (0.905,1.503)	-
Imp Amb*Egalitarianism	-	-	-	0.911** (0.418,1.411)
Constant	2.071 (0.667,3.475)	1.633* (0.281,2.985)	-2.171* (-3.341,-1.002)	4.991*** (3.970,6.011)
AIC	1026.632	1028.706	1019.751	1020.788
Log Likelihood	-503.316	-503.353	-498.875	-499.394
N	753	753	753	753

Coefficients are logit coefficients with 95% confidence intervals in parentheses. ***p < .001; **p < .01; *p < .05; [‡]p < .10.

Chapter 5

Applicant Descriptions

[All descriptions were preceded by the phrase, "A man that:", which was shown constantly at the top of the screen below the response options]

has never left the United States.	almost never eats vegetables.
has 3 pet birds.	cannot tell his left from his right.
likes to travel to exotic places.	chews with his mouth open.
claims to have seen a ghost.	is a vegetarian.
is eccentric.	once hallucinated from a bad fever.
loves to go to parties.	prefers ice cream in a cup rather than a cone.
never learned to ride a bike.	has a habit of nail biting when nervous.
is known to march to his own drummer.	has been described as very sensitive.
has a hard time remembering birthdays.	enjoys playing sports.
enjoys making jokes at family gatherings.	enjoys playing video games.
winks a lot.	has never seen the ocean.
hums when eating.	is deeply religious.
has an unusually high-pitched voice.	collects insects.
tends to follow the crowd.	has a stutter.
prefers heat over cold weather.	enjoys science fiction novels.
talks to his mother once per week.	was once bitten by a venomous spider.
hates the feel of velvet.	believes in many conspiracy theories.
only wears brown shoes.	performs magic as a hobby.
enjoys dressing in many different styles.	enjoys solving crossword puzzles.
hates people who complain a lot.	wanted to be an astronaut as a child.
has never had a cavity.	reads many philosophy books.
tends to mumble.	does not like solving puzzles.
talks very fast.	was in a rock band as a teenager.
talks with his hands.	was in the marching band in high school.
has a lot of body hair.	is very flirtatious.
accidentally stares at people often.	spends a lot of time looking in the mirror.
believes he has a long lost brother.	never wears shorts.
will be the best man at a wedding.	is prone to severe allergies.
makes odd noises while sleeping.	is very shy around strangers
has a loud voice.	has had many romantic relationships.
has poor hand-eye coordination.	is known to be quirky.
has never been in a fist fight.	does not believe in any religion.

has a large scar on his shoulder.	thinks tattoos are disgusting.
has no sense of rhythm.	is color-blind.
is allergic to cats.	lost his left foot to diabetes.
was once on TV as an extra in a sitcom.	keeps his opinions to himself.
takes a long time to make decisions.	wanted to be a movie star as a child.
stays informed about sports news.	often speaks in rhymes.
enjoys extreme sports.	hates the idea of bottled water.
enjoys romantic comedies.	loves classic movies.
is obsessed with the size and shape of chairs.	wore braces throughout high school.
whistles all of the time.	is fascinated by history.
is a very picky eater.	knows every state capital.
prefers Coca-Cola over Pepsi.	cannot swim.
has won several pie-eating contests.	claims not to fear death.
is a huge comic book fan.	prefers to be indoors.
is said to make the best grilled cheese.	has never been kissed.
regrets not exercising more.	is an only child.
dated a woman with 11 toes.	has four siblings.
plays chess once per week.	prefers dogs over cats.
thinks surfers are very courageous.	had a pet pitbull as a teenager.
spent last weekend at a funeral.	plays drums in a jazz band.
makes bird houses.	has written several romance novels.
still remembers his high school prom.	writes poems on napkins.
knew someone who wound up being famous.	is afraid of germs.
enjoys the fine arts.	is afraid of open spaces.
has an outie belly button.	has a collection of leather jackets.
can run a mile in 6 minutes.	hates puns.
plays video games with his nephew.	is afraid of thunder.
enjoys writing fiction in his spare time.	falls in love easily.
did not like the movie Titanic.	has trouble choosing clothes that match.
always thinks about the big picture.	spends a lot of time thinking about life.
is a sculptor in his spare time.	often daydreams about flying.
is considered a tree hugger.	has never won a game of bingo.
plays many musical instruments.	prefers not to wear shoes.
smoked cigarettes as a teenager.	likes to analyze people.
used to be nicknamed Smiley.	enjoys spicy food.
spent last weekend helping a friend move.	squints a lot.
took karate lessons as a child.	always has mints in his car.

prefers driving over being a passenger.	cannot hear out of his left ear.
received a speeding ticket two months ago.	does not believe love is real.
tends to be vague when speaking.	is great at impressions.
likes looking at the clouds.	collects scented candles.
has an odd fashion sense.	prefers camping over staying in a hotel.
always falls asleep during movies.	
is always making nicknames for people.	
enjoys pulling pranks on people.	
does not like roller coasters.	
stands close to people when talking to them.	
believes time travel will exist soon.	
enjoys standup comedy.	
is bad at keeping secrets.	
often wears sunglasses indoors.	
does not need glasses but wears them anyway.	
has a large collection of action figures.	
believes strongly in Zodiac signs.	