PERSPECTIVE TAKING, STEREOTYPING, PREJUDICE, AND BEHAVIORAL EXPLANATIONS: WHEN, WHY, AND HOW PERCEIVERS TAKE ON THE ATTITUDES OF A TARGET

by

SEAN MICHAEL LAURENT

A DISSERTATION

Presented to the Department of Psychology
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

June 2010

University of Oregon Graduate School

Confirmation of Approval and Acceptance of Dissertation prepared by:

Sean Laurent

Title:

"Perspective Taking, Stereotyping, Prejudice, and Behavioral Explanations: When, Why, and How Perceivers Take on the Attitudes of a Target"

This dissertation has been accepted and approved in partial fulfillment of the requirements for the degree in the Department of Psychology by:

Sara Hodges, Chairperson, Psychology
Bertram Malle, Member, Psychology
Ulrich Mayr, Member, Psychology
Mia Tuan, Outside Member, Education Studies

and Richard Linton, Vice President for Research and Graduate Studies/Dean of the Graduate School for the University of Oregon.

June 14, 2010

Original approval signatures are on file with the Graduate School and the University of Oregon Libraries.

© 2010 Sean Michael Laurent

An Abstract of the Dissertation of

Sean Michael Laurent for the degree of Doctor of Philosophy

in the Department of Psychology to be taken June 2010

Title: PERSPECTIVE TAKING, STEREOTYPING, PREJUDICE, AND BEHAVIORAL EXPLANATIONS: WHEN, WHY, AND HOW PERCEIVERS TAKE ON THE ATTITUDES OF A TARGET

Approved: _____
Sara D. Hodges

A growing body of research has focused on how perspective taking leads people to perceive themselves as "merging" with the target of perspective taking, in terms of how they cognitively represent themselves and the target. In turn, this merging has been shown to facilitate social coordination between perceivers and targets and results in reduced stereotyping of the target's group. Using this past research as a starting point, this dissertation asks a related but new question: Does perspective taking lead perceivers to take on the attitudes of the target of perspective taking, even when these attitudes are socially reprehensible? Specifically, this dissertation tested whether taking the perspective of a racist target leads perspective takers to show greater racism and stereotyping. In Study 1, 102 participants took the perspective of racist male target (or

wrote about a day in his life without taking his perspective or about a day in their own lives), learning about his attitudes from visual information alone. No main effect for perspective taking was found. However, for perspective takers only, greater self-target merging predicted higher explicit racism scores. Also among perspective takers, greater internal motivation to respond without prejudice also ironically led to greater implicit stereotyping. In Study 2, 101 participants took the perspective of a female target who was generally likable but had subtly racist attitudes. Once again, no main effect of perspective taking was found, but for perspective takers, greater external motivation to respond without prejudice led to higher explicit racism scores. In Study 3, 101 participants took the perspective of the same target used in Study 1, but were given information about the genesis of the target's attitudes. The combination of perspective taking and information led to higher explicit racism scores, and this effect was mediated by self-target merging (and not by greater positive regard for the target). Under many circumstances, perspective takers appear to reject taking on a racist target's socially undesirable attitudes, adopting them only when they have been given some reason for why the target holds those attitudes. In addition, motivation to respond without prejudice may lead ironically to greater prejudiced responses.

CURRICULUM VITAE

NAME OF AUTHOR: Sean M. Laurent

PLACE OF BIRTH: Washington, D.C.

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene
University of Massachusetts, Amherst
Greenfield Community College, Greenfield, MA

DEGRESS AWARDED:

Doctor of Philosophy, Psychology, 2010, University of Oregon
Master of Science, Psychology, 2006, University of Oregon
Bachelor of Arts, Psychology, 2004, University of Massachusetts
Associate of Arts, Liberal Arts, 2002, Greenfield Community College

AREAS OF SPECIAL INTEREST:

Perspective Taking and Empathy
Empathic Accuracy
Stereotyping and Prejudice
Psychology of Gender

PROFESSIONAL EXPERIENCE:

Reviewer, Personality and Social Psychology Bulletin, 2006-present

Reviewer, Psychology of Women Quarterly, 2009-present

Reviewer, Sex Roles, 2008-present

Reviewer, Book Proposals, Sage Publications, 2008-present

Reviewer, Book Proposals, Worth Publishers, 2008-present

Honors Thesis Advising, 2008-present

Graduate Teaching Fellow, Department of Psychology, University of Oregon, 2005-2010

Graduate Admissions Committee Member, Department of Psychology, University of Oregon, 2007-2008

Graduate Student Member, Social Affective Neuroscience Faculty Search, Department of Psychology, University of Oregon, 2006-2008

Undergraduate Teaching Assistant, University of Massachusetts, 2003, 2004

GRANTS, AWARDS, AND HONORS:

Donald and Darel Stein Graduate Teaching Award, University of Oregon, 2009

Diversity Dissertation Fellowship, Ford Foundation, 2009-2010

Travel Award, Society for Personality and Social Psychology, 2009

Graduate Research Award, University of Oregon Graduate School, 2009

Marthe E. Smith Memorial Science Fellowship, University of Oregon, 2008

Research Grant, Bend Behavioral Research Center, University of Oregon, 2007

Norman D. Sundberg Fellowship, University of Oregon, 2008

Travel Grants, Psychology Department Graduate Education Committee, University of Oregon, 2006, 2007, 2008, 2010

Research Grants, Psychology Department Graduate Education Committee, University of Oregon, 2006, 2007

Research Grant, Center for the Study of Women in Society, University of Oregon, 2006

Fighting Fund Fellowship, University of Oregon, 2005

Summa cum Laude, University of Massachusetts, 2004

Phi Beta Kappa, University of Massachusetts, 2004

Golden Key International Honour Society, Greenfield Community College, 2003

Phi Kappa Phi, Greenfield Community College, 2002

Ethel Case Award for Outstanding Academic Achievement, Greenfield Community College, 2002

Award of Scholastic Excellence, Greenfield Community College, 2001

Phi Theta Kappa, Greenfield Community College, 2001

PUBLICATIONS

Hodges, S. D., Laurent, S. M., & Lewis, K. L. (in press). Specifically motivated, feminine, or just female: Do women have an empathic accuracy advantage? In J. L. Smith, W. Ickes, J. Hall, & S. D. Hodges (Eds.), *Managing interpersonal sensitivity: Knowing when – and when not – to understand others.* NY: Nova Science Publishers.

Laurent, S. M., & Simons, A. D. (2009). Sexual dysfunction in depression and anxiety: Conceptualizing sexual dysfunction as part of an internalizing dimension. *Clinical Psychology Review, 29,* 573-585.

Laurent, S. M., & Hodges, S. D. (2009). Gender roles and empathic accuracy: The role of communion in reading minds. *Sex Roles, 60,* 387-398.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my adviser, Sara Hodges, for all of the advice and insight she has shared with me over the years. Under her tutelage, I have learned more than simply how to become a better researcher, writer, and thinker; she has also shared innumerable bits of wisdom along the way about how the ins and outs of professional life actually work and how to juggle the many scholarly activities inherent in being a productive researcher. On more than one occasion she has pulled me down out of the clouds and back into reality, and I also thank her for that. Most of all, I thank her for all the patience and support she has shown me over the years. I could not have done this without her.

I would also like to thank all the other members of my dissertation committee for providing thoughtful comments, suggestions, and insight into this research. In particular, I want to express my heartfelt gratitude to Bertram Malle, who has always been there for me – as a fellow researcher, statistics mentor, conversationalist, and lover of the finer things in life. Likewise, I would like to thank Sanjay Srivastava and Gerard Saucier, who along with Bertram, have served on many of my committees, and helped me sharpen my analytical and methodological skills.

All of the graduate student members of the Hodges' Lab have also been a great resource through the years, including Adam Kramer, Karyn Lewis, Ezra Markowitz, and J. Anna Bell. One former member in particular, Dr. Michael Myers, was and continues to

be an excellent sounding board, collaborator, and close friend. Also deserving of thanks are all of the undergraduate members of the Hodges' Lab, without whom none of this work could have been completed, the Sunday poker regulars, who helped sustain me by providing one of my few recreational outlets, and James Warmels, a great friend, dinner mate, and fellow thinker.

All of my love and gratitude goes out to my family, whose understanding and support have continued to sustain me through good times and bad. This includes my "other" parents, Alison, Helmut, and Rosemarie – your kindness, warmth, and encouragement have meant the world to me. In particular, I would like to give eternal thanks to my mother, who has always, without fail, believed in me and supported me, regardless of the paths I have chosen, my father, whose pride in my accomplishment warms my heart, and my sister, who has always been a protector and friend.

Last: There is not enough room in all of the pages of this document to detail the ways that my best friend, adviser, and wife, Heidemarie, has helped me. Lacking the space, I will simply say that she has mentored me in matters statistical, methodological, spiritual, and ontological. She has been my sounding board, my confederate, and my savior in more ways than one. I can only hope that I have returned a fraction of what she has given to me, and I thank her with the entirety of my being.

I would like to dedicate this dissertation to my uncle, Ennis Dexter Locke, Jr. Dexter was a brilliant man who died unexpectedly during my tenure at the University of Oregon, and I wish he could be here to share this moment with me.

TABLE OF CONTENTS

Chapter Page

Chapter Page

LIST OF FIGURES

Figure Page

LIST OF TABLES

Table Page

CHAPTER I

GENERAL INTRODUCTION

In the film *Avatar* (Landau & Cameron, 2009), *Neytiri* – a female humanoid inhabiting a distant planet – teaches the human Jake Sully the traditional greeting in her language, which translates as "I see you." For those who have not seen the film, this phrase is not meant to be mundane, as in "I see your physical being with my eyes." Instead, it is meant to signify the essential connectedness that can occur between thinking beings, the idea that one might be able to see *into* another person, looking inside this person's mind and being connected to everything that exists there, including those things which the other is connected to. In fact, this simple phrase is a translation of the Sanskrit word *Namaste*, which literally translates as "The God in me sees the God in you," or more aptly, "I see myself, in your eyes" (Michaelson, 2009).

In essence, these words capture much of what is meant by perspective taking, at least in terms of the psychological experience of trying to understand another person. As Ruby and Decety (2004) aptly put it: "Adopting another person's perspective involves more than simply focusing our attention on the other. It involves imagining how that person is affected by his or her situation…" (p. 988). Taking another's perspective, we feel their pain, both physical and emotional (Batson, Early, & Salvarani, 1997; Jackson, Brunet, Meltzoff, & Decety, 2006; Lamm, Batson, & Decety, 2007), but also feel their joy (Smith, Keating, & Stotland, 1989). We come to value the other (Batson, Eklund,

Chermok, Hoyt, & Ortiz, 2007), feeling more nurturant toward them (Batson, Lishner, Cook, & Sawyer, 2005) and similar to them (e.g., Cialdini, Brown, Lewis, Luce, & Neuberg, 1997; Goldstein & Cialdini, 2007; Maner, Luce, Neuberg, Cialdini, Brown, & Sagarin, 2002; Neuberg, Cialdini, Brown, Luce, Saragin, & Lewis, 1997; Williams, Parker, & Turner, 2007), which can lead to valuing the same things that the other values (Hilmert, Kulik, & Christenfeld, 2006).

At the outset, one issue should be addressed: At times, experimental manipulations which have asked participants to take a target's perspective have been called empathy inductions or manipulations of empathy (e.g., Batson, Klein, Highberger, & Shaw, 1995), and as discussed below, the term "empathy" has been used by Batson and colleagues (e.g., Batson, Fultz, & Schoenrade, 1987) to describe feelings of concern for a person other than the self. Other researchers and theorists have also defined this term in related but distinct ways. For example, empathy has also been described as the feeling that one understands another person's emotional state – which may include feeling similar to how the other feels (e.g., Eisenberg & Fabes, 1990; see also Dovidio et al., 2004; Finlay & Stephan, 2000) – or as an automatic "sharing" of others' positive and negative emotions, which may occur without deliberation or awareness (e.g., Decety & Meyer, 2008).

Perspective taking, on the other hand, can be thought of as something distinct from empathy (e.g., Hodges & Wegner, 1997), even though taking another's perspective often leads to concern for the other (e.g., Batson et al., 1987). Specifically, even though perspective taking may come about automatically (e.g., spontaneously adopting another's

"point of view") or deliberatively (e.g., trying to "get inside someone's head), it may be best characterized as a cognitively complex process of trying to understand what another person – whether real or imagined – is thinking or feeling (or might *typically* think or feel). It is form of thinking about others that may involve trying to imagine the self *as* the other, reasoning about the other, direct knowledge of the other, using stereotypes about the other's social category membership(s), or imagination, all as ways to try and understand the thoughts, feelings, motives, and behavior of a person other than the self. Thus, even though some researchers have called this complex form of trying to understand others "empathy" (e.g., Ickes, 1997), and perspective taking has both cognitive and affective consequences, the term "perspective taking" is used here as described above. That is, it is used to describe an effortful and cognitive form of empathy where one person tries to imagine another person, using whatever tools are available to do so.

One consequence of perspective taking is a psychological merging, at least to some extent, with the target whose perspective has been taken (e.g., Davis, Conklin, Smith, & Luce, 1996; Smith, Coats, & Walling, 1999), whether this merging is implied by bottom-up processes such as mimicking the other's behavior (Chartrand & Bargh, 1999), or by top-down processes such as simulating another person's mind (Ames, 2005; Decety & Grèzes, 2006). It is also sometimes operationalized as expanding the self to include the other (Aron, Aron, Tudor, & Nelson, 1991), or as using the self to understand the other (Davis et al., 1996). Furthermore, this merging is not only psychological (in terms of our thoughts and feelings becoming more connected to the other), but

neuropsychological in that research on perspective taking has found similar patterns of neural activation for perceivers who are thinking about themselves and others (e.g., Ames, Jenkens, Banaji, & Mitchell, 2008; Decety & Jackson, 2006). This does not imply that there is psychological *indistinguishability* between the self and the other (Aron, Aron, & Smollan, 1992; Batson, 1997) – in fact, some type of regulatory mechanism is needed in order to differentiate the self from the other, and merging does not entail a complete loss of the self in the other (Decety & Chaminade, 2003; Decety & Hodges, 2006; Decety & Sommerville, 2003; Jackson et al., 2006; for reviews, see Decety & Chaminade, 2003; Mason & Macrae, 2008). It does imply, however, that physiological and psychological merging are but two possible consequences of perspective-taking among many, and that this merging of identities may have profound implications for our interactions with the social world.

Overview of the Dissertation

A substantial portion of research on perspective taking has focused on the positive consequences of putting oneself in another's shoes and experiencing the world through this person's eyes. For example, taking the perspective of a target has been shown to influence attitudes toward the target and the group to which the target belongs, leading to greater offers of help for both. The primary purpose of this dissertation, however, is to explore a darker side of perspective taking, showing that it can have adverse effects as well as beneficial outcomes.

After taking the perspective of a target who holds negative attitudes toward a group, the perspective taker might temporarily come to share these same attitudes to

some extent. Specifically, across three studies, I explore the consequences of having a perceiver take the perspective of a target who presumably has negative attitudes toward African Americans (blacks) on the perceiver's subsequent explicit and implicit attitudes toward the same group. As part of this work, I identify and explore theoretically-based variables that could limit or exacerbate this effect, such as the nature of the target's attitudes (blatant or subtle), the perceived similarity of the target to participants, and whether any reasons have been provided that could explain the target's attitudes. Because previous research has shown how motivation to respond without prejudice can sometimes lead to rebound effects after stereotype suppression, I also investigate motivation to respond without prejudice as a potentially important moderator. This is because after taking the perspective of a racist target, participants may try to banish any thoughts of prejudice from their own minds. A final goal that I address in Study 1 is to test whether taking a target's perspective leads participants to explain this target's behavior in ways that more closely resemble the ways actors describe their own (versus others') behaviors, and test whether this effect is moderated by the racial content of the situations being explained.

In the following sections, I first review the literature on the beneficial effects of perspective taking and then examine the literature specifically concerned with perspective taking and intergroup attitudes. Next, I consider examples of when perspective taking has negative consequences. Following that, I outline and present results and discussion from each of the studies I conducted.

The Positive Consequences of Perspective Taking

Perspective taking is an important part of what it means to be human (Malle & Hodges, 2005). Starting from a very early age, human beings begin to demonstrate their sensitivity at decoding and understanding other minds, and developmental psychologists have documented this sensitivity in multiple ways (Moses, 2005; Trevarthen, 1979; Zahn-Waxler, Radke-Yarrow, & Brady-Smith, 1977). For example, soon after birth humans begin to imitate other's actions (Meltzoff, 1995) and attribute intentions and mental states to others (Johnson, 2000; Malle, 2005), which may lead to early forms of prosocial behavior (Repacholi & Gopnik, 1997).

In adults, perspective taking continues to be an important part of positive social functioning, and has been linked to higher accuracy in person perception (Bernstein & Davis, 1982; but see also Hodges, 2005), lower interpersonal aggression (Richardson, Hammock, Smith, Gardner, & Signo, 1994), greater sensitivity to other's plights (Clore & Jeffery, 1972), higher relationship satisfaction (Franzoi, Davis, & Young, 1995), and greater willingness to forgive another's transgressions (McCullough, Worthington, & Rachal, 1997). Perspective taking has also been shown to increase environmental awareness, leading to pro-environmental moral reasoning, concerns, attitudes, and behavior (Berenguer, 2007, 2010; Schultz, 2000; Sevillano, Aragonès, & Schultz, 2007). At work, in groups, and when competing with others for common resources, perspective taking is related to greater cooperation (Parker & Axtell, 2001), successful conflict resolution (Galinsky, Maddux, Gilin, & White, 2008), decreased egocentric biases and more realistic judgments of fairness (Caruso, Epley, & Bazerman, 2006), and a reduction

in the impact of first offers on outcomes (which are usually biased in favor of who makes the first offer) in business settings (Galinsky & Mussweiler, 2001).

The positive outcomes which grow out of perspective taking clearly abound. And the body of work by Batson and his colleagues (for reviews, see Batson, 1987, 1991, 1998, 2002, and 2010) is a testament to how perspective taking can positively influence social outcomes. In an extensive program of research conducted over the last 30 years, Batson has documented the effects of perspective taking on affect, cognition, and behavior. No area, however, has been studied by Batson as much as the effects of perspective taking on subsequent helping for a target (e.g., Batson, 1997; Batson & Ahmad, 2001; Batson, Batson, et al., 1995; Batson, Chang, Orr, & Rowland, 2002; Batson et al., 1988, 1989, 1991, 1995, 2003, 2005, 2007; Batson, Early, & Salvarani, 1997; Batson & Moran, 1999; Batson, O'Quin, Fultz, Vanderplas, Isen, 1983; Batson, Polycarpou, et al., 1997; Batson, Sager, et al., 1997; Batson, Turk, Shaw, & Klein, 1995; Batson & Weeks, 1996; Coke, Batson, & McDavis, 1978; Fultz, Batson, Fortenbach, McCarthy, & Varney, 1986).

Perspective Taking and Helping Behavior

Although a substantial portion of Batson and colleagues' research has concerned questions about the nature and existence of altruism and altruistic helping behavior, and this dissertation is not specifically concerned with either altruism or helping, it seems prudent to at least briefly discuss the overarching theoretical models derived from this work. This is because studies generated to test the empathy-altruism hypothesis have led directly and indirectly to the present research.

In early work Batson et al. (1987) discussed how state-specific emotional responses to a person in need can be broken up into what they called "empathy" – other-directed emotions such as "warmth" and "sympathy" (at the trait level, also called "empathic concern;" Davis, 1983a) – and "personal distress," which involves self-directed emotions such as "worry" and "alarm." Later, distinctions were also drawn between asking participants to imagine the other and their situation versus imagining oneself in the others' place (Batson, Early, & Salvarani, 1997).

The first form of perspective taking, where experimental instructions ask participants to imagine the target and how the target feels, is hypothesized to lead to empathy, which in turn leads to altruistic helping (i.e., helping the other without benefit to the self). The second form, which is found when a participant is instructed to imagine him or herself in the target's situation, is hypothesized to lead to a mix of empathy and personal distress. While this second form of perspective taking is thought to be potentially even more effective at motivating helping behavior (Batson, Early, & Salvarani, 1997, p. 757), the self-focused emotions which are aroused may interfere with true altruistic helping.

These distinctions between "self-directed" and "other-directed" emotions as proximal causes of helping, while advanced to address the question of altruistic helping, also led to a substantial amount of additional research. Competing models were put forth which suggested that the effects of perspective taking on helping were driven by cognitive merging between perceivers and targets, and that helping inspired by perspective taking was not truly altruistic. These new models in turn spurred even more

research and a search for variables representing other consequences of perspective taking, such as greater perceived similarity to and valuing of targets.

In 1997, a debate took place in the pages of the *Journal of Personality and Social Psychology* over what was causing the relationship between perspective taking and helping, capturing these differences of opinion. Cialdini et al. (1997; Neuberg et al., 1997; see also Maner et al., 2002) argued that the psychological merging that takes place between perceivers and targets precludes true altruism (i.e., if perceivers and targets are to some extent psychologically merged, then helping a target is in effect helping the perceiver). On the other hand, Batson and colleagues (e.g., Batson, Sager, et al., 1997) continued to argue that because taking the perspective of a target in distress leads to *other*-oriented feelings, such as increased valuing of the target and empathic concern, increased helping is due to altruistic motives. Without taking a stand on the nature of altruism, two conclusions can be drawn from the research which resulted from this debate. First, perspective taking is associated with greater empathic concern and personal distress, and empathic concern can lead to greater helping. Second, perspective taking is also associated with perceiver target merging, greater perceived similarity to targets, and greater valuing for targets.

A point raised above about perspective taking instruction should be elaborated further here, as it is directly relevant to the present research. As Batson, Early, and Salvarani (1997) pointed out, asking participants to imagine themselves *as* the target leads to more self-directed emotion than asking participants to imagine the target and their situation. In the present case, the question of interest concerns whether taking a

target's perspective can lead participants to express attitudes similar to those that the target might be expected to express. In effect, this means that the expectation is that perceivers will become more like targets. Because of this, it seems likely that an "imagine self *as* the other" induction would be more successful than an "imagine other" induction at achieving this goal.

Furthermore, there is no theoretical reason to believe that increasing empathic concern (such as by using an "imagine other" instruction) would lead to greater overlap in attitudes between perceivers and targets. It does, however, seem likely that using "imagine self" instructions would lead to a greater possibility of perceiver-target merging (also called self-other overlap or "oneness;" e.g., Aron et al., 1991; Batson, Sager, et al., 1997; Davis et al., 1996; Galinsky et al., 2005) and perceived similarity (e.g., Batson, Turk, et al., 1995; Cialdini et al., 1997; Davis et al., 1996; Goldstein & Cialdini, 2007). In light of this, an "imagine self" instruction again seems called for, because the central question here is whether perceivers' attitudes toward a social category become more similar to (or shared/merged with) those of a target. Put another way – when a target is assumed to possess particular attitudes about a social group, does perspective taking cause perceivers to feel the same way as the target about this group?

Perspective Taking and Improving Intergroup Attitudes

A relatively small but growing literature has investigated the beneficial effect that perspective taking can have on intergroup attitudes. Along with this has come a variety of models to explain the effect. Batson was also one of the pioneers of this research. In an early study, Batson, Polycarpou, et al. (1997) demonstrated that perspective taking can

improve attitudes toward the stigmatized groups to which a target belongs. In a series of experiments, participants were asked to take the perspective (versus remain objective) of a young woman with AIDS (Study 1), a homeless man (Study 2), and a convicted murderer (Study 3). In each case, perspective taking led to improved attitudes not only toward the target, but toward the targets' groups (i.e., people with AIDS, homeless people, and convicted murderers in general).

Batson, Polycarpou, et al. (1997) also tested the moderating effect of victim responsibility. That is, a stigmatized target might or might not be responsible for his or her plight (e.g., contracting HIV by having unprotected sex versus by having a blood transfusion). Although the moderating effect of responsibility was mixed, the idea that targets might be responsible for their situation is an important one to consider (e.g., Weiner, 1980; Weiner, Perry, & Magnusson, 1988). This is because when a target is from a stigmatized group and is thought to be responsible for his or her situation or attitudes (that is, the situation or attitudes are under the control of the target), empathy toward the target might be limited, and participants' might be less able or willing to take that target's perspective (for a discussion of down-regulation of empathy, see Hodges & Biswas-Diener, 2007; Hodges & Klein, 2001).

In another study by Batson et al. (2002), perspective taking not only affected attitudes toward a stigmatized target (a convicted heroin addict and dealer) and his group, but also influenced subsequent helping for that group (drug addicts). This was true even when the helping behavior could in no way benefit the target. This study thus contributed two noteworthy findings that are applicable to the present research. First, it showed that

even when a target will receive none of the help offered by perceivers, perceivers still help members of the target's group. In essence, this demonstrates how the effects of perspective taking can generalize beyond targets themselves and extend to the larger groups to which targets belong. Second, perceivers not only helped the target's group, but helped this group at the at the expense of other worthy groups. That is, helping in this study was indexed by the amount of funding (out of a fixed amount) that participants asked be allocated to the target's group. According to the cover story, this fixed amount of funds was currently (and had previously been) supporting other groups, so participants' decisions to allocate money to the target's group meant that they helped the target's group even when they were aware that to do so, other groups would have to lose their funding. Thus, even when perspective taking is causing positive outcomes (e.g., helping the target's group), it is also having negative effects (taking resources away from other needy and deserving groups).

In other research on perspective taking and intergroup attitudes, Finlay and Stephan (2000) found that participants who read about discriminatory acts against African Americans or took the perspective of victims of discrimination later reported more parallel empathy (e.g., anger, frustration) and had smaller differences in the way that they viewed ingroups versus outgroups. Dovidio et al. (2004) found that self-other merging was increased for perspective takers, and that a form of parallel empathy mediated the effects of condition on a decrease in prejudice from an earlier testing session. Similarly, Vescio, Sechrist, and Paolucci (2003) found that irrespective of the

target's manipulated stereotypicality, perspective taking (versus taking an objective perspective) led to greater endorsement of pro-black attitudes.

In several studies, Galinsky and colleagues also found support for the idea that perspective taking can increase positive attitudes toward the group to which a target belongs. For example, Galinsky and Moskowitz (2000) found that perspective taking was a better strategy than stereotype suppression for reducing stereotypes of the elderly; in a suppression condition, there was a significant rebound effect on a lexical decision task (i.e., an indirect/implicit measure of elderly stereotypes) even though an earlier essay by these participants about a day in the life of an elderly target was no more stereotypical than perspective takers' essays (essays in both of these conditions were less stereotypical than an essay written in a control condition). For perspective takers, a second essay about a different elderly target was also more positive than those in either of the other conditions, as was a third essay written about a different group (African Americans). In a second and third study, Galinsky and Moskowitz showed that a trait overlap measure accounted for differences in the stereotypicality of essays, and that even when the groups involved were minimal groups, perspective taking led to less bias in ingroup and outgroup ratings.

Galinsky and Moskowitz's (2000) findings were discussed in terms of the self-concept becoming activated and applied to the target during perspective taking, with resultant self-other merging that leads to less stereotyping of the target's group. That is,

increased accessibility of the self-concept after perspective taking may result in the self-concept "winning out" over stereotypes when evaluating members of the stereotyped group.

Additional studies by Galinsky and colleagues (Galinsky & Ku, 2004; Galinsky et al., 2005; Galinsky, Wang, & Ku, 2008) demonstrated that decreased stereotyping of the target's group can be moderated by self-evaluation and that not only do perspective takers come to stereotype the target and the target's group less, they ironically come to act in ways that are consistent with the very stereotypes they avoid in their essays. This again highlights how perspective taking can simultaneously lead to prosocial benefits (such as stereotyping a target's group less) and negative consequences (such as acting in stereotypical ways). In the next section, this idea is discussed further.

A Darker Side of Perspective Taking

While the majority of perspective taking research has focused on its prosocial benefits, perspective taking can lead to negative outcomes as well. For example, consider that perspective taking is one form of empathy, and that in a very general sense, empathy can be a costly affair that people try to regulate or sometimes avoid (e.g., Hodges & Biswas-Diener, 2007; Hodges & Klein, 2001; Shaw, Batson, & Todd, 1994). Consider also that perspective taking leads not only to empathic concern, but to distress for the self (e.g., Batson et al., 1987, 1983; Batson, Early, & Salvarani., 1997; Davis, 1983b; Davis et al., 1999; Maner et al., 2002), or to perceivers taking on the assumed mood states of

others, such as embarrassment (Miller, 1987). In addition, knowing that one has not been successful in an empathy-induced helping attempt leads to negative mood (Batson & Weeks, 1996).

There are also several examples of when the same perspective taking manipulation simultaneously leads to both positive and negative effects. As noted above, Batson et al. (2002) found that taking the perspective of a drug addict and dealer positively influenced perceivers' attitudes and helping behavior toward drug dealers in general, but that this prosocial benefit for drug dealers came at the expense of other programs that were also worthy and were already being funded (including a program that helped mentally retarded children learn new skills, an educational program for single parents, an environmental advocacy and action group, and an animal rights group). In a similar study (Batson, Klein, et al., 1995), taking the perspective of a terminally ill child led participants to move her up on a treatment list, even though other children on the list were more in need, had less time to live, or had applied earlier for assistance. Clearly, in both of these cases, helping the target also hurt others.

These studies are not alone in showing undesirable effects of perspective taking. Epley, Caruso, and Bazerman (2006) found that even though perspective taking led to more realistic judgments about fair resource allocation for the self versus others, reactive egoism (expecting others to act in a selfish way) led participants to actually claim a higher share of community resources. Caruso et al. (2006) similarly found that perspective taking decreased egocentric biases in beliefs about contributions to work groups, but also resulted in decreased satisfaction with the group work and less desire for

future collaboration. In other examples, perspective takers have made irrational choices for themselves in a single trial prisoner's dilemma (Batson & Ahmad, 2001; Batson & Moran, 1999), have acted in ways that conflict with justice (Batson, Klein, et al., 1995), have increased positive outcomes for the target at their own expense (Batson et al., 2003), have favored targets to the detriment of the collective (Batson, Batson, et al., 1995), and have experienced greater stereotype threat (Marx & Stapel, 2006).

Taking the perspective of a target who is a member of a stigmatized group may also cause problems for perceivers. As noted above, the perceiver may come to act in ways that are congruent with the stereotypes of the target (Galinsky et al., 2008). For example, while taking the perspective of a cheerleader may lead one to like cheerleaders more and to maybe even contribute to a pep-rally fund, it can also lead one to perform more poorly on a test of intellectual ability (Study 4). Further, when taking the perspective of targets from stigmatized groups (such as racists), psychologically "merging" with this target might lead to "stigma by association" (Galinsky et al., 2005, p. 119). Last, and most importantly for the current research, the perspective taker might come to share some of the same intergroup beliefs and attitudes as the target, subsequently expressing these attitudes and "distancing [themselves] from the rest of the world" (Galinsky et al., 2005, p. 120).

This idea, that taking the perspective of a target who holds negative attitudes may lead perceivers to share these same attitudes and subsequently express them, is a topic that has not been previously investigated, and is the focus of this dissertation. In three

studies, I attempt to answer this question: What happens when a perceiver takes the perspective of a racist?

A Final Word: Behavioral Explanations

As discussed above, Vescio et al. (2003) found that irrespective of the target's manipulated stereotypicality, taking the perspective of a black target (versus remaining objective) led to greater endorsement of pro-black attitudes. Unlike the empathic concern model and the self-other merging models, Vescio et al. found that perceivers' attributions mediated the effects of perspective taking on attitudes. This raises another intriguing question and calls for a brief digression: Does perspective taking lead perceivers to think about and explain a target's behaviors similarly to how the target might explain those same behaviors?

The roots of this question can be found in traditional attribution theory, which describes an actor-observer asymmetry in how people explain their own versus others' behavior. According to this theory, actors explain their own behavior in terms of situational factors, while observers explain actors' behavior in terms of the actors' stable dispositions, or factors which reside within the person (Jones & Nisbett, 1972). Malle (2004) presents a cogent discussion of how this work appears to be based on a misinterpretation of earlier work by Heider (1958), and in a recent meta-analysis (Malle, 2006), provides compelling evidence that this theory – which is widely thought to be an established fact, and appears in many social psychology textbooks – lacks empirical foundation except under certain circumstances, such as when actors and observers are describing negative events.

The work by Vescio et al. (2003) is not the first to examine the effects of perspective taking on behavioral explanations, using the Jones and Nisbett (1972) framework as a theoretical basis. In short, the idea is that through an empathy induction or a shift in visual perspective, perceivers should explain an actor's actions similarly to how the actor would explain the same actions – as if the perceiver is "inside" the actor, looking out on the world.

A variety of experiments have been set up to test this effect with results that are not inconsistent with what might be hypothesized given the traditional actor-observer asymmetry hypothesis. For example, Reagan and Totten (1975) found that perspective taking instructions caused perceivers to make more situational attributions, and Galper (1976), using participants' free-responses and responses on rating scales, found that instructions to empathize with a target led to fewer internal and greater situational explanations for the target's behavior,. Other researchers have also found similar results (e.g., Brehm & Aderman, 1977; Gould, & Sigall, 1977). Visual perspective taking research has also demonstrated similar findings. For example, one often cited paper by Storms (1973) showed that when actors watched a videotape of an interaction they had with another participant from the same perspective they took during the interaction, they ascribed relatively more situational attributions to their behavior. Reversing their perspective to match that of an observer, however, led to relatively more dispositional attributions (however, see Malle, 2006, 2010, for a discussion of the lack of replication of this effect).

One issue with drawing strong conclusions from this research is that studying behavioral explanations in this way has been called into question on several grounds. First, forcing participants to respond on measurement scales does not accurately reflect the ways that people typically explain behavior. Second, coding of explanatory free responses into disposition-situation causes may overly rely on linguistic surface features of the explanations and ignore the actual intent of these explanations (Malle, 2006; Malle, Knobe, O'Laughlin, Pearce, & Nelson, 2000). A promising alternative for examining behavioral explanations can be found in Malle and colleagues' "folk-conceptual" theory of behavioral explanation (e.g., Malle, 1999). This coding system relies primarily on coding of free responses rather than using rating scales, and addresses important distinctions in the ways that people actually explain behavior, such as distinguishing between explanations of intentional versus unintentional or observable versus unobservable behaviors. By avoiding some of the potential pitfalls of traditional coding systems (e.g., rating scales that force perceivers to choose between person and situation causes of behavior), use of this coding system can provide additional insight into how perspective taking influences behavioral explanations.

Outside of perspective taking research, a substantial and growing amount of research has used this theoretical framework to explain actor-observer asymmetries in behavioral explanation. For example, actors choose to explain their unintentional and unobservable behaviors more than their observable, intentional behaviors, while the reverse is true for observers (Malle & Knobe, 1997; Malle & Pearce, 2001). Although both actors and observers tend to use more reasons than other explanation modes for

intentional behaviors (Malle, Knobe, & Nelson, 2007), actors also tend to use more reasons (i.e., the reasons in light of which the actor acted) and fewer causal histories of reasons (i.e., factors that precede an action, and describe what might have brought about the reasons for the action; cf. Malle, 1999) than observers, while observers tend to use relatively fewer reasons and more causal histories of reasons to explain actors' behavior (Malle et al., 2007). Furthermore, actors use relatively more belief reasons (e.g., "I think, I felt," etc.) and fewer desire reasons (e.g., "I need, I wanted," etc.) than observers, and use fewer explicit mental state markers (e.g., "Why did you walk the dog?" *"He had to go to the bathroom"* versus "*I* thought *he had to go to the bathroom"*) than observers (Malle et al., 2007).

As stated by Malle et al. (2007, p. 493): "Providing a reason is [itself] an act of perspective taking…" Therefore, an additional goal of Study 1 in this dissertation is to use the folk-conceptual framework to explore whether perspective taking will lead perceivers (i.e., observers of a target's behavior) to explain a target's behavior in ways that the target (i.e., an actor) might explain his or her own behavior. Furthermore, because the target in this first study is not one whom most perceivers will like, participants may try to distance themselves from the target's behaviors – which is opposite from what one would expect from a person explaining their own behaviors. In other words, participants' impression management goals may undermine the effects of perspective taking, making for a particularly strong test of perspective taking hypotheses. Research that has shown how impression management goals affect the ways that people explain their own and others' behaviors can shed some light on this assertion.

Malle et al. (2000) demonstrated that when participants were given an instruction to appear self-centered versus rational, they cited marked desire reasons, "thus highlighting their wants and needs rather than their thinking" (p. 322). In another study in the same paper (Study 6), participants read one of four versions of a vignette that provided an actor's ostensible reasons for why his girlfriend did not eat dessert. When the actor's reason cited an unmarked belief ("She's been gaining weight") as opposed to a marked belief ("She *thinks* she's been gaining weight"), participants liked the actor significantly less (p. 323). Further, evidence was presented that this effect on liking for this actor was related to the idea that in the marked belief version, he was distancing himself from his girlfriend's belief, suggesting that he himself did not believe that she was gaining weight.

In sum, taking a target's perspective may cause perceivers to explain the target's behavior in ways that are more like the ways the target would explain his or her own behavior. In the present studies, because the target is from a group that is not well-liked (i.e., racists), if these predicted asymmetries occur it will provide strong evidence of the influence of perspective taking on behavioral explanations. On the other hand, impression management concerns may be especially salient, leading participants who take the perspective of a racist target to try and subsequently distance themselves from the target, and it remains to be seen which of these forces will win out in the end.

CHAPTER II

STUDY 1: TAKING THE PERSPECTIVE OF A RACIST FROM VISUAL INFORMATION ALONE

Introduction

As discussed earlier, a small but growing body of work has shown that perspective taking can be used to improve attitudes towards not only targets, but towards their groups (Batson et al., 2002; Batson, Polycarpou, et al., 1997; Dovidio et al., 2004; Finlay & Stephan, 2000; Galinsky et al., 2005; Galinsky & Ku, 2004; Galinsky & Moskowitz, 2000). Similarly, research has shown that after perspective taking (Galinsky et al., 2008), similarity inductions (Goldstein & Cialdini, 2007), or simply writing about a "day in the life" of a target in the first person (Wheeler, Jarvis, & Petty, 2001), perceivers can come to act in ways that are concordant with stereotypes of the target's group (although, see Sedikides, 1992, and Smeesters, Wheeler, & Petty, 2009, for examples of when using a third person writing task can also induce a focus on the other which may be similar to perspective taking).

These studies, taken together, represent a variety of theoretical models. For example, Finlay and Stephan (2000) and Dovidio et al. (2004) focused on how perspective taking leads to parallel empathy rather than reactive empathy while Vescio et al. (2003) examined the effects of perspective taking on attributions and the relationship of these attributions to positive attitudes toward the target's group. Batson and

colleagues' work (e.g., Batson et al., 2002) focused primarily on how perspective taking leads to empathic concern, and on demonstrating that the effects of perspective taking are explicitly *not* due to self-other merging. In another camp, Galinsky et al. (e.g., 2005) and Goldstein and Cialdini (2007) focused specifically on the ideas of self-other merging, similarity, and shared identities rather than emotional responses as a consequence of perspective taking (for other examples of similar themes, see Davis et al., 1996; Maner et al., 2002; see also Smith & Henry, 1996, for a discussion of ingroup merging with the self, and Aron et al., 1991, 1992; Smith et al., 1999, for discussions on self-other merging with a relationship partner).

Departing from these accounts and based on a separate research literature concerned with behavioral priming (e.g., Bargh, 1994; Bargh, Chen, & Burrows, 1996; Wheeler & Petty, 2001) comes yet another theoretical model accounting for the effects of perspective taking. In this account, Wheeler and colleagues (e.g., Wheeler, DeMarree, & Petty, 2005; 2007; Wheeler & Petty, 2001) suggest that there is no need for conscious or deliberative perspective taking or for an emotional connection to others for perspective taking to affect behavior. Instead, perspective taking might be thought of as just another form of priming (where the target is the prime). Because the self-concept is fluid and malleable, perspective taking might activate prime-related information from within the self-concept, or might even introduce new information related to the prime into the self-concept.

In some ways, this account is not much different from other accounts of perspective taking or self-other merging, in that the attributes of a target can become at

least temporarily activated within the self-concept, or the self-concept can be applied to targets. In other ways, however, it differs by focusing primarily on priming effects and automatically activated behavior, and does not consider cognitively "becoming one" with another necessary – or perhaps it redefines what "becoming one" means.

Although this theory does not take into account available data on self-other merging and sets aside the fact that perspective taking is often goal directed (e.g., explicitly and purposively trying to understand what is going on in another person's mind), it nonetheless brings to light another aspect of the complexity of perspective taking. By framing a perspective taking attempt in terms of incidentally activated knowledge about the self or the other (e.g., stereotypes), and suggesting that this knowledge might become a temporary part of the self-concept and subsequently applied in behavior, it highlights in a new way the fact that there may be quite unintentional consequences of perspective taking (Davis, 2005). That is, it makes even more clear that perspective taking can invoke not only controlled but automatic processes (e.g., Hodges & Wegner, 1997), and might invoke them simultaneously (e.g., Davis et al., 2006).

Considering all of the documented effects of perspective taking and all of the theoretical models that have been proposed to explain its effects, it seems rather unlikely that any model represents "one true" reality. Instead, it seems reasonable to assume that perspective taking can have a host of different effects depending on who or what the target is, what situation the target is in (if any is specified), or what group is the target from (and what stereotypes, attitudes, and judgments are available or salient about the group). Effects might also be determined by characteristics of the perceiver, such as how

deeply a perceiver is willing to think about the target's mental states and intentions, the mental state of the perceiver (e.g., whether the perceiver has available cognitive resources to spend on understanding the target), how motivated the perceiver is to understand the target, what the perceiver's perspective taking goals are, and the perceiver's attitudes about the target. Equally likely to be important are perceiver-target interactions, captured in facets such as the perceived attitudes and beliefs of the target, perceptions of the target's motivations or goals, how much the target is liked, how similar the target is to the perceiver (or how similar the perceiver *thinks* the target is), and the relationship (or perceived relationship) between the perceiver and target. While this list suggests a variety of effects, it is in no way exhaustive, and merely presents a sample of the ways that perspective taking will influence emotion, cognition, and behavior.

Of note, the current research does not consider affective responses toward the target such as empathic concern, as the majority of research which has considered these responses has focused on helping behavior for targets who are in distress (e.g., Batson, 2010). It does, however, consider three potential mediators that have been advanced and examined in the literature (similarity, positive regard, and perceiver-target merging), and also examines the potentially moderating effect of motivation to respond without prejudice.

Mediators and One Moderator

Similarity (e.g., Goldstein & Cialdini, 2007), self-other merging (e.g., Galinsky et al., 2005), and positive regard for the target (e.g., Batson et al., 2007; Batson, Polycarpou, et al., 1997) have all been investigated as mediators of the effects of perspective taking

on behavior, and each is also considered here. That is, perspective taking is expected to lead to increased perceived similarity to and positive regard for the target, and also to greater perceiver-target merging. If this is the case, and if condition is also found to directly influence participants' attitudes, then each of these variables will be tested as a mediator.

In addition to these potential mediators, motivation to respond without prejudice (Plant & Devine, 1998) is considered as a possible moderator of the effects of condition on attitudes. This is primarily because college students (i.e., the participants in the current study) are very aware of stereotypes about African Americans and feel substantial pressure to avoid using these stereotypes (Judd, Park, Ryan, Brauer, & Kraus, 1995). Moreover, motivation to respond without prejudice has been shown to influence participants' responses on measures of racial attitudes (e.g., Devine, Plant, Amodio, Harmon-Jones, & Vance, 2002). In the present case, because participants (particularly perspective takers) are thinking about and writing about a very racist target, their goals to distance themselves from the target may make their desire to avoid appearing prejudiced especially salient. That is, even though previous research has used targets similar to the target used in the present study (e.g., Macrae, Bodenhausen, Milne, & Jetten, 1994), no research thus far has had participants take the perspective of a person who is obviously a racist and then examined the effects of the manipulation on participants' evaluations of the same group about which the target holds negative attitudes.

Research has shown that higher external motivation to respond without prejudice (e.g., a motivation based on fear that other will respond negatively to one's prejudiced

attitudes) is associated with lower explicit prejudice, and that greater internal motivation to respond without prejudice (e.g., a motivation to be nonprejudiced because of an egalitarian self-concept) combined with low external motivation is associated with the lowest levels of implicit prejudice (e.g., Plant et al., 2002). Other research, however, has found that greater motivation can lead to greater stereotyping after stereotype suppression. For example, Wyer (2007) found that high motivation to control prejudice and high external but low internal motivation to respond without prejudice both led to greater stereotyping following attempts to suppress stereotypes about blacks. Considering that suppression is one way that people try to banish unwanted thoughts about others from mind, and that suppression can lead to subsequent rebound of these unwanted thoughts (Macrae et al., 1994; see also Plant & Devine, 1998; Wyer, Sherman, & Stroessner, 2000), it seems plausible that participants who have just spent time imagining themselves *as* a racist might subsequently try to banish negative thoughts about blacks from their minds (i.e., suppress them). This might be particularly true for those participants who are highest in motivation to avoid appearing prejudiced.

Hypotheses

Based on the above, a first hypothesis concerns the potential mediators outlined above. Specifically, compared with other conditions (described fully in the Method section, below), it is predicted that perspective takers will rate the target as more similar, have more positive regard for the target, and feel more merged with the target (Hypothesis 1). Exploratory analyses will also examine whether any of these variables directly affect participants' attitudes toward African Americans as a group.

As part of the experimental manipulation, all participants will write about a day in the life of the target. Similar to other perspective taking manipulations (e.g., Galinsky & Moskowitz, 2000), some participants will write about a day in the life of the target as if they "were" that person. Other participants will simply write about a day in the life of the target (without taking his perspective) or will write about a fictional day in their own life. Because participants have no specific instruction to suppress stereotypes of the target, and because no particular norms are in place against holding stereotypes about a racist target (e.g., Wyer et al., 2000), Hypothesis 2 is that participants in the perspective taking condition will write essays that portray the target as more racist than participants in the matched control condition. In effect, writing about a day in the life of the target from the target's perspective may be similar to the "stereotype expression" condition used by Galinsky and Moskowitz (2007). Greater expressions of racist content in participants' essays on behalf of the target might then represent how perspective takers are thinking in more complex ways about what it means to be a racist when they imagine themselves as one. Furthermore, higher racist content in perspective takers' essays might also be a way for participants' to comply with instructions, but also to "punish" the target for his beliefs (i.e., to present him in as negative way as possible).

Third, and representing the primary hypothesis (Hypothesis 3), I predict that perspective takers who write about a day in a racist target's life as if they *were that person* (see Galinsky & Moskowitz, 2000) will come to express attitudes that are more similar to those the target might be expected to express (i.e., greater stereotyping and racism), compared to participants who either write about a day in the life of the same

target without taking his perspective, or participants who write about a day in their own life. These attitudes might be best captured through direct or explicit attitude measures (e.g., Vescio et al., 2003), indirect/implicit measures of attitudes and stereotyping (e.g., Galinsky & Moskowitz, 2000), or a behavioral measure of social distancing (e.g., Macrae et al., 1994).

One reason to consider both implicit and explicit measures is simply that previous perspective taking research has examined both (e.g., Dovidio et al., 2004; Galinsky & Moskowitz, 2000). Others are that these measures may truly be tapping different constructs (e.g., Greenwald & Banaji, 1995), that condition may strengthen automatic negative associations about blacks but leave controlled responding unchanged, or vice versa (e.g., Devine, 1989), that implicit measures will better capture those attitudes which participants are unwilling to admit (e.g., Payne, Burkley, & Stokes, 2008), or that both implicit and explicit attitudes will be affected by condition, but that measurement issues may obscure a relationship between them (e.g., Cunningham, Preacher, & Banaji, 2001). Either way, it seems prudent to include both direct and indirect measures of participants' attitudes, because previous research upon which this dissertation is based has not shown a clear preference for either type of attitude measurement strategy, and further because this research is to some extent venturing into unexplored territory.

When considering the possible effects of condition on participants' perceptions of the target and their own racial attitudes, however, one possibility to consider is that perspective takers might not want to engage with the target. That is, no additional

information is given about the target except that which is conveyed by his appearance (which makes clear his membership in a racist group, as described in the Methods section below). This was done in order to match similar instructions used in previous research which was ostensibly about interest in people's impressions of others which they get "from visual information alone" (Galinsky & Moskowitz, 2000, p. 711; see also Macrae et al., 1994). This raises the possibility, though, that very few participants will like, feel similar to, or merge with the target. In this case, it is also possible that the hypothesized main effects for condition on participants' attitudes will not be found.

Still, as outlined above, it is conceivable that perspective takers who are high in motivation to respond without prejudice might try to suppress thinking in stereotypical ways about African Americans after imagining themselves as a racist target. Therefore, the possibility that condition will interact with motivation to respond without prejudice will be explored, with a tentative prediction advanced that participants in the perspective taking condition who are more motivated to not appear prejudiced will actually show more negative racial attitudes.

One last hypothesis concerns participants' explanations of the targets' (or their own) behaviors during the scenarios included in their essays. Specifically, perspective takers should explain the target's behavior in ways that are more similar to ways that actors explain their own behavior (Hypothesis 4). This may be moderated, however, by the valence of the behaviors which are being explained. All participants were asked to include three scenarios in their essays – two which involved the target engaging in behaviors that were planned to be somewhat ambiguous with respect to racial attitudes

(i.e., the target's behavior could have been propelled by the target's negative feelings toward blacks, or might have been for other reasons), and one which was neutral with respect to race. Therefore, it is possible that perspective takers will try to distance themselves from the target's (potentially) racist behaviors and will not differ from control subjects in the ways they explain his behavior for the scenarios which involve race. Because a "write about self" condition was also included in the design, it will be possible to test whether perspective takers explain the target's behaviors in ways that are more similar to the ways that participants actually explained their own behavior for the same scenarios, compared with the third person control.

Finally, one additional test will concern participants' memory for their essays, memory for their use of race as an organizing theme in the essays, and memory for their actions during the three behavioral scenarios. That is, participants may choose to "intentionally" forget or suppress memories of their essays and the themes they used in constructing these essays because these actions are things they do not want to think about (Macrae, Bodenhausen, Milne, & Ford, 1997). If so, this might explain any differences in outcome measures that are found. To make sure this is not the case, they will be tested for their explicit knowledge of the content of their essays.

Method

Participants

Participants were 102 undergraduate students (75 female, 25 male, 2 did not report gender) from the University of Oregon who participated in partial fulfillment of

course requirements. Participants were allowed to choose multiple ethnic or racial categories, and 11 participants choose to do so, resulting in 86 who identified as white, 1 as black, 13 as Asian, 5 as Latino/Latina, 2 as Middle Eastern, 5 as Native American, and 9 as "other." Ages ranged from 18 – 46 (M_{age} = 19.49, SD_{age} = 3.73). Four participants reported that English was not their first language, and reported speaking English between 3 – 17 years. Analyses with these participants excluded did not differ from those where they were included, so they were retained. The one participant who self-reported as black was excluded from further analyses, as the primary hypothesis concerned attitudes toward African Americans.

Procedure

On arrival, the experimenter (in all cases a white female) explained to participants that they would be taking part in a variety of pilot studies, under the pretext of choosing the best study to be used in future research. Participants in two of three conditions were told that the first task they would do involved forming an impression of a randomly chosen target person, and were asked to choose a number from 1 – 20, which would determine which target they formed an impression of (for a similar method, see Monteith, Spicer, & Tooman, 1998). In front of the participant, the experimenter sorted through 20 envelopes clearly marked #1 - #20, found the appropriately numbered envelope and gave this to the participant. In fact, all the envelopes contained the same target: the photograph pictured a white male with a shaved head, wearing a sleeveless t-shirt, black jeans, black boots, and sunglasses, sitting on a beverage cooler, smoking a cigarette. On his shirt, there was a drawing of an eagle, along with the words "WHITE PRIDE – Save Our

Culture" clearly printed, along with the word "Deutschland" repeated twice. Participants in a third condition also "randomly" chose this target and rated him along several dimensions (discussed below), but did not do so until later in the study.

Next, participants who saw the target photograph first were asked (via instructions on a computer) to either take the perspective of the target person from the photograph and write a first person essay about a day in his life as if they were that person, looking at the world through his eyes and walking through the world in his shoes (perspective taking condition; Galinsky & Moskowitz, 2000), or were instructed to write a third person essay about a day in the target's life (third-person condition). In the third condition (i.e., participants who had not yet seen the photograph), participants were instructed to write about a fictional day in their own life (self condition). (See Appendix A for the full set of instructions for all conditions.)

In all three conditions, participants were also given three scenarios to incorporate into their essays. Two of these involved situations where the target (or the self, in the self condition) had social interactions that were somewhat ambiguous with respect to race. That is, the behaviors in these scenarios could have easily been interpreted as being driven by the target's negative attitudes toward blacks, or could have been interpreted neutrally. In the first scenario, participants were asked to imagine that the target (or self) approached a black clerk who was working in a hardware store for directions, but the clerk was busy helping another black customer and the target (or self) then walked away without getting the directions. In the second scenario, before leaving a large tip, the target (or self) joined in on a conversation with several employees of a restaurant who were

laughing about their black manager and talking about affirmative action. The behaviors in the third scenario did not have a racial component (and involved solitary behaviors rather than social interactions), and was therefore considered neutral with respect to racial attitudes (i.e., the target (or self) grabs a book from a table at home and goes for an evening walk, and then sits on a bench reading). These scenarios are described more fully in Appendix A.

Following the essay writing task, participants completed a series of measures in the following order: a behavioral measure of social distance, followed by a measure of affective state and set of implicit measures designed to assess prejudice and stereotyping of African Americans. This was followed by questions about the target and the essay writing task, explicit measures of motivation to respond without prejudice, and explicit measures of prejudice. Next, participants answered questions that asked them to explain the behavior of the target (or the self) during the scenarios described above and then completed questions that assessed their memory for the essays they wrote. Finally, participants were probed for suspicion. Although most participants were aware that each of the studies components had something to do with race and racial attitudes, none reported that they thought the essay task influenced subsequent responses.

Dependent Variables

Behavioral Measure of Social Distance. Following research by Macrae et al. (1994), after the primary manipulation participants were asked to take a seat in a row of 7 chairs while the experimenter prepared materials for the next part of the study. The last chair in the row had a backpack on it, which experimenters suggested belonged to an

African American participant who was doing the same study. (Experimenters asked participants if they had seen another participant, described as a "tall black male wearing a hooded sweatshirt," who was also doing the study, then mentioned that "he must have stepped out of the room for a moment" and would "be right back" because his backpack was still there.) After directing the participant to take a seat, the experimenter briefly left the room to gather materials, then came back, noted the chair that the participant had chosen (values ranged from 1 = the chair closest to the backpack to 6 = the chair furthest to the backpack), and had the participant follow her to the next part of the study. Thus, higher numbers reflect participants sitting at a further distance from the chair that was ostensibly being used by the African American participant, and indicate greater social distance. Means and *SD*s for this variable and all other variables are reported in the Results section, below.

Implicit Associations Test. The *Implicit Associations Test* (IAT) is a computer task that measures the relative strength of associations between pairs of concepts (for a complete description of the IAT, see Greenwald, McGhee, & Schwartz, 1998). In the present case, all participants completed two IATs. One of these IATs (*prejudice IAT*) measured associations of African American (black) names (e.g., Jamal) and European American (white) names (e.g., Todd) with pleasant and unpleasant words (e.g., sunshine, filth). The second (*stereotype*) IAT measured associations of black and white names with evaluative stereotypes of blacks and whites: Black stereotypes were negatively valenced stereotypes of black American men (e.g., hostile) and white stereotypes were positively valenced stereotypes of white American men (e.g., successful).

For the stereotype IAT, a first block had participants classify names as either African American or European American using left ("A" key) and right ("L" key) hand key presses on a standard keyboard. In a second block, participants used the same keys to distinguish between positive and negative traits. In the third and fourth blocks (combined blocks), participants used the same key to classify black names *and* negative stereotypes, using the other key to classify white names and positive stereotypes. In a fifth block, participants practiced classifying names again, with the categories reversed (e.g., if black names were originally classified using the "A" key and white names using the "L" key, then white names would be classified using the "A" key and black names with the "L" key). The final two blocks were combined blocks again, with the pairings reversed from those of the third and fourth blocks. That is, if black names were classified using the same key as negative stereotypes in the earlier blocks, then white names and negative stereotypes were classified with the same key in the final blocks. The prejudice IAT used the same format, only substituting positive and negative words for traits.

Which IAT was completed first was counterbalanced (i.e., half of participants completed the stereotype IAT first followed by the prejudice IAT; the other half completed the prejudice IAT followed by the stereotype IAT). In one version of each IAT, black names were first paired with negative stereotypes (or unpleasant words); in the other version, black names were first paired with positive stereotypes (or pleasant words), and whether participants completed the first or second version was also counterbalanced.

IAT scores were computed using the *D* statistic (see Greenwald, Nosek, & Banaji, 2003, for a detailed description of this scoring algorithm and discussion of why this statistic is preferable to previous scoring algorithms). This method uses untransformed response latencies from both practice and critical combined categorization blocks, eliminates trials with latencies > 10,000 ms, and eliminates subjects for whom more than 10% of trials have latencies < 300 ms (no participants met this criterion). Means are then computed for latencies to correct responses within each block, as are separate pooled *SDs* for practice and critical blocks. Average latencies for congruent blocks (i.e., white names paired with positive stereotypes or evaluatively positive words) were then subtracted from latencies for incongruent blocks (i.e., black names paired with positive stereotypes or evaluatively positive words) separately for practice blocks and critical blocks, and then practice and critical blocks were divided by their respective pooled *SD*s. These quotients (i.e., for practice and critical blocks) were then averaged. This final score comprises the *D* statistic.

Explicit Measures

In all cases, unless otherwise reported, scales were computed by averaging appropriate constituent items (reverse-scored where appropriate). All scales were additionally coded so that higher scores would reflect more of the construct in question (for example, higher numbers for the Positive and Negative Affect scales reflected higher positive and negative state affect, respectively). Scales that assessed participants' racial attitudes were coded so that higher scores would indicate more negative racial attitudes (e.g., greater racism). In cases where attitudes toward both whites and blacks were

measured (i.e., semantic differential and feeling thermometers), in addition to computing separate average indices measuring attitudes for whites and blacks, comparative measures were created that reflected differences in attitudes or opinions for whites and blacks (similar to the IAT; see Gawronski, Peters, Brochu, & Strack, 2008). In each case, these variables were scored so that higher numbers reflected more positive attitudes/assessments of whites relative to blacks. Appropriate measures of reliability are reported for each measure below, and means and *SD*s are reported in the Results section.

Positive and Negative Affect Scales (PANAS; Watson, Clark, & Tellegen, 1988). The PANAS consists of 20 words that describe feelings and emotional states, and is organized into two scales that measure global positive ($\alpha = .86$) and negative ($\alpha = .77$) affect. Examples of items from the positive affect scale are "interested" and "excited." Examples from the negative affect scale are "distressed" and "upset." For each item, participants were instructed to "Indicate to what extent you feel this way right now, that is, at the present moment." Each scale consists of 10 items, measured on a scale where 1 = "very slightly or not at all," 2 = "a little," 3 = "moderately," 4 = "quite a bit," and 5 = "extremely." Higher numbers therefore reflect a higher level of affect.

Internal and External Motivation to Respond Without Prejudice (IM and EM; Plant & Devine, 1998). The IM scale ($\alpha = .83$) concerns internal, personal reasons that people have for not acting prejudiced. An example item is "I attempt to act in nonprejudiced ways toward Black people because it is personally important to me." The EM scale ($\alpha = .80$) concerns external, social reasons that people have for not appearing prejudiced. An example item is "I attempt to appear nonprejudiced toward Black people

in order to avoid disapproval from others." Each scale has 5 items and all items are on a 9-point scale, where 1 = "Strongly disagree" and 9 = "Strongly agree."

*Old-Fashioned and Modern Racism (*McConahay, 1986). These scales each consist of seven items intended to tap into prejudiced attitudes toward African Americans. The Old-Fashioned scale (α = .88) uses items which are thought to reflect the ways that prejudice was expressed in the past, but which are no longer acceptable attitudes to endorse. For example, "Black people are generally not as smart as whites." The Modern scale (α=.63), on the other hand, was designed to more subtly tap into explicit but ambivalent attitudes toward African Americans in a way that respondents would be more likely to endorse, using items such as "Discrimination against blacks is no longer a problem in the United States." All items are on a 7-point scale where 1 = "Strongly disagree" and 7 = "Strongly agree."

Attitudes toward Blacks (Brigham, 1993). Similar to the Old-Fashioned and Modern Prejudice Scales described above, the 20-item Attitudes toward Blacks scale (α = .76) attempts to capture respondents' attitudes toward African Americans. Sample items are "I think that black people look more similar to each other than white people do" and "It would not bother me if my new roommate was black." All items are on a 7-point scale where 1 = "Strongly disagree" and 7 = "Strongly agree."

Semantic Differential (e.g., McConnell & Liebold, 2001). Using a 7-point scale, where higher numbers indicate less favorable attitudes, participants were asked to rate their feelings toward blacks and whites along five dimensions (anchors were: "beautiful/ugly" [i.e., 1 = "beautiful;" 7 = "ugly"], "good/bad," "pleasant/unpleasant,"

"honest/dishonest," "nice/awful"). Reliabilities (alphas) for the white and black semantic differentials, respectively, were .86 and .89. An additional variable was computed that measured less favorable attitudes toward blacks relative to whites by subtracting scores for whites from those for blacks. Positive scores therefore reflect less favorable attitudes toward blacks relative to whites, negative scores reflect less favorable attitudes toward whites relative to blacks, and a score of "0" reflects a lack of difference in attitudes toward both groups.

Feeling Thermometer (e.g., McConnell & Liebold, 2001). Participants were presented with a picture of a thermometer that was labeled "Very warm" at 100^{o} and "Very cool" at 0^{o}, and instructed that "higher temperatures indicate more warmth (more favorable feelings) and lower temperatures indicate more coolness (less favorable feelings)." Participants were then asked to choose a number between 0-100 to indicate their feelings toward, among other groups serving as distracters, African Americans (blacks) and European Americans (whites). Thermometer scores were then reverse-coded so that a score of 0 = "maximum warmth" (positive feelings) and a score of 100 = "maximum coolness" (negative feelings). As was done for the semantic differential variables, an additional variable that measured coolness toward blacks relative to whites was also computed by subtracting the score for whites from the score for blacks. For this variable, a score of "0" indicates no relative preference for either racial category. Positive scores indicate greater coolness toward blacks relative to whites, and negative scores indicate the opposite.

Target and Essay Measures

Participants responded to several questions which assessed their feelings about the target, their similarity and perceived relationship with the target (i.e., self-other merging and "we-ness"), perceived racism of the target, strategies used in the essay writing task, and one question about the difficulty of the task. Dependent measures were created after examining correlations among related items, and in some cases, examining exploratory principle components analyses. These analyses (and other similar analyses) were not formal tests, and were conducted simply to identify whether items meant to comprise distinct measures clustered appropriately; exact results of these analyses are therefore not reported (although reliabilities and correlations are reported where appropriate). These measures included the following:

Regard for the Target. Participants were asked 4 questions about their positive regard for the target ($\alpha = 82$). All questions were asked on 7-point scales where higher numbers reflected higher agreement (1 = "not at all"; 7 = "very much"). The questions were: "How much do you like the person in the photograph?" "To what degree do you care about the person in the photograph?" "To what degree would you want to spend time with the person in the photograph?" "Does the person in the photograph seem like someone who could be a close friend?"

Similarity to and Merging with the Target. Participants were asked one question that assessed their perceived similarity to the target, and were also asked to respond to two measures of self-target merging. All questions were asked on 7-point scales where higher numbers reflected higher agreement. The similarity question was: "To what degree

are you and the person in the photograph similar?" (1 = "not at all" and 7 = "very"). The first of the two merging measures was the *Inclusion of Others in Self scale* (IOS; Aron et al., 1992). This pictorial scale contains 7 pairs of circles representing "self" and "other." Anchored at "1" on one end, the two circles do not touch, representing lack of overlap. Each additional set of circles shows progressively greater overlap, with the set represented by "7" showing a substantial amount of overlap (i.e., higher numbers reflect greater self-other overlap). The second measure was adopted from Cialdini et al. (1997). Participants were asked to indicate on a 7-point scale (1 = "not at all"; 7 = "extremely") the extent to which they would use the term "we" to characterize themselves and the person in the photograph.

The IOS scale was highly correlated ($r = .71$) with the "we" measure, and the two were averaged to form a measure of perceiver-target merging ($\alpha = .84$). The "similar" item was correlated at .18 with the IOS, .31 with the "we" measure, and .26 with the aggregated "merging" measure, so it was retained as a single item measure of similarity.

Perceived Racism of the Target. Participants were asked two questions about their perception of the target's attitudes toward African Americans. Both questions were asked on a 7-point scale, where higher numbers reflected greater perceived racism. The first question was: "How racist do you think the person in the photograph is?" (1 = "Not at all racist," 7 = "Very racist"). The second question was: "Do you think the person in the photograph likes or dislikes African Americans?" (1 = Does not like at all," 7 = "Likes very much"). These items were aggregated ($\alpha = .75$, $r = .59$) and higher numbers reflect a greater perception that the target was racist.

Essay Writing Strategies and Difficulty of the Essay Writing Task. Participants were asked to rate the extent to which they used several different strategies when writing their essays, and were also asked questions about the target. (In the "self" condition, participants first "randomly" selected the same photograph seen earlier by participants in the other conditions, and were invited to spend a few minutes thinking about the target before answering these questions.) All questions were on 7-point scales where higher numbers reflected greater agreement (1 = "very little," 7 = "very much"). The questions were: "To what extent did you try to imagine how you would feel if you were the person in the photograph?" "To what extent did you try to imagine how the person in the photograph was feeling?" "How much effort did you spend trying to imagine what the person in the photograph was like?" "How much effort did you spend trying to figure out what the person in the photograph was thinking?" "How much effort did you spend trying to figure out how the person in the photograph would act during the events in the course of his day?" Participants were also asked to rate how difficult the essay writing task was on a 7-point scale, where 1 = "very difficult" and 7 = "very easy."

With the exception of the "difficulty" item, each of the other 5 items were highly correlated, and were thus aggregated ($\alpha = .87$) as a general measure of "effort spent imagining the target." The "difficulty" item was uncorrelated with each of these items (*r*s ranged from -.12 to .12, average $r = -.02$), and was retained as a separate measure of how difficult participants found the task.

Behavioral Explanations. Participants were asked to explain the behavior of the target (or the self) during the scenarios which were described above. That is, participants

were asked several open-ended questions relating to the content of their essays, based on the specified scenarios. Two of the scenarios (described above and in Appendix A) invited the possibility that the target's actions were motivated by negative attitudes toward blacks, but also allowed for reasoning that avoided issues of race. The target's actions in the third scenario were planned to be neutral with respect to race. To elicit explanations, participants were asked two questions each about the first and second scenarios and three questions about the third scenario. The salience of racial attitudes in scenarios was considered as a moderating variable.

The questions included: Scenario 1 – "While at the hardware store, why did the person in the photograph ["you" was substituted for "the person in the photograph" in the control condition] approach that particular clerk?" "Why did the person in the photograph walk away before getting directions?" Scenario 2 – "At the restaurant, why did the person in the photograph join in the conversation?" "At the restaurant, why did the person in the photograph leave a large tip?" Scenario 3 – "Why did the person in the photograph go for a walk that evening?" "Why did the person in the photograph bring a book with him?" "Why did the person in the photograph sit on the bench?"

Participants' explanations were coded for content by two trained coders who were blind to condition and hypotheses, using Malle's (1998/2009) *F.Ex* coding scheme. Specifically, the principle investigator and coders first collaborated in breaking participants' explanations into codable segments. This was done because given the free-response nature of the task, participants might provide multiple explanations for each question, but also might include information that was not relevant or asked for. After

identifying codable explanations, coders examined whether each explanation cited a reason, a causal history of a reason, or a mere cause (no mere causes were found, and this category is not discussed further). Reason explanations were further coded for whether participants cited a belief, desire, or a valuing (valuings are not considered further), and beliefs were further coded for presence or absence of a mental state marker (e.g., "He thought," "I felt"; for a discussion, see Malle et al., 2007). Causal histories were also further coded for whether they cited a background factor that resided in the person (e.g., a trait) or the situation. Causal histories that did not fit clearly into either of these two categories (e.g., person x situation interactions) were coded as "other." (The "other" category is also not considered further.) The principle investigator resolved all coding differences that arose.

Coders' responses to each of these categories were then aggregated within each question, each scenario, and also within the broader categories representing whether race was salient in the behavioral scenarios which were explained. Coder reliability was indexed by percent agreement and *kappa* statistics, and is reported in the Results section, below (no reliabilities are reported for number of explanations because the number of explanations generated for each participant for each question was decided through discussion among coders and the principle investigator).

Memory Questions. Participants were asked several questions about their memory for the essays they wrote earlier in the study. Responses were coded by two trained coders (blind to condition and hypotheses) who used the following criteria: 0 = "content listed is very different from essay content (i.e., they didn't remember);" 1 = "content

listed is somewhat similar to essay content;" 2 = "content listed very accurately reflects essay content." Coder reliability (r and α) are reported below.

All questions were prefaced by the instruction "Think back to the essay you wrote earlier." The memory questions were: 1) "Try to recall each of the events you were asked to write about in as much detail as possible. In other words, recall the full circumstances of the events - exactly as they were written in the instruction sheet given to you." 2) "Try to recall how the target reacted to each of the events you just described. Briefly, summarize how the target responded to each event." 3) "Try to recall as many words or phrases as possible that you used in this essay concerning race or ethnicity. In other words, if any part of your essay contained words or phrases relating to race or ethnicity, please list them below. List only those words or phrases you used in the essay." 4) "Try to recall as many words or phrases as possible that you used in this essay concerning the way the target felt (i.e., the target's feelings or emotions). In other words, if any part of your essay contained words or phrases relating to how the target was feeling, please list them below. List only those words or phrases that you used in the essay."

Responses were averaged across coders (Question 1 $\alpha = .85$, $r = .74$; Question 2 $\alpha = .76$, $r = .62$; Question 3 $\alpha = .85$, $r = .75$; Question 4 $\alpha = .82$, $r = .70$; across all questions $r = .73$, $\alpha = .84$). Perhaps not surprisingly given the vastly different nature of each of the questions, the 4 memory items were not very highly correlated (average $r = .15$, although memory question 1 correlated significantly at .30 and .29 with questions 2 and 3), and did not seem to be tapping "general memory" for the essay, but specific

memory for the content each question represented. Therefore, each memory question was individually retained.

Essay Content Analysis.

Participants' essays were coded along several dimensions (see Galinsky & Moskowitz, 2000, for a similar method) by two trained coders who were blind to condition and hypotheses. Codes were then averaged across coders (reliabilities are reported below after discussion of variable creation). Some of the coding involved counting particular types of words or stereotypes. Other coding was more subjective, such as the creativity of the essays. The categories are listed below:

Word/Phrase Counts. First, the number of neutral and negative racial category membership labels for African Americans was counted (e.g., black, nigger). Next, name calling, or referring to an African American in a negative way, was counted (e.g., "those jerks" [in reference to African Americans]). Words or phrases that reflected stereotypes of African Americans were also counted (taken from Devine, 1989). These included negative stereotypes such as "aggressive" or "unintelligent," positive stereotypes such as "athletic" or "musical," and neutral stereotypes such as preferences for certain types of food, physical characteristics, etc.

Target Stereotypicality. This code reflected the extent to which the author used stereotypes (versus some other means of imagination) to create their narrative of the target (Galinsky & Moskowitz, 2000). Codes were 0 = "negligible or no use of (target) stereotypes" 1 = "a small amount of stereotypes," 2 = "moderate use of stereotypes," and 3 = "high use of stereotypes/target is extremely stereotypical." For example, a target who

lived in the South and drove a pickup truck with a gun rack might be considered as a "1" while a target who lived in the South, drove a pickup truck with a gun rack, worked in a factory, and played "war games" with his friends in his spare time might be a "2." A "3" might show evidence of all of the above, but also live in a trailer and also belong to an "exclusive [i.e., no women, no minorities] men's club."

Target Prejudice. This code captured how prejudiced the target seemed to be as portrayed by the author. Codes were 0 = "not at all prejudiced/no evidence of prejudice," 1 = "some evidence of prejudice/mixed content," 2 = "fairly prejudiced," and 3 = "extremely prejudiced."

Target's Attitudes toward Blacks. This code is similar to the code above, except that it allowed for the potential expression of positive attitudes toward African Americans. The codes were 0 = "very positive," 1 = "somewhat positive," 2 = "neutral/mixed," 3 = "somewhat negative," and 4 = "very negative."

Essay Creativity. This code related to how creative or imaginative the essays were. In other worlds, it tried to capture how much effort participants put into creating a "world" for the target that was rich and full of detail. Codes were 0 = "not at all creative," 1 = "a little bit creative," 2 = "fairly creative," and 3 "extremely creative."

Examination of the frequency of count variables revealed that several of these counts did not vary significantly enough for further inclusion, with most participants' essays not showing any evidence of using the measured categories. These variables included negative labels, name calling, and neutral and positive stereotypes. Thus, these variables were excluded from further analyses. Similarly, target stereotypicality was

limited in variability, with over 87% of essays showing no evidence of target stereotypicality. (This latter finding might seem unlikely, given that the target's racism was perceived as relatively high, but coders were explicitly instructed to avoid using the target's racism as a criterion for this code, as this was clearly captured in other codes.) Of the remaining codes, the target's perceived prejudice (coder $\alpha = .95$, $r = .90$) and attitudes toward blacks (coder $\alpha = .88$, $r = .81$), and negative black stereotype counts (coder $\alpha = .90$, $r = .81$) were clearly related (inter-item rs ranged from .61 to .85, average $r = .71$). Because these variables were measured differently, responses to each were standardized and aggregated to form a measure of "Essay Racism Content" (inter-item $\alpha = 88$). Relations among the remaining codes were weak or inconsistent, so "References to Race" (coder $\alpha = .99$, $r = .98$) and "Creativity" (coder $\alpha = .84$, $r = .73$) were retained as individual variables.

Results

Descriptive Statistics

Below, descriptive statistics for all variables are provided. Table 1 provides descriptive statistics for the indirect measure of social distance (seating measure), Implicit Associations Tests (IATs), explicit variables (e.g., measures of racial attitudes), perceptions of the target and the essay writing task, and codes derived from participants' essays. Reliabilities are also provided where appropriate, as are possible score ranges for each variable.

Table 1.
Descriptive Statistics for Implicit and Explicit Measures – Including Perceptions of the Target, Perceptions of the Essay Writing Task, and Codes Derived from Participants' Essays (Including Memory for the Essay)

Variable	α	*M (SD)*	Possible Ranges
Seating Measure	-	3.49 (1.50)	1 to 6
Prejudice IAT	-	.29 (.33)	-
Stereotype IAT	-	.28 (.31)	-
Positive Affect	.86	2.47 (.69)	1 to 5
Negative Affect	.77	1.41 (.44)	1 to 5
Internal Motivation to Respond without Prejudice	.83	7.29 (1.55)	1 to 9
External Motivation to Respond without Prejudice	.80	5.20 (1.82)	1 to 9
Old-Fashioned Racism	.88	1.90 (.84)	1 to 7
Modern Racism	.63	2.36 (.99)	1 to 7
Attitudes Toward Blacks	.76	2.25 (.84)	1 to 7
Semantic Differential (Whites)	.86	3.06 (.84)	1 to 7
Semantic Differential (Blacks)	.89	3.05 (.96)	1 to 7
Semantic Differential (Blacks - Whites)	-	.23 (.86)	-7 to 7
Feeling Thermometer (Whites)	-	25.15 (18.29)	0 to 100
Feeling Thermometer (Blacks)	-	31.32 (18.44)	0 to 100
Feeling Thermometer (Blacks - Whites)	-	6.18 (18.00)	-100 to 100
Positive Regard for the Target	.82	1.84 (.93)	1 to 7
Perceived Similarity to Target	-	1.73 (1.15)	1 to 7
Perceiver-Target Merging	.84	1.58 (.91)	1 to 7
Perception of Target as Racist	.75	6.21 (1.02)	1 to 7
Effort Expended During Essay Task	.87	4.16 (1.30)	1 to 7
Difficulty of the Essay Writing Task	-	4.12 (1.62)	1 to 7
Essay Racism Content	.88	.00 (.90)	-
Essay References to Race	-	3.11 (2.01)	-
Essay Creativity	-	1.07 (.98)	0 to 3
Essay Memory Question 1	-	1.80 (.45)	0 to 2
Essay Memory Question 2	-	1.73 (.57)	0 to 2
Essay Memory Question 3	-	1.61 (.63)	0 to 2
Essay Memory Question 4	-	.95 (.74)	0 to 2

Table 2 (below) provides statistics for behavioral explanations, and includes information on coder reliability (*kappa* and percent agreement). In this table, *n*s refer to the total number of each explanation type within each question and scenario, and across all questions. Next to this, the percent symbol (%) indicates the percentage that the *n* represents out of the total appropriate category. That is, because explanations could cite

either reasons or CHRs, the (%) values next to the reason and CHR *n*s represent the percentages out of the total number of explanations that reasons or CHRs were cited. Similarly, the (%) values for beliefs and desires are out of the total number of reasons (because valuings are not considered here, the percentages do not always equal 100), the "marked" category is out of the total number of beliefs, and the person and situation categories are out of the total number of CHRs (because some CHRs were coded as "other," percentages do not always equal 100). Next to the *kappa* statistics, the percent symbol (%) represents coders percent agreement.

Codes are number of explanations (EXPL), reasons (REASON), and causal histories of reasons (CHR). Within reason categories, there were beliefs (BELIEF), which could be marked (B-MARK) or unmarked, and desires (DESIRE). For causal histories of reasons, there could be factors residing in the person (C-PERS), or situation (C-SIT).

Table 2.
Descriptive Statistics for Behavioral Explanations

Variable	κ (%)	*n* (%)	*M (SD)*	Variable	κ (%)	*n* (%)	*M (SD)*
Question 1				Question 2			
EXPL		138	1.30 (.59)	EXPL		160	1.54 (.64)
REASON	.91 (97)	110 (80)	1.06 (.65)	REASON	.98 (99)	119 (74)	1.15 (.71)
CHR	.95 (99)	24 (17)	.23 (.44)	CHR	.98 (99)	40 (25)	.39 (.60)
BELIEF	.94 (97)	75 (68)	.73 (.66)	BELIEF	.89 (94)	69 (58)	.67 (.66)
B-MARK	.98 (99)	12 (16)	.12 (.32)	B-MARK	.79 (89)	39 (57)	.28 (.54)
DESIRE	.98 (99)	35 (32)	.34 (.52)	DESIRE	.92 (97)	42 (35)	.41 (.55)
C-PER	.66 (99)	1 (4)	.01 (.10)	C-PER	.98 (99)	37 (93)	.26 (.59)
C-SIT	.76 (99)	5 (21)	.05 (.22)	C-SIT	(100)	3 (7)	.03 (.17)

Table 2 (continued).

Variable	κ (%)	*n* (%)	*M (SD)*	Variable	κ (%)	*n* (%)	*M (SD)*
Question 3				Question 4			
EXPL		165	1.60 (.68)	EXPL		157	1.52 (.70)
REASON	.81 (91)	110 (67)	1.07 (.72)	REASON	.80 (94)	136 (87)	1.32 (.74)
CHR	.81 (91)	55 (33)	.53 (.67)	CHR	.75 (93)	21 (13)	.20 (.47)
BELIEF	.90 (96)	50 (45)	.49 (.58)	BELIEF	.91 (97)	78 (57)	.76 (.72)
B-MARK	.84 (92)	20 (40)	.19 (.42)	B-MARK	.81 (91)	19 (24)	.18 (.46)
DESIRE	.93 (97)	36 (32)	.35 (.54)	DESIRE	.90 (98)	27 (20)	.26 (.52)
C-PER	.76 (92)	30 (55)	.29 (.48)	C-PER	.75 (94)	19 (90)	.18 (.46)
C-SIT	.81 (95)	22 (40)	.21 (.48)	C-SIT	.27 (97)	1 (5)	.01 (.10)
Question 5				Question 6			
EXPL		145	1.42 (.66)	EXPL		113	1.11 (.44)
REASON	.67 (93)	134 (92)	1.30 (.71)	REASON	.76 (97)	110 (97)	1.07 (.45)
CHR	.61 (93)	11 (8)	.13 (.33)	CHR	.59 (97)	3 (3)	.03 (.17)
BELIEF	.64 (90)	29 (22)	.28 (.54)	BELIEF	.60 (86)	33 (30)	.32 (.53)
B-MARK	.88 (95)	8 (28)	.08 (.27)	B-MARK	.91 (96)	10 (30)	.10 (.33)
DESIRE	.88 (95)	99 (74)	.96 (.83)	DESIRE	.83 (91)	61 (55)	.59 (.55)
C-PER	.72 (97)	8 (73)	.08 (.27)	C-PER	-- (97)	1 (33)	.01 (.10)
C-SIT	.24 (96)	3 (27)	.04 (.19)	C-SIT	.49 (98)	1 (33)	.01 (.10)
Question 7				Total			
EXPL		140	1.36 (.67)	EXPL		1018	9.85 (2.31)
REASON	.71 (92)	132 (94)	1.28 (.71)	REASON	.80 (94)	851 (84)	8.26 (2.38)
CHR	.65 (91)	8 (6)	.08 (.27)	CHR	.79 (94)	162 (16)	1.59 (1.22)
BELIEF	.80 (94)	43 (33)	.42 (.59)	BELIEF	.85 (93)	377 (44)	6.52 (2.03)
B-MARK	.83 (93)	2 (5)	.02 (.14)	B-MARK	.84 (93)	110 (29)	1.07 (1.17)
DESIRE	.83 (92)	84 (64)	.82 (.76)	DESIRE	.91 (96)	384 (45)	3.73 (1.99)
C-PER	.78 (96)	2 (25)	.02 (.14)	C-PER	.76 (95)	98 (60)	.95 (1.02)
C-SIT	.32 (94)	5 (63)	.04 (.22)	C-SIT	.66 (97)	40 (25)	.40 (.65)
Total Race Salient				Total Race Neutral			
EXPL		620	5.97 (1.64)	EXPL		398	3.77 (1.16)
REASON	.83 (93)	475 (77)	4.61 (1.70)	REASON	.71 (94)	376 (94)	3.65 (1.24)
CHR	.83 (94)	140 (23)	1.36 (1.11)	CHR	.63 (93)	22 (6)	.23 (.29)
BELIEF	.91 (96)	272 (57)	2.64 (1.53)	BELIEF	.69 (90)	105 (28)	1.02 (1.10)
B-MARK	.82 (91)	90 (33)	.87 (1.00)	B-MARK	.88 (94)	20 (19)	.19 (.51)
DESIRE	.93 (98)	140 (30)	1.26 (1.24)	DESIRE	.85 (93)	244 (65)	2.37 (1.41)
C-PER	.78 (95)	87 (62)	.84 (.90)	C-PER	.70 (96)	11 (50)	.11 (.37)
C-SIT	.77 (96)	31 (22)	.30 (.57)	C-SIT	.32 (96)	9 (41)	.10 (.33)

Data Examination and Reduction

Before conducting primary analyses, two additional goals were considered. A first goal was simply to examine distributional characteristics for all variables. A second goal was to reduce the large number of explicit variables for further analyses to a more manageable number. Discussion of both of these issues follows.

Most variables appeared to be normally distributed, although the similarity and merging variables appeared to be strongly positively skewed. Further examination showed that for each of these variables, most participants indicated an absence of agreement with the variable (percentages of participants who indicated a complete lack of agreement with these variables ranged from 47% for merging to 56% for similarity, with another 25% of participants indicating the next lowest level of agreement for both variables, and relatively few participants indicating higher agreement, thus creating the strong skew). Because of this, these variables were dichotomized, with a "0" indicating lack of similarity or merging and a "1" indicating at least some agreement with each variable.

Several explicit variables also showed moderate positive skew. Additional analyses were thus conducted to explore whether some of these variables could safely be aggregated to normalize distributions, and to determine which variables if any might need to be transformed before inclusion in further analyses.

In a first step, correlations among all explicit variables that reflected racial attitudes were examined. This included established scales which measure attitudes towards blacks but not whites (e.g., Modern Racism Scale), scales or items where

responses were examined for both blacks and whites (e.g., semantic differential scales), and relative variables that were created from these latter scales that were meant to examine relative differences in attitudes toward blacks and whites (e.g., black - white thermometer ratings).

Two clear trends emerged: For variables where attitudes toward blacks but not whites were examined, strong positive correlations were found among all variables. For variables which examined attitudes toward both blacks and whites, there appeared to be multiple sources of variance, reflecting a general trend for participants to rate blacks and whites similarly (i.e., indicating more negative attitudes toward blacks was associated with more negative attitudes toward whites). This trend likely reflected impression management concerns and response sets, as well as actual attitudes toward both groups.

On the basis of these analyses, as well as an exploratory principle components analysis, the following decisions were made: first, because the established scales correlated so strongly (average r = .73), these were standardized and aggregated to form a single "Explicit Racism" scale (α = .89), which was normally distributed. Second, to explore measures that reflect relative rather than absolute explicit bias, the relative thermometer and semantic differential scales were retained and the absolute scales (which correlated only modestly with other measures of racism) were removed from further analyses.

Closer examination of the relative thermometer and semantic differential scales revealed that two influential cases were causing the skew (z-scores ranging from 2.5 to 5.3) on these items. Removal of these participants normalized distributions, so further

analyses using these two variables were tested with these participants excluded and included. Analyses did not differ based on their exclusion, so they were retained. Table 3 (below) lists the correlations among all retained race-related dependent variables, and also gives descriptive statistics for each.

Table 3.
Descriptive Statistics and Correlations for Final Explicit Measures

Variable	*M* (*SD*)	2.	3.	4.	5.	6.
1. Explicit Racism	.00 (.90)	.22*	.43**	.17	.15	.28**
2. Relative Thermometer (White - Black)	5.00 (15.73)		.20*	-.16	.16	.13
3. Semantic Differential (White - Black)	-.09 (.90)			-.09	.29*	.08
4. Seating Measure	3.51 (1.5)				-.07	.11
5. Prejudice IAT	.29 (.33)					.24*
6. Stereotype IAT	.28 (.31)					

* $p < .05$, ** $p < .01$

Primary Analyses

Analyses were first conducted on all potential control variables (positive and negative affect, effort spent during the essay writing task, perceived difficulty of the essay writing task, perceptions of the target as racist, and participants' memory for their essays). Following this, all mediators were examined (i.e., similarity, positive regard, and perceiver-target merging) for effects of condition (Hypothesis 1), and the relationships between these variables and participants' racial attitudes were explored. Next, analyses focused on the influence of condition on participants' essays (e.g., expressed racist content in the essays; Hypothesis 2) followed by tests of the primary hypothesis that condition would influence participants' racial attitudes (Hypothesis 3). After this, the

interaction between condition and motivation to respond without prejudice was explored. Finally, the effect of condition on behavioral explanations (Hypothesis 4) was examined.

In most cases, variables were tested using factorial analyses of variance, with participant sex examined as an additional factor (when no effects involving sex were present, this variable is not discussed further). For similarity and merging, which were dichotomized, χ^2 independence tests were used, and follow up tests examined correlations between these variables and participants' attitudes. A slightly different analytical strategy was pursued in the section on behavioral explanations, and this strategy is discussed in that section.

Potential Control Variables. To rule out whether participants' state positive and negative affect, effort spent trying to imagine a day in the target's life, perception of the essay task as difficult, perceptions of the target as racist, or memory (for the scenario instructions, behavior during the scenarios, race related words used during the essay, or the emotions of the target/self) might have influenced subsequent responding on primary dependent variables, each of these variables were tested in separate models. The only significant effect which emerged was for the variable which measured effort in trying to imagine the target and a day in his life. For effort, there was a main effect of condition, $F(2, 94) = 14.44$, $p < .001$, $\eta^2 = .24$ (all reported η^2 are "partial η^2"). Based on the means, it appears that people in the perspective taking (PT; $M = 4.40$, $SE = .21$) and third person (3P; $M = 4.57$, $SE = .23$) conditions put in similar amounts of effort, while those in the (write about) self condition (SELF; $M = 2.88$, $SE = .26$) put in significantly less

effort trying to understand the target than both other conditions, $ps < .001$. A significant sex effect also emerged, $F(2, 94) = 7.24$, $p = .008$, $\eta^2 = .07$, with women ($M = 4.31$, $SE = .13$) reporting spending more effort than men ($M = 3.59$, $SE = .23$).

Hypothesis 1 – Potential Mediators. In the event that the expected effects of perspective taking on participants' attitudes were found, perceived similarity to and positive regard for the target, as well as perceiver-target merging[1] were considered as potential mediators. An initial test thus examined the effects of condition on these variables.

Contrary to prediction, neither similarity nor positive regard was significantly impacted by condition. For merging, an unexpected effect emerged. In the PT condition, a smaller proportion of participants than would be expected by chance (13/35) indicated some merging with the target, while in the SELF condition, a larger proportion than expected indicated feeling merged with the target (23/34; in the 3P condition, approximately half of the participants [17/32] participants indicated feeling merged), $\chi^2 (2, N = 101) = 6.44$, $p = .04$.

Follow up analyses examined correlations between each of these variables and each of the measures of participants' racial attitudes (i.e., seating measure, prejudice and stereotype IATs, explicit racism composite, and relative semantic differential and

---

[1] It should be noted that when the terms "merged," "positive regard," and "similarity" are used here (and throughout this dissertation), these terms might actually be best represented by the phrases "less psychologically distanced," "less *negative* regard," and "less dissimilarity." That is because across all conditions, merging with, regard for, and perceived similarity to the target received uniformly low scores. On a scale where the lowest possible response was "1" and the highest possible "7," average responses to all of these items were less than 2. Therefore, use of these terms here and throughout is not meant to actually reflect high levels of merging, positive regard, and similarity. The reader is therefore cautioned to be aware that for each of these variables, participants provided very low scores.

thermometer variables). The only correlation that reached significance was a correlation between the dichotomized merging variable and the explicit racism composite, $r\,(100) = .24, p = .02$, indicating that a sense of feeling merged with the target was associated with somewhat higher explicit racism. Further analyses revealed that this correlation was significant in the PT condition ($r\,(35) = .38, p = .03$), but not in the 3P ($r\,(32) = .15, p = .41$) or SELF conditions ($r = .17, p = .35$), although r to z transformations showed that these correlations did not significantly differ. Thus, there is some hint that a sense of being at least somewhat merged with the target after perspective taking is predictive of more racist attitudes.

Hypothesis 2 – Essay Codes. Hypothesis 2 was that perspective takers would write essays with significantly more racist content than participants in other conditions. Other codes from participants' essays were also examined, such as use of neutral race-related words (e.g., "African American") – which would indicate that participants were organizing their essays around racial themes – and essay creativity, which should not differ across conditions.

As expected, a main effect of condition was found on racism in participants' essays, $F\,(2, 101) = 50.97, p < .001, \eta^2 = .50$. Follow up t-tests showed that participants in the PT condition ($M = .70$, $SE = .11$) wrote essays with significantly more racist content than participants in the 3P condition ($M = .11$, $SE = .11$), who in turn wrote significantly more racist essays than SELF participants ($M = -.82$, $SE = .11$; all $ps < .001$). For the counts of neutral references to race, condition was again significant, $F\,(2, 101) = 12.51, p < .001, \eta^2 = .20$, with PT ($M = 3.81$, $SE = .30$) and 3P ($M = 3.67$,

SE = .32) not differing from each other (p = .95), but both making significantly more references to race than SELF (*M* = 1.86, *SE* = .31; ps < .0005). For essay creativity, no significant effects emerged. Taken together, these results indicate that even though perspective takers and third person participants organized their essay themes equally along racial lines (i.e., no differences in references to race), perspective takers still evidenced greater racist content in their essays. This result could be interpreted as having occurred because perspective takers were simply doing what they were expected to do – portraying the target as a racist. However, another interpretation is that participants may have been trying to show their dislike for the target and for having to take his perspective by portraying him as extremely racist. In effect, they may have been trying to "punish" the target for his attitudes as part of reluctant compliance with being asked to "become" the target.

Hypothesis 3 – Participants' Racial Attitudes. Contrary to predictions, no direct effects of condition on participants' racial attitudes were found. This was true for the indirect measure of social distance (distance from the "black participants backpack"), both IAT measures, the explicit racism composite, and semantic differential and thermometer variables that assessed participants' attitudes toward blacks relative to whites. Reasons for why this might have occurred are discussed below, in the Discussion section.

Motivation to Respond Without Prejudice. Recall that an exploratory prediction was that condition might interact with motivation to respond without prejudice to predict participants' racial attitudes. To test this hypothesis, moderated multiple regression

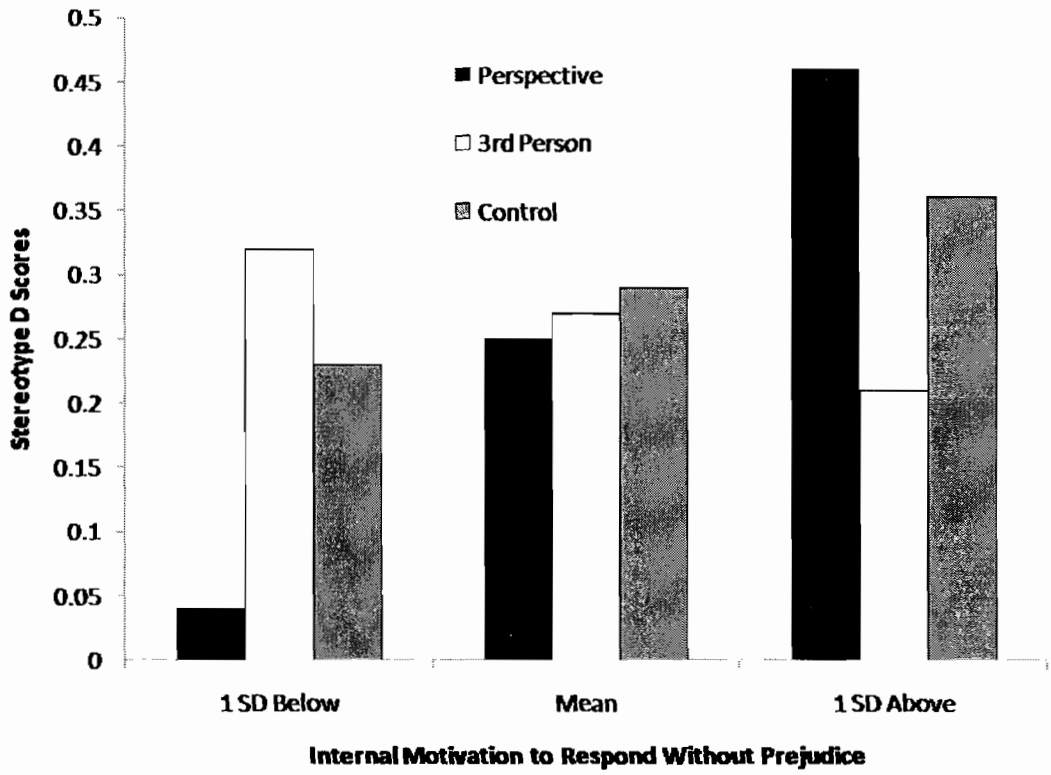
analyses were used to test each dependent variable (i.e., all measures of participants' racial attitudes), and if significant interactions were found, they were explored with simple effects tests where condition was treated as the focal independent variable (cf. Aiken & West, 1991). In this approach, two dummy codes were computed to test the perspective taking condition separately against the 3P and SELF conditions. Internal (IM) and external (EM) motivation to respond without prejudice variables were grand mean centered, and interaction terms were computed by multiplying the moderators by the dummy codes. These terms included all 2-way interactions (i.e., condition by IM, condition by EM, and IM by EM) and the 3-way interaction. (Sex was also tested as a factor, including all possible interactions with sex, but because no effects involving sex were found, this variable is not discussed further.) Variables were entered in blocks (i.e., dummy codes and the centered moderators were entered in a first step, followed by all 2-way interactions in a second block, and 3-way interactions in a third block), and significance of interactions was assessed by changes in R^2.

Although results of several of these tests were suggestive, the only effect that reached significance was an interaction between IM and condition, $F\Delta$ (2, 94) = 7.27, p = .001, $R^2\Delta$ = .13 (see Figure 1, below). Simple effects revealed that at the mean for IM, there were no differences in stereotyping D scores between PT (M = .25, SE = .05), 3P (M = .27, SE = .05), and control (M = .29, SE = .05) conditions. At 1 SD below the mean on IM (i.e., participants who were less motivated to respond without prejudice), however, PT (M = .04, SE = .08) had lower IAT D scores (i.e., less implicit negative stereotyping of blacks relative to whites) than 3P (M = .32, SE = .06) and SELF (M = .23,

SE = .09). At 1 *SD* above the mean on IM, this effect reversed, with PT (*M* = .46, *SE* = .07) showing greater implicit stereotyping of blacks than 3P (*M* = .21, *SE* = .08) and SELF (*M* = .36, *SE* = .07). No differences were present between 3P and SELF participants in any analysis. Thus, one way of conceptualizing this interaction is that IM was a significant predictor of stereotype *D* scores only in the PT condition (B = .14, β = .68, $p < .0005$). In the 3P (B = -.04, $p = .21$) and SELF (B = .04, $p = .25$) conditions, IM had no significant effect.

Thus, some support was found for the hypothesis that among perspective taking participants, those who were most concerned (in this case, for internal reasons) with appearing prejudiced were those who automatically stereotyped blacks the most and those who were least concerned with appearing prejudiced automatically stereotyped blacks the least. Importantly, this effect was only found for perspective takers, who, if they were concerned about being prejudiced, might have tried to suppress their thoughts of black stereotypes after imagining themselves *as* the target in order to distance themselves from the target and his attitudes – and in so doing, actually had greater stereotypical associations of blacks than perspective takers who were not concerned with appearing prejudiced.

Figure 1
Interaction of Condition with Internal Motivation to Respond without Prejudice on Stereotype IAT Scores.



Behavioral Explanations. As discussed above, participants were instructed to incorporate three different scenarios into their essays. Two of these scenarios had a potentially negative interpretation with regard to racial attitudes (i.e., the target [or self] might have acted the way he or she did because of his or her racial attitudes or might have had other reasons for acting in that way), while the third scenario was, at least in theory, neutral with regard to race. Participants were then asked a total of four questions about the two scenarios where race was potentially salient, and three questions about the

neutral scenario. All questions asked participants to describe in their own words why the target (or self) had acted the way he or she did.

Several options for data analysis were available, given the structure of the data. Because the major question of interest here is not participants' attitudes toward African Americans, but whether taking the psychological perspective of a target caused participants to explain the target's behaviors in ways that were more similar to the way participants explained their own behavior, the following strategy was used. Code counts (e.g., number of explanations) were aggregated within participants and individual questions and then averaged separately across race-related scenarios and race-neutral scenarios. These averages were then treated as within-person levels of a repeated-measures factor, and tested using mixed factorial analyses of variance where condition was a between-subjects factor and salience of race in the scenario was a within-subjects factor. Although other data analysis strategies were possible, this analysis seemed relatively straightforward, and allowed parsimonious tests of major hypotheses. Sex was also examined as an additional factor. When no effects (main effects or interactions) involving sex were present, sex was removed and was not considered further. (Because of this, *df* differ across some analyses.)

For number of explanations generated, there was a significant effect of salience of race, with more explanations generated for the race-related scenarios ($M = 1.47$, $SE = .05$) than for the neutral scenario ($M = 1.25$, $SE = .04$; $F(1, 91) = 16.80$, $p < .001$, $\eta^2 = .16$), as well as a significant effect for sex, which showed that women ($M = 1.44$, $SE = .03$) generated more explanations than men ($M = 1.27$, $SE = .06$; $F(1, 91) = 6.76$,

$p = .01$, $\eta^2 = .07$). A marginally significant effect for condition was also present, $F(2, 91) = 2.69$, $p = .07$, $\eta^2 = .06$ ($M_{PT} = 1.35$; $M_{3P} = 1.46$, $M_{SELF} = 1.26$, *SE*s, respectively = .05, .06, .07), showing that participants in the 3P condition generated the most explanations and the self condition the fewest, with perspective takers between the two.

For reason explanations, the only significant effect that emerged was for sex, $F(1, 91) = 5.74$, $p = .02$, $\eta^2 = .06$, with women ($M = 1.24$, $SE = .04$) providing more reasons than men ($M = 1.05$, $SE = .07$). This finding is not surprising given the correlation between number of explanations and number of reasons, $r(101) = .84$, $p < .001$.

For CHRs, a similar pattern emerged as for overall explanations, even though the correlation between CHRs and total explanations generated was small, $r(101) = .18$, $p = .08$. That is, a main effect for salience of race emerged, $F(1, 98) = 66.76$, $p < .001$, $\eta^2 = .41$, with more CHRs generated for the race-related scenarios ($M = .34$, $SE = .03$) than for non race related scenarios ($M = .08$, $SE = .02$).

Turning to the belief explanations category, there was a main effect of salience of race for beliefs, $F(1, 98) = 45.38$, $p < .001$, $\eta^2 = .32$, with a higher number of belief explanations generated for the race salient scenarios ($M = .67$, $SE = .04$) than for the neutral scenarios ($M = .35$, $SE = .04$). An interaction between race-salience and condition was also found, $F(2, 98) = 3.13$, $p < .05$, $\eta^2 = .06$. In the race-salient scenarios, belief explanations did not differ much as a function of condition ($M_{PT} = .69$, $M_{3P} = .63$,

M_{SELF} = .68, all *SE*s = .07, $F < 1, p > .76$). In the race-neutral scenario, however, the SELF (M = .52, SE = .06) condition generated more belief explanations than the other PT (M = .24, SE = .06) and 3P (M = .28, SE = .06).

Several options were available to test whether the presence or absence of mental state markers for belief statements differed across conditions. One option would be to use counts or averages of counts across questions, but an issue with this is that markers are confounded with the presence or absence of beliefs. That is, participants who made fewer (or no) belief explanations will automatically have fewer marked beliefs. A better option is to use the proportion (out of total belief statements) that also had mental state markers. One issue with this approach, however, is that sample size will decrease by the number of participants who have did not use at least one belief explanation for race-salient questions and also at least one for race-neutral questions. To deal with this, three proportions were created. First, the proportion of marked beliefs out of total beliefs within race-salient and race-neutral scenarios was computed, and tested using a strategy similar to the strategy used above in order to examine whether condition and the salience of race had an impact on use of marked beliefs. A total proportion of marked beliefs to total beliefs was also computed across *all* scenarios, and tested in separate between-subjects models.

When salience of race was considered as a within-subjects factor, two effects were significant for marked beliefs. These included a main effect for salience of race, F (1, 61) = 9.52, p = .003, η^2 = .13, with a higher proportion of marked beliefs for race-salient (M = .37, SE = .04) than for race-neutral scenarios (M = .19, SE = .04), and a main effect for condition, F (2, 61) = 4.16, p = .02, η^2 = .12. Follow-up tests revealed a

significant difference between 3P (M = .39, SE = .06) and SELF (M = .18, SE = .05; p = .02) conditions, with the PT condition (M = .28, SE = .04) not significantly differing from either of these (ps > .31). No other effects were found.

Collapsing across salience of race (in effect, the same analysis as above, but with greater degrees of freedom), a main effect was again found for condition, F (2, 98) = 3.87, p = .02, η^2 = .07. Participants in the 3P condition (M = .20, SE = .03) again differed from the SELF condition (M = .10, SE = .03; p = .02), with PT not differing from either (M = .17, SE = .03, ps > .15).

Turning to desire reasons, the only significant effects were for salience of race (F (1, 91) = 50.09, p < .001, η^2 = .36) and sex, F (1, 91) = 4.61, p = .03, η^2 = .05. Overall, there were more desires cited when race was not salient (M = .76, SE = .05) than when it was (M = .31, SE = .04), and women (M = .61, SE = .03) cited more desires than men (M = .46, SE = .06).

The last two codes under investigation concern whether participants' CHRs more often cited factors residing in the person (the target or the self) or in the situation. Similar to the investigation of marked beliefs above, a first step examined the proportion of CHRs that cited person or situation factors within race-salient and race-neutral scenarios. One issue that quickly emerged was that the number of participants who had cited at least one CHR for both race-salient and race-neutral scenarios was very low, mostly because participants cited significantly fewer CHRs across the race-neutral questions (total race-neutral CHRs = 22, M = .23) compared to the race-salient scenarios (n = 140, M = 1.35; p < .001). Therefore, analyses of person and situation factors were collapsed

across the salience of race factor, although it should be kept in mind that the reported tests are weighted heavily toward race-relevant scenarios. Also, it should be noted that because some participants did not cite any CHRs across any of the 3 scenarios, *df* are fewer in these analyses.

A main effect for condition was found for person factors, with SELF participants ($M = .44$, $SE = .07$) citing fewer person factors than PT ($M = .69$, $SE = .08$) or 3P ($M = .65$, $SE = .07$) participants, $F(2, 76) = 3.12$, $p = .05$, $\eta^2 = .08$, although post hoc tests again did not show any significant differences among conditions.

For situational factors, a main effect was again found for condition, $F(2, 79) = 3.71$, $p = .03$, $\eta^2 = .09$. Post hoc tests showed that SELF participants provided more situational factors than PT participants, $p = .02$, but that there were no differences between PT and 3P or between 3P and SELF ($M_{SELF} = .43$, $M_{PT} = .15$, $M_{3P} = .26$, all SEs = .07). Table 4 (below) provides a descriptive summary of raw means for all codes, broken down by condition and by whether race was salient in the scenarios participants were asked to explain.

Discussion

Study 1 set out to test the hypothesis that participants who took the perspective of a racist target – learning about his attitudes from visual information alone (e.g., Galinsky & Moskowitz, 2000) – and wrote an essay about the day in the target's life *as if they were that person* would come to express attitudes that were in line with those the target might

Table 4.
Raw Means of Behavioral Explanation Codes across Condition

Behavioral Explanation Code	Perspective Taking			Self			Third Person		
	Race	Neutral	Total	Race	Neutral	Total	Race	Neutral	Total
Number *†	1.51	1.23	1.39	1.38	1.37	1.38	1.63	1.33	1.5
Reason	1.23	1.16	1.2	1.04	1.29	1.15	1.23	1.24	1.23
Belief *‡	.69	.24	.92	.68	.52	.98	.63	.28	.93
Marked *†	.26	.03	.16	.13	.08	.11	.27	.09	.19
Desire *	.36	.86	.57	.29	.66	.45	.38	.91	.61
CHR *	.29	.07	.19	.35	.07	.23	.40	.10	.27
Person †	.19	.05	.12	.18	.01	.09	.27	.05	.16
Situation †	.04	.01	.02	.12	.04	.08	.08	.05	.07

Note: Codes marked with (*) showed significant main effects for race-salience at $p < .10$. Codes marked with (†) showed significant main effects for condition $p < .10$. Codes marked with (‡) showed interactions between race-salience and condition at $p < .10$.

be expected to express (i.e., more stereotyping and racism). To test this hypothesis, participants were provided with a photograph of a white male wearing a "white pride" t-shirt, and then wrote about a day in the life of this target from a first-person (perspective taking condition) or a third-person (third-person condition) perspective, or wrote about a fictional day in their own life (self condition). In the self condition, participants later viewed the same target and provided their impressions of him.

As part of these essays, and important for understanding the results of Study 1, participants were also asked to incorporate three different scenarios into their essays (described in the Method section and in Appendix A). The inclusion of these three scenarios was important not only because participants later explained the target's (or their own) behavior during these scenarios (to address hypotheses about the ways people explain a target's behavior after perspective taking); but they were also important in that

race could become a salient issue even for those participants who did not take the perspective of a racist target.

A variety of methods were used to tap into participants' state of mind both during and after the essay task. Essays were coded along several dimensions that included descriptions of the target (or the self) acting in racist ways, use of race-related words, and use of African American stereotypes. Participants also answered questions about the target, such as how racist he was, how similar they felt to him, how psychologically merged they were with him, and how much they liked him or would consider him someone that could be a close friend. Other variables tapped into participants' internal and external motivations to not be prejudiced, their affective state after the task, and their memory for their essays and for the essay instructions. Several methods were also used to assess participants' attitudes toward blacks.

Only limited support was found for hypotheses. For the first hypothesis (i.e., Hypothesis 1) – which was that perspective takers would regard the target more positively, perceive greater similarity to him, and merge with him to a greater extent than participants in other conditions – no support was found. In fact, in addition to the lack of effect of condition on similarity and positive regard, fewer participants than would have been expected by chance felt merged with the target in the perspective taking condition, while a greater number than expected by chance felt merged in the self condition.

Despite this, it was discovered that merging was correlated with greater explicit racism, and further analyses revealed that this correlation was significant only in the perspective taking condition. Although no evidence was found that merging was

interacting significantly with condition to predict explicit racism – in fact, a test of the difference between correlations was not significant – it does provide limited and descriptive evidence that the overall relationship between merging and participants' racial attitudes appeared to be somewhat stronger for perspective takers.

The second hypothesis, which predicted that perspective takers would write essays with more racist content than participants in other conditions, was supported. Participants in the perspective taking condition wrote the most racist essays, followed by those in the third-person condition, and then followed by those in the self condition. This was true even though the third-person condition made as many references to race as perspective takers (both conditions made more references to race than those in the self condition), showing that participants in both perspective taking and third-person conditions clearly thought that race was salient, and included content that highlighted their focus on race as an organizing theme.

This demonstrates a qualitative difference between taking the perspective of a racist target and thinking about this target from an outside perspective. That is, the target was the same in both conditions, as were the potentially racially charged scenarios, and in neither condition were participants explicitly instructed to make the target seem racist. The result, however, was different. Being asked to write about a day in the life of the target as if they actually *were* the target versus being asked to write about the target from an outside perspective resulted in perspective takers creating more racist content – in their roles *as* the target. It should be kept in mind that part of this effect might have simply reflected a deeper engagement with the target on the part of perspective takers,

and a stronger commitment to translating the little information they had about the target into a cohesive narrative. Even if this is the case, it does not diminish the fact that in accessing and expressing greater racist content on behalf of the target, perspective takers might have also had more information brought to mind that was inconsistent with egalitarian goals. Future work might further explore the consequences of writing essays with greater racist content. Perhaps they include behavioral consequences in interracial interactions or effects on measures of racial attitudes which are more indirect than those used here.

When it came to taking on the attitudes of the target as their own (Hypothesis 3), no support was found – at least in terms of main effects of condition on racial attitudes. One possibility (aside from the consideration that taking the perspective of a racist target might not affect participants' own attitudes) is that this represents a floor effect, in that there was uniformly low explicit racism across all conditions. That is, explicit measures are relatively easily controlled, and when presented with questions about their own racial attitudes, participants may have simply and uniformly refused to admit to (or did not feel) any racial prejudice. For more indirect measures, such as the measure of participants' seating distance from the "black participant's" backpack or the Implicit Associations Tests (IATs), a potential floor effect, while possible, is certainly not as obvious.

For the former, one possibility is that participants might not have paid attention to the experimenter's claim that the backpack belonged to an African American participant who would "be right back," or might not have thought that they would have to encounter this participant. In fact, when asked during debriefing "Did you think that you would

encounter another participant during the study?" several participants answered "No." Furthermore, some participants noted during debriefing that they thought it was strange that the experimenter mentioned the race of the other participant. Thus, the seating measure might not have actually been a measure of "social distance from an African American," or at best, was a very noisy measure of social distance. Considering that participants were never explicitly told that they would be interacting with this participant (as in Macrae et al., 1994), this seems even more likely. Some participants might not have realized that there was "another participant" and others might have been suspicious when the experimenter mentioned that he was black. Others might have been aware of this participant and his race, and yet still never thought that they would encounter him.

For the IATs, such an explanation is not possible. What is possible is that simply thinking about race or racial issues (recall that participants in the self condition were also asked to include scenarios in their essays that specifically mentioned race and racial issues such as affirmative action) primed automatic associations of blacks with negative stereotypes and unpleasant words equally for all conditions. Still, for at least one of these measures (stereotype IAT), there was evidence that perspective taking had a different effect than other conditions. That is, internal motivation to respond without prejudice had an effect on perspective takers that it did not have on other conditions.

One likely possibility is that after writing essays that included quite racist content (on behalf of the target), perspective takers – especially those with the most motivation to avoid appearing prejudiced – may have tried to banish thoughts of racism from mind, *suppressing* these thoughts in order to distance themselves from a distasteful task and

reassert their egalitarian goals. Thus, the desire and "permission" to express racist thoughts during the essay task might have resulted in higher automatic stereotyping because of post-task suppression, in much the same way documented by Macrae et al. (1994) and others (e.g., Galinsky & Moskowitz, 2000; Wyer, 2007; Wyer et al., 1998).

Why this effect did not extend to the prejudice IAT is uncertain, but speculatively, it may have to do with the possibility that while some participants actually associate blacks with the unpleasant words used in this measure (e.g., poison, filth) to a greater extent than they associate whites with these words, many do not, and the essay writing task did not really strengthen these associations. Another possibility is that the sometimes large measurement error associated with the IAT (i.e., low reliability) was also present here, making the effect, if present, very difficult to find.

For Hypothesis 4 – that perspective taking would lead participants to explain the target's attitudes in ways that are more similar to the ways participants in the self condition explained their own behaviors – very little support was found. Specifically, it was predicted that perspective taking (versus writing about the target in the third-person) would lead to greater use of reason explanations, of which more would be beliefs (which would in turn more often be unmarked), to less use of causal histories of reasons (CHRs), and when CHRs were provided, to a greater proportion of them that would cite situation rather than person factors.

In fact, only two analyses revealed any support for this hypothesis. The first of these was merely suggestive, and was represented in an interaction between condition and race salience on belief explanations. For the race-salient scenario, perspective takers

used more belief explanations than third-person participants and slightly more than participants who wrote about themselves. For the race-neutral scenario, a clear preference for belief explanations was evidenced for participants who wrote about themselves, with fewer belief explanations for both perspective takers and third-person participants.

The second significant effect, which was perhaps more clear, was that in two different analyses (one analysis took the relevance of race in the scenarios into account, and another collapsed across the relevance of race), perspective takers provided a number of marked beliefs that was somewhat greater than participants in the self condition but somewhat fewer than participants in the third-person condition. That is, while a main effect for condition was found for marked beliefs, and follow-up tests revealed that self and third-person conditions were significantly different from each other, perspective taking did not differ significantly from either of the other conditions. Examining the means for marked beliefs in the analysis that considered the salience of race, the linear effect can be clearly seen. In the self condition ($M = .18$), very few marked beliefs were provided. Perspective takers used slightly more marked beliefs ($M = .28$), and third-person participants used the greatest number ($M = .39$).

Although this effect for marked beliefs was found in relative isolation – meaning that no other strongly supportive effects of perspective taking on behavioral explanations were found – it does suggest that perspective takers may have tried (or some may have tried) to make the target appear more rational and less self-centered. Perhaps because the target was so clearly distasteful, however, they were not able to take this far enough. This explanation is consistent with Malle et al. (2000), who showed that participants who were

trying to appear self-centered provided more marked reasons than participants who were trying to appear rational. It is also consistent with work by Epley, Keysar, Van Boven, and Gilovich (2004), who showed that perspective takers, when imagining how others will act, first anchor on the self (e.g., how would *I* act) and then adjust away from this anchor (the target it not *me,* so how would *he* or *she* act?). Epley et al. also demonstrated that the adjustment away from the self is often insufficient. That is, it is difficult for people to get away from their own egocentric perspective. In the current case, at least for marked beliefs, this may have also been the case. Perspective takers might have started their explanations imagining why they *themselves* would have acted in particular ways, and then adjusted away from that anchor to account for the fact that it was not *them* who performed the action – it was the target.

Considering the fact that perspective takers were also explaining the target's behavior in scenarios where he may have been imagined as responding to blacks in racist ways provides further insight. Under normal circumstances, perspective takers might tend to explain a target's behavior more similarly to the way that people explain their own behavior, but in the present case, perspective takers might have used their explanations of the target's behavior (particularly during the race-relevant scenarios) as an opportunity to distance themselves from the "target's" racist beliefs by making his actions especially heinous.

For all other hypotheses, as noted above, perspective taking did not have any clear effect. In fact, in many cases, perspective takers and third-person participants looked almost exactly the same in the ways they explained the target's behavior, which was

often different from those in the self condition. An important question then becomes: Why?

A closer examination of the methods used may provide some insight. First, and relevant to both race-salient and race-neutral scenarios, is that while participants in the perspective taking condition were instructed to write their essays in the first person, as if they were the target (all participants complied), the behavior explanation questions were asked about the target in the third-person. For example, the first question participants were asked was "Why did the person in the photograph [in the control condition, *you*] approach that particular clerk?" Because this wording should linguistically force participants out of a "self-as-target" frame of reference and into a "target-as-other" frame, asking the question this way provides a more stringent test of hypotheses than asking participants, for example "As the target, why did you…?" – a formulation that is more consistent with the linguistic frame of the experimental manipulation.

Another obvious consideration related is that because the target was a racist (and not well-liked), the very behaviors which were being explained – particularly in the race-relevant scenarios – were quite different in the perspective taking and third-person conditions compared with the self condition. That is, the behaviors that participants in the self condition explained were more often their attempts to *avoid* appearing racist in the race-related scenarios, while the behaviors that participants in other conditions explained had been crafted to demonstrate the target's racism.

For example, consider the following content from participants' essays (taken from the scenario where participants were asked to imagine that they were in a restaurant and

overheard employees laughing about their black manager and affirmative action, and where the target (or the self) then joined in the conversation before leaving a large tip). In the self condition, one participant wrote: "As I was waiting to order at the counter I overheard the employees talking about the new manager. I noticed that the manager was black and they seemed to be cracking some jokes about it. I was really getting upset because I would rather someone come over to take my order than make jokes about their new black manager. So I asked if someone could help me then I got my food and left a large tip and walked out." This can be contrasted with a relatively subdued but typical response to this scenario in the perspective taking condition: "Of course the man was black. If they're not running for president, they're taking good, hardworking Americans jobs away from them. We got into a brief discussion about Affirmative Action and how that's the only reason blacks can get anywhere in this country. My spirits boosted, I finished my meal, left a big tip, and left."

For the race-neutral scenario, the reason for the lack of a finding is not as clear. Should not the lack of race as a theme in this scenario allow perspective taking to have its effect on behavioral explanations? Several considerations would suggest that in the present study, the answer is still "not likely." First of all, there was also a very consistent trend for participants in all conditions to give very similar, short, direct answers to the question prompts for this scenario. For example, the question "Why did the person in the photograph [you] bring a book with him [you]?" led to responses such as "Because he [I] felt like reading" or "He [I] likes to read." To the question, "Why did the person in the photograph sit on the bench?" almost half of participants in all conditions answered

something like "He [I] was tired" or "He [I] felt like sitting down." Thus, there was very little variability in the types of responses that were given, which probably further limited a reasonable test of the perspective taking hypothesis.

Furthermore, the behaviors of the target were often still related to race even in the race-neutral scenario. For example, several participants in the perspective taking and third-person conditions mentioned that the book the target took with him was on white supremacy or other racial topics, and one perspective taker suggested that the target was taking a walk to get away from anywhere he might see a black person. This was not the case for participants in the self condition, so the very behaviors being explained in the neutral scenario were not always of the same nature.

Considering all of the facts described above – the different linguistic framing of question prompts for perspective takers versus those in the self condition, the inclusion of race-related scenarios that provided perspective takers an opportunity to distance themselves from the target and his actions, the fact that because the target was clearly racist, the types of behaviors being explained were different for perspective takers and participants in the self condition, and the lack of variability in types of responses for the race-neutral scenario – lead to a very stringent test of perspective taking hypotheses, and very little chance of supporting them. Given all of the reasons for perspective takers' explanations to look very different from those of participants who wrote about themselves, what is perhaps most surprising is that *any* effects emerged.

One further idea is that in all conditions people *imagined* the target's (or their own) actions, and then provided hypothetical explanations for each of the imagined

actions. When trying to explain everyday actions, explanations might be expected to typically follow cultural scripts, particularly when little information is available about a target person to provide richer detail. Because of this, it might also be expected that linguistic markers might be one way participants could indicate a perspective taking attempt – which is what was found here. One reason this claim can be made is that cultural scripts do not dictate whether a perceiver will mention a target's mental state – but the perceiver's perspective and attitudes toward the target will likely affect whether the perceiver will try to embrace or distance himself or herself from the target's actions.

Because the effects of perspective taking on behavioral explanations did not represent the primary hypothesis in the present research, the design was clearly less than optimal. The present research did provide some limited evidence, however, that even under these less than optimal circumstances, perspective taking can have an effect on people's behavioral explanations – even though the only evidence was found concerned the use of mental-state markers for belief reasons. Future research could fruitfully examine similar hypotheses as were advanced here and could even use similar methods, perhaps varying target characteristics and behavioral scenarios to see if whether taking the perspective of a target who is liked (versus disliked, as was the case here) and is performing social (versus solitary) behaviors has the effect of shifting participants' explanations so that they look more like those of people explaining their own behaviors.

Moving Forward

Aside from being a lengthy and probably tiring study for participants, Study 1 was clearly also psychologically complex. Participants' motivation to comply with

instructions was probably influenced by their own attitudes and beliefs about African Americans – and by the strong proscriptions against appearing racist (e.g., Judd et al., 1995) – but also by the target and his attitudes. This likely meant that some probable goals (such as compliance) were at odds with other goals (such as avoiding getting into the head of a target whom they clearly did not like).

It is no wonder then that across many dependent variables and sets of related hypotheses, very limited support was found for the idea that perspective takers would come to share the negative attitudes of the target. Furthermore, a few findings departed from the usual results of perspective taking research. For example, the simple fact that the target was not liked at all, including by those in the perspective taking condition, departs from the typical finding that perspective taking leads to greater valuing of the target (e.g., Batson et al., 2007; Galinsky et al., 2008). In itself, this may have important consequences for perspective taking. That is, when the target is not valued, it is possible that people will be reluctant to try to imagine actually being that person.

When speculating about reasons why participants had such low merging and perceived similarity with the target after perspective taking, it seems a highly plausible explanation that this is simply because the target was hard to like. His very obviously strong and unsavory attitudes, dissimilar to those of the participants generally, might have stopped any real attempts to take his perspective. Rather than actually trying to get inside the head of the target, participants in the perspective taking condition might have for the most part simply relied on the fact that he was a racist to generate their narratives. Study 2 addresses this possibility, using a target who is more similar to the majority of

participants (i.e., a female target is used rather than a male), and whose attitudes, while still reflecting racial bias, are much more subtle than in Study 1.

Another important variable to consider is that participants were not given any further information about the target except what he looked like, and his attitudes were deduced from that visual information alone. In the absence of any other information, it seems highly likely that participants would not try to come up with any mitigating factors that might have played a role in the development of the target's attitudes. In effect, they probably considered him fully culpable for attitudes that are under his control. Study 3 addresses this possibility by manipulating the information that participants have about the target, giving him reasons that might make participants think his attitudes are less controllable and more understandable, thus making him seem less responsible for his beliefs.

CHAPTER III

STUDY 2: TAKING THE PERSPECTIVE OF A WOMAN WITH SUBTLY RACIST ATTITUDES

Introduction

Study 1 failed to find direct support for the hypothesis that taking the perspective of a target who clearly has negative attitudes about a racial group will affect the way that participants respond toward that same racial group. Still, this work is important for several reasons. First, it demonstrated that there are some limits to when perspective taking leads to greater positive regard for a target (e.g. Galinsky et al., 2008). That is, perspective takers did not like a racist target any more than participants who wrote about a day in the life of the target without taking his perspective, or participants who only briefly thought about the target. Furthermore, perspective takers constructed narratives that did not paint the target in a very flattering light, writing essays that made him appear more racist than other participants (who did not take his perspective), even though these other participants had the same information about the target's attitudes.

Despite the lack of direct support for the perspective taking-attitude change hypothesis, Study 1 did provide evidence that *some* perspective takers' attitudes were more like those of the target, which is the first time this has been reported. That is, potentially as an ironic result of perspective taking, perspective-taking participants who were more internally motivated to respond without prejudice scored higher in implicit

stereotyping. Furthermore, among perspective takers, perceiver-target merging and explicit racism were correlated to a larger (though not significantly larger) extent than was found in other conditions.

However, as noted above, the results from Study 1 were generally not consistent with hypotheses. Among other things, Study 1 also did not support previous models where self-target merging has mediated the results of perspective taking (e.g., Galinsky & Moskowitz, 2000; Galinsky et al., 2008; Maner et al., 2002). One issue with Study 1 is that the target was an obvious racist, and probably because of this, he was not well-liked and participants did not feel at all similar to him. Study 2 tackles both of these issues at the same time. Below, the questions of similarity and positive regard are addressed first before turning to the related issue of the nature of the target's attitudes (i.e., that they were quite overt and objectionable).

In Study 1 the target was judged as very dissimilar on average to participants. As an indication of this, on a 7-point scale, the average amount of perceived similarity participants reported having with him on the single item similarity measure was only 1.72. Furthermore, a composite measure of positive regard for the target was also quite low ($M = 1.84$). Considering that similarity and positive regard were positively correlated ($r = .48$), that similarity leads to valuing the other (e.g., Batson, Turk, et al., 1995), and that similarity and positive regard have both been investigated as important variables in perspective taking research (e.g., Batson, Polycarpou, et al., 1997; Goldstein & Cialdini, 2007), it seems important that people did not like this target, and certainly did not feel similar to him.

When considering *why* participants did not like this target and did not feel (or were generally unwilling to admit feeling) similar to him, one does not have to look far. As discussed earlier, there are strong norms against appearing racist (e.g. Wyer et al., 2000). Many researchers have claimed that "old-fashioned" or ""blatant" racism (e.g., McConahay, 1986; Pettigrew & Meertens, 1995), while not gone, has been to a large part replaced by more indirect forms called "subtle," "modern," "symbolic," or "aversive" (Gaertner & Dovidio, 1986; McConahay, 1986; Gaertner & Dovidio, 1986; Pettigrew & Meetrens, 1995; Sears, 1988). These forms of racism are not necessarily acceptable, but are thought to be less unacceptable to endorse. The target used in Study 1 was a fairly straightforward, if not extreme, example of an old-fashioned type of racist. The fact that he wore a white pride t-shirt indicated a willingness or even a desire to share his views with the world without apology. Participants therefore had a good reason to not like him and to rate themselves as very dissimilar from him.

Another issue to consider, at least in terms of similarity, is that the target was male. This takes on meaning when considering that the majority of participants in Study 1 were female (approximately 70%). While sex was not a major factor for the results in Study 1, the vast majority of participants – females – would likely feel more similar to a target who is the same sex as they are.

Therefore, the target was changed in several ways for Study 2. The first change is that the photographed target presented to participants was of an attractive blond female rather than a male. To further encourage liking, similarity, and merging, information was included about the target that portrayed her as relatively likable and friendly – the kind of

person one might want as a friend. Second, her attitudes toward African Americans were presented as less harsh and more subtle than those of the Study 1 target.

One element of racism that may be particularly subtle is intolerance for interracial romantic relationships. While many might not even consider this to be an issue of "racism" per se, research has shown that disagreement with items such as "I would not hesitate to date [marry] an African American" loads on a racism factor along with racist responses to other items from scales such as McConahay's (1986) Modern Racism Scale, such as "Because of affirmative action, many African Americans have gotten jobs they are not qualified for," "I feel comfortable talking to African Americans," "African American culture has a lot of problems which keep African Americans down," and "It's [not] easy to understand the anger of African Americans" (Baldwin, Day, & Hecht, 2000). Therefore, information about the target not wanting to date a black man because she had never been comfortable around blacks (see Appendix B) was woven into a short narrative about the target, included at the end so as to not unduly influence participants' early perceptions of the target as generally likable.

Another subtle form of racism is disapproval for African Americans holding positions of power. Prior to the data collection phase of this study, a United States presidential election was held where one of the candidates (and now president, Barack Obama) had a mixed racial background, and is considered by many to be an African American (Swarns, 2007). While being outspoken against this candidate might not be construed as necessarily driven by racial attitudes, in conjunction with other information such as learning that the target has never felt comfortable around blacks it could help

create the sense that this person has negative attitudes toward blacks. Furthermore, these racial attitudes might be of a form more acceptable to many participants. Therefore, a comment which suggested that the target did not generally like to "talk politics" but was vocal in her opposition to Barack Obama was included.

Like Study 1, participants in Study 2 were asked to take the perspective of the target and write about a day in her life. Unique to this study was that rather than having other participants simply write about a day in the target's life with no further information (e.g., Galinsky & Moskowitz, 2000), some participants were explicitly told to be "objective" when writing their essay about the target. This instruction was adapted from control conditions used by many perspective taking researchers (e.g., Batson, Sager, et al., 1997; Galinsky et al., 2008; Vescio et al., 2003), and was included as a way to make participants remain more objective about a relatively compelling target. Also unique to Study 2 was an additional condition in which participants did not write about themselves and instead took the perspective of the target, but without being given any information about the target's attitudes (i.e., no information about the target's attitudes with respect to African Americans was given). This allows a test of taking the perspective of a racist target against both an objective condition where information about the target's attitudes is also present, and another perspective taking condition where this information is not available.

Study 2 used some of the dependent variables from Study 1, but greatly reduced the number of explicit measures, and focused on two global measures of attitudes towards blacks (the semantic differential scale and the feeling thermometer). By using fewer

measures, the study was greatly simplified, making it less obvious that the study was about "perceptions of race." (Additional measures were also taken to make this less obvious; see below.)

Study 2 also included a new measure, taken in full from Wyer et al. (2000). This measure involved having participants read a short vignette about a person of unspecified race named "Robert." In the course of the story, Robert engages in three ambiguous instances each of hostile, passive, unintelligent, and intelligent behaviors. Thus, Robert's behavior could be interpreted as consistent (or not) with African American stereotypes. After reading this story, participants were asked to rate Robert along several dimensions including stereotype-relevant traits (e.g., rude, unintelligent) that are typically associated with African Americans. Because Robert's race was not specified, if participants rated him higher on these stereotypical negative traits, this might indicate a greater activation of traits associated with negative black stereotypes.

Other changes were that Study 2 was presented as two separate studies, one which had participants form an impression of a target and answer a few questions about that target and themselves, and another that presented the IAT as a separate experiment about how people make rapid classifications of different categories of information (e.g., Galinsky & Moskowitz, 2000). In the first "study," participants either took the perspective of the target and wrote about a day in her life, or wrote about a day in her life from an objective perspective. No instructions were given about writing in the first versus third person, allowing a manipulation check of whether participants linguistically complied with perspective taking instructions. It was expected that perspective takers

would write in the first person and objective participants would write in the third person (Galinsky & Ku, 2004).

Also, because examining behavioral explanations was not a part of this study, participants were not given any instruction as to what to include in their essays. This also allowed for stronger inferences about the content of their essays (e.g., evidence of racism, use of racial labels), because they were not asked to write about scenarios that specifically involved race. Therefore, any racial content would presumably be driven by perceptions of the target rather than by specific behaviors that were asked to be included in the essays.

A final change was that participants provided responses to several measures in an unrelated internet-based study earlier in the term. This was done for two reasons. First, this allowed the possibility of ruling out whether participants' motivation to respond without prejudice would again interact with condition when it was not measured in the experimental context. A second reason for collecting measures earlier was that it would allow the chance to examine changes in racial attitudes from a baseline when there were no particular pressures to respond without prejudice (i.e., because participants responded to measures on their own time, via the internet, there should be less social pressure to respond in positive ways about racial attitudes). Measuring attitudes again after the experimental induction would allow for a test of whether attitude change took place.

Hypotheses

As in Study 1, the first hypothesis (Hypothesis 1) was that taking the target's perspective would lead to greater perception of similarity to the target, greater positive

regard for the target, and greater perceiver-target merging. That is, it was expected that similarity, positive regard, and merging would be greater – regardless of whether participants had information about the target's racial attitudes – when participants took the target's perspective compared to when they wrote about a day in her life from an objective perspective. Exploratory analyses were also again examined to see whether any of these variables were associated with participants' own racial attitudes.

A second hypothesis (Hypothesis 2) was that as in Study 1, taking this target's perspective (when information was available about her racial attitudes) would lead participants to craft essays that included more racist content than participants in other conditions. This was expected even though the target's attitudes were presented as relatively subtle, no race-salient scenarios were included, and despite the fact that no specific instructions were given explicitly prompting participants to organize their essays along racial lines. If this prediction is supported, it would provide further evidence that perspective takers were clearly and vividly imagining what it would be like to be the target, and that they paid attention to the information regarding the target's racial attitudes, even though these attitudes were somewhat understated.

The primary hypothesis, similar to Study 1 (Hypothesis 3) was that perspective takers who had information about the target's racial attitudes would have elevated scores on either implicit or explicit measures (or on both) that assessed their own negative racial bias. Last, the possibility that participants' motivation to respond without prejudice would interact with condition to predict their racial attitudes was again explored, with the expectation that those participants in the perspective taking (racist target) condition who

are higher in motivation to respond without prejudice might experience rebound effects (e.g., Macrae et al., 1994; Wyer, 2007) and end up ironically demonstrating more prejudiced responses.

Method

Participants

Participants were 101 undergraduate students (66 female, 35 male) who participated in partial fulfillment of course requirements. As in Study 1, participants were allowed to identify with one or more racial/ethnic groups. Seventy-nine self-identified as white, 3 as black, 16 as Asian, 6 as Latino/Latina, 1 as Middle Eastern, 2 as Native American or Alaskan Native, and 5 as mixed (non-black). Two participants declined to respond. Ages ranged from 18-31 ($M = 20.01$, $SD = 2.00$). Thirteen students responded that English was not their first language, and indicated that they had been speaking English for 2-18 years ($M = 10.00$, $SD = 5.51$). Excluding these participants did not substantively affect the results from any analyses, so they were retained. The three black participants were excluded from further analyses.

Procedure

Only participants who voluntarily completed an online general survey prior to signing up for the study were eligible to participate. The general survey was typically completed early in the term, sometime between 2 and 9 weeks before the experimental session, and contained a variety of unrelated measures used by multiple researchers within the same department, including the ones used here.

On arrival in the laboratory, participants were told that to save time, they would be taking part in two unrelated studies (e.g., Dovidio et al., 2004), and that they would be debriefed about both following the second study. The first study was presented as an impression formation task, with procedures similar to Study 1 (e.g., all participants chose a "randomly" selected target; e.g., Montieth et al., 1998). After selecting a target photograph, participants in two perspective taking (for a subtly racist target – "PT-racist;" for a target where no attitude information was given – "PT-control") conditions were asked to take the perspective of the person in the photograph and write about a day in that person's life as if they were that person. (Complete perspective taking instructions can be found in Appendix A.) In the objective condition, participants were asked to write about a day in the person's life from an "objective" perspective. Specifically, participants in this condition were instructed:

> After studying the photograph the experimenter gave you, and reading the information about the person in the photograph, you should spend 5-10 minutes writing a short essay about a typical day in the life of that person. While writing this essay, try to be as objective as possible when imagining what is happening to this person and what their day is like. Try not to let yourself get caught up in imagining what this person has been through or how the person feels. Simply describe a day in the life of this person as objectively as possible.

Following this, the next screen on the computer presented a vignette about the person in the photograph ("Amanda"). In the PT-racist and objective conditions, information about the target's attitudes was included (described above, see Appendix B for full vignette); in the PT-control condition, participants also took the perspective of this target, but the information about her racial attitudes was not presented. Following the essay-writing task, participants responded to questions about the essay task and their

perceptions of the target, completed several explicit measures (described below), and provided demographic information. Next, they were told that they were done with the first study, but that they would be debriefed after completing the second study.

The second study was presented as being related to "hemispheric dominance," attention, and rapid categorization. Participants filled out a different informed consent form (written in a different style, with a different font, and listing a different experimenter than the first study). To bolster the cover story (and also serving as distracter tasks), participants were first asked to trace a series of complex shapes with their non-dominant hand, and information on handedness was clearly recorded by the experimenter. They then completed an unrelated measure (the Need for Cognition Scale; Cacioppo, Petty, & Kao, 1984).

After completing these two distracter tasks, but before doing the IATs, the experimenter "noticed" that she had forgotten to give participants a measure from the first study. That is, she told participants that while they were doing the tracing task, she noticed that she had clipped an additional impression formation task under some other papers on her clipboard, and that she had forgotten to give them this task, which was part of the previous study. She apologized and asked them if they could do it now, telling them that it was really short, and involved reading a paragraph about a person and giving some impressions of that person. All participants complied, and none later expressed suspicion. Participants then read the "Robert" story and rated him on a set of traits (described below).

Participants then did the IATs (described below), were probed for suspicion, and fully debriefed. No participants reported awareness of the connection between the two studies.

Measures

As in Study 1, participants were asked several questions about their perceptions of the target and the essay writing task, as well as their perceived similarity and merging with the target. All four "regard for target" items (i.e., liking and caring about the target, wanting to spend time with the target, and the target's friendship potential; see Study 1) were again highly correlated (average $r = .53$), except in the current study, the similarity question was also highly correlated with this set of questions (average r of similarity with each other item = .52). All five of these variables were thus averaged to form a measure of "positive regard" ($\alpha = .85$). A merging variable ("merging") was created by averaging responses to the IOS and the "we" question ($r = .77$, $\alpha = .87$). The target's perceived racism ("target racism") was again computed from two items that asked how racist the target was and how much the target likes African Americans ($r = .67$, $\alpha = .80$). Effort exerted trying to imagine the target ("effort") was measured with the same five items used in Study 1 ($\alpha = .76$), and a single item difficulty measure was retained as a measure of essay-writing task difficulty ("essay difficulty").

To assess whether participants in the objective condition actually remained objective when they created their narratives (and whether perspective takers took the target's perspective), two items were included here that were not included in Study 1. One item measured the extent to which participants took the perspective of the target

("To what extent did you try to "take the perspective" of the person in the photograph?"), and one item measured the extent to which participants stayed objective ("To what extent did you try to remain objective about the person in the photograph?"). Both were measured on 7-point scales, where 1 = "not at all" and 7 = "very much." To create a composite measure of objectivity, the "take the perspective" item was reverse-scored and averaged with the objective question (r = .48, α = .64) to form an "objective" measure.

As in Study 1, essays were coded along several dimensions by two trained coders who were blind to hypotheses and condition (for all codes, coder r = .89, α = .94). These codes included whether the essay was written in the first person, which was simply a note of whether the essays were written from a first person perspective (e.g., "I woke up and went to the store") or a third person perspective (e.g., "He woke up and went to the store"). In addition, codes included counts of neutral and negative racial labels (e.g., words like "black" or "nigger," respectively), and negative and positive stereotypes of blacks. Subjective ratings of the target's racism and attitudes toward blacks were again coded (for target racism, 0 = "no evidence of the target being racist," 1 = "ambiguous or mixed evidence," 2 = "target is somewhat racist," and 3 = "target is very racist;" attitudes toward blacks allowed for the possibility of positive attitudes toward blacks, and codes were 0 = "very positive (attitudes toward blacks)," 1 = "somewhat positive," 2 = "neutral/mixed," 3 = "somewhat negative," and 4 = "very negative"), as was a code that assessed the creativity and imagination of the essays (0 = "not very creative," 1 = "slightly creative," 2 = "fairly creative," 3 = "extremely creative").

As in Study 1, the essay creativity/imagination code ("essay imagination") was retained as a single-item measure. Unlike Study 1, very few participants included any negative black stereotypes (8%) or neutral category membership labels such as "black" (20%). Furthermore, 76% of essays showed no evidence of the target being prejudiced, and 85% of the essays also were neutral with respect to the target's feelings about blacks, showing either mixed evidence or no evidence of negative attitudes toward blacks. It is instructive to note that for each of these codes, there were significant differences (all $ps \leq .001$) among condition, and follow-up tests indicated that the perspective taking/racist target condition had significantly greater evidence of racist content and race-related themes than both other conditions for all codes (most $ps \leq .005$, except for the code concerning number of racial category labels, where the comparison between PT-racist and objective was $p = .03$). Objective and PT-control did not differ on any of codes. Because they were highly correlated, these items (negative stereotypes, neutral racial labels, target racism, and target's attitudes toward blacks) were therefore first standardized, and then aggregated to form a measure of "essay racism" ($\alpha = .89$).

Participants again completed two IAT measures (prejudice and stereotyping), although these tests were slightly changed from Study 1. As noted by Devine et al. (2002), even though research has shown that IAT effects are not driven by differential familiarity of black and white names (Dasgupta, McGhee, Greenwald, & Banaji, 2000), using faces rather than names to represent racial categories may be more compelling as faces are clearly and easily categorizable. Thus, for this reason, and to provide generalizability of any IAT effects, 6 black and 6 white faces (3 male and 3 female from

each racial group) were used as the racial categories rather than using black and white names (as in Study 1). The same positive and negative stereotypes (stereotype IAT) and pleasant and unpleasant words (prejudice IAT) used in Study 1 were again used, and scoring was again with the *D* metric. The photographs of faces were taken from Project Implicit (Nosek, Banaji, & Greenwald, 2006), a website devoted to IAT research.

Internal and External Motivation to Respond without Prejudice was again assessed in the current study, but participants' responses on this measure were collected as part of a general survey available to the subject pool. Collecting these data at a time unrelated to the study was meant to stop any potential influence of condition on responses. Both measures were adequately reliable (for both internal and external motivation, αs = .79).

The feeling thermometers for blacks and whites that were used in Study 1 were again used here, and were again reverse-coded to reflect coolness (rather than warmth) toward whites and blacks. Scores on the white thermometer were then subtracted from scores on the black thermometer so that positive scores reflected (as in Study 1) greater coolness toward blacks relative to whites, and negative scores indicated the opposite.

The same semantic differential scale used in Study 1 was again used in Study 2, but participants' responses were collected twice. As with motivation to respond without prejudice (described above), participants completed this scale earlier as part of a general survey, via the internet. They completed the same measure a second time during the experimental session, after the essay writing task, in order to test whether their responses differed from baseline as a result of experimental condition. Semantic differential

difference scores were created by first averaging responses to the five items ("good/bad," "nice/awful," "honest/dishonest," "beautiful/ugly," and "pleasant/unpleasant") for the general survey (for blacks and whites, respectively, αs = .91, .87) and for the experimental session (for blacks and whites, respectively, αs = .84, .86) separately (higher scores reflecting more negative attitudes), and then subtracting baseline scores from experimental session scores. This variable ("semantic differential") thus reflects both negativity in attitudes toward blacks relative to whites and change from baseline, with positive scores meaning higher relative negativity toward blacks during the second measurement and negative scores reflecting greater positivity.

As described above, one additional dependent variable used here was taken from Wyer et al. (2000), and is fully described there. After reading a short vignette describing a day in the life of a character named "Robert" (whose race was not specified), where Robert acted in ways that were ambiguously hostile, passive, intelligent, and unintelligent, participants rated him on the following traits, using a 9-point scale: rude, hostile, aggressive, respectful, passive, intelligent, and unintelligent. Seven other non-relevant traits, (e.g., "funny," "anxious") also served as distracters from the purpose of the task.

For responses to the Robert story, the items rude, hostile, aggressive, respectful (reverse-coded), and passive (reverse-coded) were aggregated as a measure of Robert's rudeness ("rude," α = .78) and the items intelligent (reverse-coded) and unintelligent were aggregated as a measure of Robert's lack of intelligence ("unintelligent," r = .45, α = .61). Rude and unintelligent were minimally related (r = .20, p = .05).

Results

Descriptive Statistics

Table 5 (below) provides descriptive statistics (including reliability statistics where appropriate) for all variables which were retained. Also provided are possible score ranges for each variable.

Table 5.
Descriptive Statistics for Study 2 Measures

Variable	α (*r*)	*M (SD)*	Possible Ranges
Positive Regard	.84	3.63 (1.19)	1 to 7
Merging	.87 (.77)	2.62 (1.38)	1 to 7
Target Racism	.80 (.67)	64 (1.32)	1 to 7
Effort	.76	5.06 (.93)	1 to 7
Essay Difficulty	-	4.34 (1.51)	1 to 7
Objective	.64 (.48)	4.32 (1.31)	1 to 7
Essay Imagination	-	1.81 (.74)	0 to 3
Essay Racism	.89	0.00 (.87)	-
Prejudice IAT	-	.28 (.32)	-
Stereotype IAT	-	.27 (.37)	-
Internal Motivation to Respond without Prejudice	.79	7.58 (1.27)	1 to 9
External Motivation to Respond without Prejudice	.79	5.13(1.89)	1 to 9
(Relative)Thermometer	-	-.50 (12.18)	-100 to 100
(Relative Pre-Post) Semantic Differential	-	.44 (1.04)	1 to 7
Rude	.78	4.78 (1.30)	1 to 7
Unintelligent	.61 (.45)	4.36 (1.26)	1 to 7

Data Examination

Again, distributions of all variables were examined. The merging variable showed very slight positive skew that was corrected by a square-root transformation. Analyses involving merging were examined using both the transformed and untransformed variables, and did not differ by which variable was used, so all reported analyses focus on the untransformed variable. The essay racism content variable was also quite skewed

because most participants had showed no evidence of including any of the relevant codes that made up the composite. No transformations were able to correct this skew. The original untransformed variable was therefore retained and tested, and the differences among conditions are meant to be primarily descriptive. An additional dichotomous variable was also created where "0" = no evidence of racism in the essay and "1" = some evidence, and tested for differences across condition with a χ^2 test.

Primary Analyses

As in Study 1, most analyses were conducted using analyses of variance (ANOVAs) treating condition as a factor and testing for sex effects. When no influence of sex was present, this term was removed from further models and is not discussed. For the essay racism content variable, an additional χ^2 test is reported in addition to the ANOVA.

Manipulation Check and Perceptions of the Target and Essay Task. First, to test whether participants in the objective condition remained more objective than participants in the PT-racist and PT-control conditions (and that participants in these conditions took the target's perspective), the "objective" variable was tested. A significant effect of condition was found, $F(2, 95) = 9.19$, $p < .05$. The objective condition ($M = 4.93$, $SE = .18$) was more objective than PT-R ($M = 3.96$, $SE = .18$) and PT-NR ($M = 4.02$, $SE = .19$) conditions, which did not appear to differ.

Next, the perceived difficulty of the essay task, perceptions of the target as racist, and the effort expended trying to imagine the target were examined. Starting with the latter first, effort spent trying to imagine the target did not differ across conditions.

Participants' perceptions of the essay writing task as difficult, however, were significantly affected by condition, (F (2, 95) = 3.20, $p < .05$), with the PT-racist condition (M = 4.15, SE = .25) reporting the task as more difficult than the PT-control condition (M = 3.77, SE = .27), who in turn thought it was more difficult than the objective condition (M = 3.11, SE = .25). This suggests that trying to take a racist target's perspective feels more difficult than taking a non-racist target's perspective, while on the whole, trying to take any target's perspective feels more difficult than writing about a day in the life of a racist target while remaining objective, even if a similar amount of effort is perceived to be expended in trying to imagine the target, regardless of condition.

For perceptions of the target as racist, the expected effect of condition did emerge F (2, 95) = 37.38, $p < .0005$, partial $\eta 2 = .44$. Participants in PT-racist (M = 5.32, SE = .17) and objective (M = 5.13, SE = .17) conditions did not differ much in their perceptions of the target as racist, and both of these conditions thought the target was more racist than the PT-control condition (M = 3.33, SE = .18), showing that even though the vignette about the target was constructed to be subtle, participants considered the target to be somewhat racist when they had relevant information about her attitudes. It should be noted that the perception of the target as racist in the PT-control condition appears to indicate that these participants may have thought the target was somewhat racist (an unintended outcome). The midpoint of this scale (a rating of "4"), however, could have been interpreted by participants as "I don't know," or "neutral" rather than a rating of the target as somewhat racist. That is, anchors for the items that make up this scale were 1 = "not at all racist" and 7 = "very racist" for one item and 1 = "likes (blacks)

a lot" and 7 = "strongly dislikes." Thus, with no information available about the target's attitudes, participants likely assumed that she was somewhere in the middle, perhaps indicating that the target had no strong attitudes about blacks one way or the other.

Hypothesis 1 – Potential Mediators. Turning to tests of main hypotheses, recall that Hypothesis 1 was that perspective takers – regardless of whether they had information about the target's racial attitudes – would feel more similar to and merged with the target, and also regard the target more positively than participants in the objective condition. Because similarity did not appear to be distinct from positive regard, it will not be discussed further (except in the context that it was one of the items which made up the positive regard variable).

As was the case in Study 1, little support was found for this hypothesis, at least for those perspective takers who had information about the target's racial attitudes. For positive regard, there was a main effect of condition, F (2, 95) = 4.77, p = .01, partial η^2 = .09. The PT-control condition (M = 4.15, SE = .21) regarded the target more positively than the PT-racist (M = 3.40, SE = .20) and objective (M = 3.35, SE = .19) conditions, which did not differ from one another. A marginally significant main effect of merging was also present (F (2, 95) = 2.67, p = .07, partial η^2 = .05), and again, the PT-control condition (M = 3.01, SE = .25) appeared to be somewhat more merged with the target than the PT-racist (M = 2.35, SE = .23) and objective conditions (M = 2.42, SE = .23) conditions.

Follow-up tests examined the correlations between positive regard, merging, and each of the dependent variables that involved participants' racial attitudes. Unlike in

Study 1, neither variable was correlated with participants' attitudes either across conditions, or within any condition. Although there was some evidence that perspective taking could influence positive regard for the target and merging, at least when no information was presented that suggested the target was racist, when the target *was* perceived as racist, positive regard and merging were no greater for perspective takers than for participants who were explicitly told to remain objective. This may demonstrate a potentially limiting condition on when perspective taking can induce more positive feelings toward a target. That is, if a target is disliked, or has even mildly reprehensible attitudes, perceivers may try to avoid taking the target's perspective, and when asked to do so, respond by asserting that they do not feel similar to, merged with, or like the target.

Hypothesis 2 – Essay Codes. As in Study 1, Hypothesis 2 was that perspective takers (who had information about the target's racial attitudes) would write essays with significantly more racist content than participants in other conditions. The creativity of participants' essays was also examined.

The essay racism composite was significantly affected by condition, $F\,(2, 98) = 15.62$, $p < .0005$, partial $\eta^2 = .24$. The PT-racist condition ($M = .60$, $SE = .13$) wrote essays with significantly more racist content than the objective condition ($M = -.19$, $SE = .13$), and the PT-control condition wrote essays with the least amount of racist content ($M = -.40$, $SE = .13$). A test using the dichotomized essay racism variable which tested the PT-racism condition against both other conditions combined was significant,

with 16/33 (48.5%) essays in the PT-racist condition and only 5/68 (7.4%) in the other two conditions combined showing some evidence of racism, χ^2 (1, N = 99) = 22.82, $p < .0005$.

Even though there was limited racism expressed in participants' essays (as discussed in the Methods section, above), support was again found for the hypothesis that taking a racist target's perspective – even when the target's racial attitudes are subtle – leads to more racist content in essays than does writing about the same target without taking the target's perspective. This lends support to the idea that when confronted with a target who appears to be racist, perspective takers may call to mind racist images and information to create their narratives

Again (as in Study 1), this does not mean that participants' *own* attitudes subsequently become more racist, but it does suggest they *can* imagine what it would be like to be a racist – probably using their own stereotypes of what a racist would be like – and use this information to portray this target as such. Moreover, as is evident here, this effect was especially strong for perspective takers (i.e., in the objective condition, almost no reference to the target's attitudes were made), which points again to differences between taking a target's perspective and simply thinking about the target.

Hypothesis 3 – Participants' Racial Attitudes. Again, contrary to prediction, condition had very little effect on participants' racial attitudes. That is, for both stereotype and prejudice IATs, and also for a variable which represented changes in negativity toward blacks relative to whites from an earlier pretest to the experimental session (i.e., semantic differential), no effects were found.

For the relative thermometer variable, however, where higher numbers represented greater coolness toward blacks relative to whites, a surprising and unpredicted effect for condition did emerge, F (2, 92) = 3.29, p = .04, partial η^2 = .07. Here, the PT-control condition scored slightly higher in relative coolness toward blacks (M = 4.0, SE = 2.18) than the PT-racist (M = -2.09, SE = 2.08, p = .11) and objective conditions (M = -2.86, SE = 2.01, p = .06).

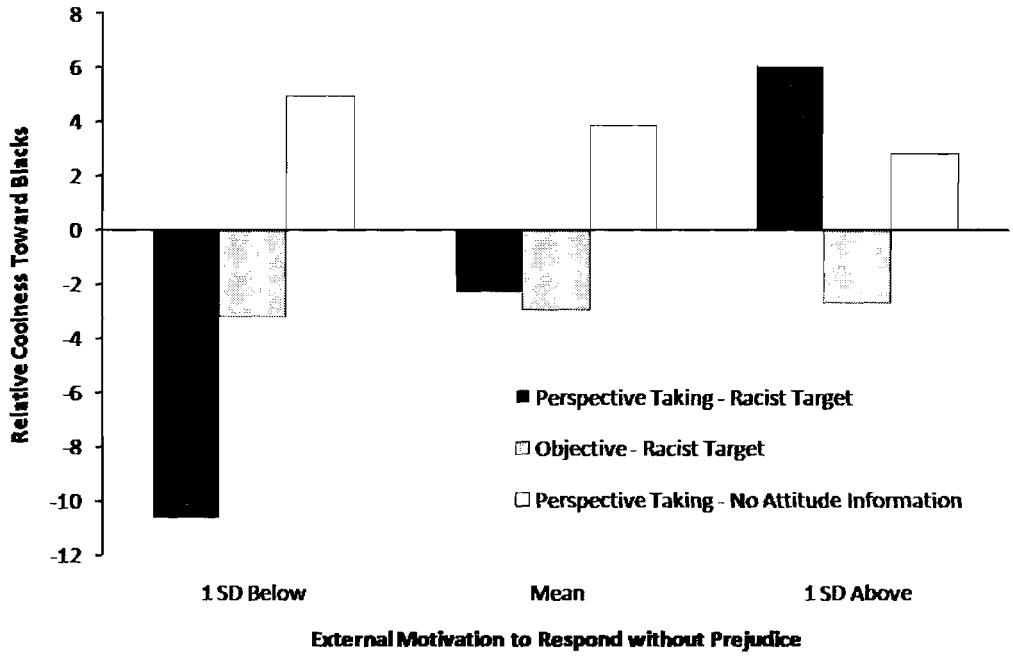
Although this effect was not predicted, and is not even in the predicted direction, it is still interesting. That is, perspective takers who had no particular reason to express (or inhibit expressing) racism had slightly elevated thermometer scores relative to participants who had just written an essay about a racist target. One possible interpretation of this pattern of results is participants who had just thought about a racist target may have been either consciously or unconsciously trying to avoid appearing racist, or were feeling slightly warm toward or protective of African Americans to a greater extent than those in the PT-control condition. It is important to realize, though, that this effect was not very large, and only represented a small (although significant) shift in relative attitudes toward blacks.

Motivation to Respond Without Prejudice. As in Study 1, the next test focused on whether motivation to respond without prejudice would interact with condition to affect participants' attitudes. To test this, two dummy codes were computed, where the PT-racist condition was always coded "0" and each other condition was coded "1" (in separate codes). Internal and external motivation to respond without prejudice were both grand-mean centered and multiplied by the dummy codes, and sets of predictors were

entered into multiple regression models in blocks (first, the set of dummy codes and the centered moderator were entered, followed by the interaction terms). Significance of interactions was assessed with changes in R^2. (Also as in Study 1, sex was tested as an additional factor in these moderated analyses, but because it was never significant in any of these analyses nor did it significantly interact with any other variables, it is not discussed further.)

Unlike in Study 1, internal motivation to respond without prejudice did not have an effect on stereotype IAT scores. In addition, no significant interactions emerged for either of the motivation to control prejudice variables on the prejudice IAT, semantic differential, or perceptions of "Robert" (the indirect measure of activation of black stereotypes created by Wyer, 2007, and used here) as rude or unintelligent. An interaction did emerge, however, between condition and external motivation to respond without prejudice (EM) on the relative coolness toward blacks measure, F (2, 86) = 4.32, p = .02, $R^2\Delta$ = .08. Exploring this interaction showed that the relationship between EM and coolness was positive and significant in the PT-racist condition (B = 4.39, β = .69, p = .003), and this relationship differed from the objective (B = -4.26, p = .02) and PT-control conditions (B = -4.95, p = .006). Further analyses showed that EM was not significant in the objective (B = .12, p = .90) or PT-control (B = -.59, p = .59) conditions. Figure 2 (below) plots mean levels of relative coolness toward blacks compared with whites within each condition at low, average, and high levels of EM.

Figure 2
Interaction of Condition with External Motivation to Respond without Prejudice on Relative Coolness toward Blacks.



A marginal interaction between condition, external, and internal motivation to respond without prejudice was also found (F (2, 79) = 2.66, p = .08, $R^2\Delta$ = .05), which showed that internal and external motivation interacted (B = -3.21, p = .06) in the PT-racist condition only (PT-racist versus objective, B = 4.61, p = .02; versus PT-control, B = 3.41, p = .07). In effect, this means that higher external motivation not only predicted higher relative coolness only in the PT-racist condition, it only predicted it when participants were at the mean or below on internal motivation. At one *SD* above the mean internal motivation, the effect for external motivation on coolness was smaller and

nonsignificant (B = 1.24, p = .60) and at 1 *SD* below the mean on internal motivation, the effect was even stronger (B = 9.12, p = .001). None of these effects (i.e., of external motivation to respond without prejudice, or of external by internal motivation on participants' relative thermometer scores) was significant in any other condition.

Discussion

Study 2 further explored the hypothesis of attitude change following perspective taking for a racist target. Unlike Study 1, where the only information participants received about the target came from a photograph which pictured the target in a "white pride" t-shirt, Study 2 took a different approach. That is, a female rather than a male target was used, information was provided in order to make her appear relatively "regular" and likable, and her attitudes were crafted to be somewhat subtle and less blatant.

This strategy was successful in creating higher overall positive regard for and merging with the target, which was higher in this study than in Study 1. The subtlety of the target's racial attitudes was also not missed by participants, who rated her as less racist overall than in Study 1 – although importantly, they still found her to be somewhat racist (i.e., in the two conditions that gave information about the target's racial attitudes, participants rated her above the midpoint on the scale measuring their perceptions of her as racist, while in the condition that did not provide any information, she was rated below the midpoint).

Still, even though the target's attitudes were not as blatant, the target was more similar (at least in terms of her gender) to a majority of participants, and she was better

liked than the target in Study 1, only limited support for hypotheses was again found. For example, the first hypothesis – that taking the target's perspective would lead to greater merging and positive regard, regardless of the target's attitudes – only received partial support. That is, while participants in the "control" perspective taking condition (where no racial attitude information was present) regarded the target somewhat more positively and felt somewhat more merged with her than the objective condition (where the target's racial attitudes were known), participants in this condition also differed from the perspective taking condition where the target's racial attitudes were known, and no differences in positive regard or merging were found between the objective condition and the "racist target" perspective taking condition. In other words, the greater positive regard and merging that both were a result of perspective taking were limited to a target who was not racist (or at least, should not have been perceived as such).

For the hypothesis (Hypothesis 2) that participants who took the perspective of the racist target would write essays that had more racist content than participants who wrote about the same target, but did so from an objective perspective (or who took the perspective of a non-racist target), clear support was again found. Participants who took the perspective of a subtly racist target wrote essays that evidenced significantly more racist content on average and used racial labels (e.g., "black") more often than participants in both other conditions. This shows again that at least some participants were attending to the information about the target's attitudes toward blacks and using this information to retrieve ideas from their own minds about what a subtly racist woman might be like.

As in Study 1, this finding for racism expressed in perspective takers' essays does not mean that participants' *own* attitudes were more racist, but the finding is still important to consider, because it suggests several things. First, it suggests that participants, when presented with the task of temporarily "becoming" a racist, could effectively do so – which means that information about racist attitudes and stereotypes of blacks was activated in participants' minds, an activation which might influence subsequent interactions with or perceptions of other people. Second, the information about the target's attitudes in Study 2 did not paint her as an overt racist; she was just someone who admitted to feeling uncomfortable around blacks, and who didn't want to date a black man. Even so, some perspective takers still chose to organize their essays around racial themes, rather than using other information available about the target in the vignette (information that served as an organizing theme for other participants, particularly those in the objective condition). Thus, participants who took the racist target's perspective appeared to find this information particularly salient in comparison to other information which was provided (i.e., which described her as a "nice" person).

For the third hypothesis, that perspective takers would take on the attitudes of the target as their own (Hypothesis 3), there was again (as in Study 1), no support for a direct effect of perspective taking across several measures of racial attitudes (including two IATs, a relatively indirect measure of stereotype activation, and two explicit measures of participants relative preference for blacks over whites). In fact, on one explicit measure (i.e., the relative thermometer that measured feelings of "coolness" toward blacks relative to whites), participants who took the perspective of a racist target or wrote about this

target from an objective perspective showed slightly *lower* relative coolness – feeling slightly *warmer* toward blacks than they did toward whites – than participants who wrote about the same target, but didn't have access to the target's attitudes. While counter to hypotheses and inconclusive, such a pattern of results would be consistent with the possibility that after thinking about a racist target, participants may become especially aware of their own attitudes toward blacks, and it may remind them of their own egalitarian goals.

Related to this, an interaction of condition with motivation to respond without prejudice was again found, although the effect was not the same as in Study 1. Rather than greater internal motivation predicting greater stereotype IAT scores for perspective takers, which was the case in Study 1, in Study 2, greater *external* motivation predicted greater relative coolness toward blacks relative to whites. Perhaps most striking about this interaction is the fact that participants who were *lower* in external motivation to respond without prejudice were substantially lower than other conditions in relative coolness toward blacks. In fact, participants who were one standard deviation below the mean on external motivation showed, on average, close to a standard deviation less coolness toward blacks than toward whites.

In Study 1, the effects of internal motivation more clearly represented the possibility of a rebound effect (e.g., Macrae et al., 1994) unique to perspective takers. After taking the perspective of an obviously racist target, perspective takers who were relatively high in internal motivation to respond without prejudice – which, according to Devine, Plant, Amodio, Harmon-Jones, and Vance (2002) represents people who are

relatively high in self-determination and have internalized reasons for being nonprejudiced – had relatively high automatic stereotyping of blacks. That is, internalized concerns about wanting to not appear prejudiced may have been associated with suppression of stereotypes about blacks for perspective takers only, which in turn was correlated with higher automatic stereotyping.

In Study 2, however, external and not internal motivation was related to perspective takers' attitudes, and the measure of racial bias was directly rather than indirectly measured. Because rebound effects are typically thought to reflect increased accessibility of automatically activated associations after an attempt to suppress stereotypic content (e.g., Macrae et al., 1994; Wyer, 2007), it is hard to argue for suppression and rebound effects in Study 2. Consistent with work by Devine et al. (2002; Plant & Devine, 1998), which shows that greater external motivation to respond without prejudice is associated with higher explicit prejudice, one possible explanation for this particular finding in Study 2 is that perspective takers who were lower in external motivation simply were lower in prejudice, generally. After taking the perspective of even a subtly racist target, these participants may have wanted to emphatically express that even if the target was, *they* were not prejudiced. Therefore, expressing a preference that favored blacks over whites may have reflected anger at their ingroup because of continuing racism (e.g., Finlay & Stephan, 2000). It may have even reflected their own guilt over thinking in racist ways on behalf of the target, which would go against their generally egalitarian goals (e.g., Devine et al., 2002).

Still, important to consider is that this effect was present only for perspective takers. That is, there is something about being "inside" the mind of the target (even when the target is an imaginary person) versus viewing the target from the "outside" that prompted participants to adjust their own attitude expression – whether this adjustment was in favor of blacks or in favor of whites. As discussed in Study 1, this suggests that perspective takers might have been experiencing an internal conflict where *as the target* they express *the target's* prejudice – but simultaneously try to inhibit their own prejudice, at least after the essay task. Those participants who wrote about the target from an objective standpoint or took the perspective of a non-prejudiced target would not have these conflicts.

It is difficult to know exactly why external rather than internal motivation was in play in Study 2, and why it affected a more controllable attitude measure as opposed to a less controllable one. One admittedly speculative possibility may have to do with the fact that the target's quite overt attitudes in Study 1 may have caused almost all participants to make responses on controllable measures that were as non-racist as possible (i.e., a floor effect), but that on a relatively less controllable measure (i.e., the stereotype IAT) the suppressed stereotypes showed the rebound effect. In Study 2, however, because the target was not as loathsome, this might have allowed for some flexibility in response on a controllable measure, but may not have affected automatic responding (or may have affected it, but the effect was obscured due to measurement error or some other issue).

Considering the overall lack of a main effect of condition, as in Study 1, the same question which was asked there is again salient: Why were there no direct effects of

condition on participants' attitudes? Of course, one possible explanation is that perspective taking does not lead to perceivers taking on a target's attitudes. Another possibility is that as in Study 1, where the target's overt racism may have stopped most participants from even trying to understand the target in a deep or complex way, the target in Study 2 may have also caused participants to shy away, even though her attitudes were not as harsh.

Two reasons for this might be posited. First, the target was female, and gender norms for women suggest that they should be warm, kind, and oriented toward others (e.g., Prentice & Carranza, 2002). It might also be important to consider that women are typically expected to be somewhat more egalitarian than men (e.g., Sidanius , Pratto, & Bobo, 1994), considering that women, like blacks, are more often the victims of discrimination. That is, even though the target was portrayed as "nice" and "friendly," the fact that she was violating gender prescriptions for understanding and equality by also being racist may have caused participants to try to distance themselves from her. In effect, the net result of these conflicting pieces of information may have been that she was not seen as a typical woman – that she was in fact not nice at all – which therefore extinguished any desire to grow closer to her (i.e., merge with her). It is not hard to imagine a thought process that was something like "She's a nice woman in some ways, but I don't think she's the type of person I'd want as a friend. I mean, she's a racist!"

Moving Forward, Again

In conclusion, Study 2 (like Study 1) again provided very limited support for the idea that participants who take the perspective of a racist target can come to express

similar attitudes as the target might be expected to express. However, as in Study 1, participants showed evidence of successfully taking the racist target's perspective, with the result that they wrote essays with more racist content when they were asked to write from the target's point of view. Thus, perspective takers are not *unable* to take a racist target's point of view, but they appear to resist merging with her in the way that people who have taken the perspective of less objectionable targets have been show to do. For example, relative to when the target was portrayed as having subtly racist attitudes, perspective takers who had no information about the target's attitudes did feel somewhat more merged with the target, and felt more positive regard for her.

One additional variable that could potentially affect participants' merging with the target is the type of information they have about the target. In the absence of any other information (Study 1), or with conflicting information about how "nice" the target is (Study 2), it seems unlikely that participants would try on their own to come up with any mitigating factors in the target's past which might explain the development of the target's attitudes. Study 3 addresses this possibility by returning to the same target used in Study 1, but changing the way information is presented about the target's attitudes so that it is not quite as vivid as in Study 1 (i.e., wearing a "white pride" t-shirt with "Deutschland" printed on it not once but twice). Furthermore, Study 3 manipulates the information that participants have about the target, giving some participants reasons about why he is racist that might make his attitudes seem less controllable and blameworthy, thus removing some of his responsibility for being racist. In this case, if participants think his attitudes are somewhat more reasonable given his history, this may provide a route for perspective

takers to feel that it is "OK" to like the target or to merge with him (or at least create more variability in responses to these measures), and thus, take on his racist attitudes.

CHAPTER IV

STUDY 3: TAKING THE PERSPECTIVE OF A RACIST WHO HAS REASONS FOR HIS ATTITUDES

Introduction

Studies 1 and 2 failed to find a direct effect of taking the perspective of a racist target on participants' attitudes. This was true when the target was a man, and his overt and extreme racial attitudes were inferred from visual information alone, and it was also true when the target was a woman and her racial attitudes were subtle and presented in conjunction with other information that portrayed her as a generally likable person.

Furthermore, a hypothesis that taking the perspective of a racist target would lead to greater positive regard for the target and greater perceiver-target merging was not supported in either study, even though prior work has shown that perspective taking leads to greater levels of both of these variables (e.g., Batson et al., 2007; Galinsky et al., 2008). Because of this, and because there was no main effect of condition on participants' racial attitudes, the finding that merging (or oneness) mediates the effects of perspective taking on downstream variables was not replicated (e.g., Galinsky & Moskowitz, 2000; Galinsky et al., 2008; Maner et al., 2002). When trying to understand why no direct effects were found, and in order to identify when they *may* be found, considering who the targets were in the first two studies is important.

In Study 1, the lack of main effects of condition on attitudes was probably driven by the strong dislike that participants felt for the target, which may have caused the uniformly low and invariant levels of merging across all conditions. With no additional information about the target, participants probably wanted to distance themselves from him and his beliefs. Unlike a real, multidimensional person, this target existed in a vacuum, where participants knew nothing about him save his appearance (which loudly and unambiguously proclaimed his attitudes). Despite this, perspective takers were willing and able to "get inside his head" as evidenced by their essays, even if doing so was so distasteful that they may have actively resisted expressing attitudes that might even hint at the possibility that they themselves were racist.

In Study 2, participants did like the target more and reported more merging overall than in Study 1, but once again there were no differences in positive regard and merging between perspective takers and objective participants who had access to the target's racial attitudes, and again, no main effects for condition on attitudes were found. One reason for this may have been that the target was inconsistent. That is, the target may have been simultaneously likable and similar (because she was portrayed this way), but also dislikable and dissimilar (because of her attitudes – which although more subtle, still reflected racist undertones).

Study 3 attempts to reconcile issues identified in Studies 1 and 2. The same target used in Study 1 was used again in order to assure that he was viewed as unambiguously racist. However, his racism was portrayed less vividly and less unabashedly, and was presented to participants in a way that might not be as affectively charged. This slight

change might lessen participants' immediate negative reaction to him, allowing a chance for merging to take place as it did not in Studies 1 and 2.

In addition to the change in the way the target was presented, and to promote greater merging and positive regard (or at least, to promote greater variability in these variables than was found in Study 1), information was presented to some participants which might allow them to "forgive" the target somewhat for his attitudes. This inclusion of information might also provide another route by which participants might feel they have "permission" to merge with him – or at least to not as actively resist merging with him. Because participants knew little about the target in Study 1 beyond his attitudes toward blacks (they only saw a picture of him, in which he may have been seen as a skinhead due to his appearance and the fact that he wore a t-shirt espousing racist values), they had no reason to value this person or to merge with him. Given that having prejudiced or stereotypical attitudes toward blacks is not desirable (e.g., Devine, 1989; Monteith, 1993; Plant & Devine, 1998), the target's racial attitudes may actually have been socially stigmatizing, particularly without any context as to why the attitudes existed. That is, wearing a shirt with "white pride" clearly printed on it gives the impression that the target not only has unacceptable attitudes toward blacks, but also that he is not trying to conceal his attitudes in any way (Devine, Plant, Amodio, Harmon-Jones, & Vance, 2002). He is clearly responsible for his attitudes and behavior.

Research on taking the perspective of a member of a stigmatized group (such as people with AIDS or convicted murderers) is not new, although this research has typically focused on changing attitudes toward the target and the target's group (Batson

et al., 1997; 2002), not changing attitudes towards other groups. An important variable considered in this research was the responsibility of the stigmatized person for his or her situation. For example, in one experiment (Batson et al., 1997, Experiment 1), a significant interaction between responsibility and experimental condition emerged, showing that participants who took the perspective of a person with AIDS felt much more empathy for her when she was not personally responsible for contracting the disease. In another study reported in the same paper, this interaction was not significant. Still, even though no further effects which involved responsibility were found, an examination of the means in each study shows that the results were always at least in the direction of demonstrating an influence of responsibility on perceivers' attitudes toward the targets' groups.

Other research (Weiner et al., 1988) has specifically focused on the controllability and responsibility of targets for their stigmas, and perceptions of those targets based on these factors. The primary finding in this research was that when factors in a target's past were given for his current behaviors (e.g., child abuse, drug addiction), participants thought the behaviors were less controllable, and they also thought the target was less responsible for these behaviors. In addition, when the stigma was associated with less controllable factors, perceivers liked the target more, blamed him less, felt less angry at him, and were more willing to give him assistance.

While the relevant groups in this past research performed concrete behaviors that contributed to their stigmatization (abusing a child or taking drugs), and the current study has more to do with the target's attitudes and not his concrete behaviors, similar effects

may be found. That is, when information is provided that shows the genesis of the target's attitudes, participants may feel that his attitudes are less controllable and that he is less responsible for them.

Specifically, the information provided to participants about the target details his history of conflict with African Americans as he grew up in an area of a large city where blacks outnumbered whites. Furthermore, it describes how the target felt that the negative interactions he had with blacks were due to *their* racial prejudice and dislike for white people. Thus, even though the target is still a member of a strongly disliked group, this potent information might make participants feel as if the target is less responsible for his attitudes, or that his attitudes are more defensible. It might also provide an independent pathway for participants to sympathize with him (i.e., regard him less negatively). As a result of this, participants with information might come to express greater racial prejudice than participants without information.

Considering both of these changes suggests that the effects for perspective taking and information might be relatively independent. Perspective taking should primarily work to increase merging with and positive regard for the target. Information, on the other hand, should be primarily associated with perceiving the target as less responsible for his attitudes and perhaps with regarding the target less negatively. When the two factors are additively combined – that is, when perspective takers perceive that the target is less responsible for his attitudes – positive regard and merging should be at their greatest, as should negativity toward blacks.

Still, it is also possible that the effects of perspective taking and information will combine synergistically such that when participants take the target's perspective but also have information about the genesis of his attitudes, merging, positive regard, and negativity toward blacks will be especially high. In this case, the two factors should interact, again leading to a prediction that the combined perspective taking/information cell will have the highest positive regard and merging and the most negative attitudes toward blacks, with values for each of these variables that would be higher than expected if their effects are only additive.

Study 3 thus set out to do several things to make positive regard, merging, and attitude change more likely for perspective takers. As described above, even though the same target photograph used in Study 1 was used in Study 3, the target was slightly toned down, in that the "white pride" slogan on his shirt was removed. In place of this, participants were simply given information that described the target as an overt racist – which was meant to be a similar but somewhat less vivid (than Study 1) portrayal of his racist attitudes. Other participants were also given the same information about the target's attitudes, but this information was prefaced by background information about the target's life that suggested reasons for his attitudes.

Hypotheses and Design

In a 2 x 2 factorial design, participants either took the target's perspective and were asked to write about a day in his life as if they were the target, or were given no further instruction other than to write about a day in his life. Some participants were given minimal information which presented the target as a racist and others were given

the same information along with information that discussed background factors in the target's past which might have led him to possess racist attitudes.

The expectation was (as in Studies 1 and 2) that perspective taking would lead to greater positive regard for and merging with the target. Most important was the prediction that as a result of valuing the target or merging with him to a greater degree, perspective takers would express attitudes that were more similar to those of the target. As in Studies 1 and 2, perspective takers were also expected to include more racist content in their essays.

It was further predicted that having information about the reasons for the target's attitudes would lead participants to think he was less responsible for these attitudes and to feel that his attitudes were more defensible. A tentative prediction was also advanced that having information would lead to greater positive regard (or less negative regard) for the target and would lead on its own to more negative attitudes toward blacks.

Perspective taking and information could have independent effects or could work together synergistically, resulting in an interaction effect, but regardless, there was an expectation that the cell which represented the effects of both perspective taking and information would have the highest overall merging and positive regard relative to all other conditions, and that participants in this cell would also show the most negative attitudes toward blacks. Furthermore, it was expected that participants in this cell would write essays with the most racist content overall. Finally, the hypothesis that greater internal or external motivation to respond without prejudice might interact with condition to predict participants' racial attitudes was again examined.

Pretest

To pretest the vignettes that were used in the main study received, an independent set of participants (n = 56; 19 male) who did not participate in the main study read one of two short vignettes about the target (and which were later used in the main study). Half of the participants – in the "no information" condition – read a short paragraph that described the target "Jon" (i.e., the man in the photograph) as someone who did not like African Americans and who avoided them whenever possible. It also mentioned that when he could not avoid them, he typically reacted with hostility and disrespect. The other participants read the "information" condition vignette, which provided the same information, but prefaced it by providing information about Jon's history (e.g., growing up in a predominantly black neighborhood as a child and getting into many fights with black kids, sometimes ending up in the hospital). Participants received extra class credit for completing the assignment at home and later returned it to the experimenter.

Participants were then asked (on 7-point scales where higher numbers indicate greater agreement) whether the target's attitudes were defensible, reasonable, and controllable, whether the target should be blamed for his attitudes, whether he was responsible for his attitudes, whether he wants to have those attitudes, and whether his attitudes are a result of his conscious choices. The defensible, reasonable, and blamed (reversed) items made up one scale (Defensible, $\alpha = .76$); the want, responsible, and choice items made up another scale (Responsible, $\alpha = .70$); and the controllable item did not correlate well with other items.

Results showed that when participants had information about why the target was racist, his attitudes were seen as more defensible (M= 3.38) and he was seen as less responsible for them (M= 4.82) than when participants had no information about the target (defensible $M = 1.42$; responsible $M = 5.54$), ts (54) = -6.75, 2.11, respectively, ps < .05. Although information did not significantly impact ratings of controllability ($p = .16$), there was a trend for participants with information ($M = 5.0$) to think the target's attitudes were less controllable than participants who had no information ($M = 5.68$).

Method

Participants

In the main study, participants were 101 undergraduate students (66 female, 35 male) from the University of Oregon who participated in partial fulfillment of course requirements. Participants were asked about their racial/ethnic identity and could check multiple categories. Seventy-five identified as white, 1 as black, 15 as Asian, 4 as Latino/Latina, 1 as Middle Eastern, 1 as Native American, and 5 as "mixed" or "other." Ages ranged from 17 – 38 ($M_{age} = 20.00$, $SD_{age} = 2.69$). Sixteen participants reported that English was not their first language, and reported speaking English for 3 – 16 years ($M = 7.73$; analyses conducted with these participants excluded did not substantively differ from those where they were included, so all reported analyses included these participants). The one participant who self-reported as black was excluded from further analyses, as the primary hypotheses were concerned with attitudes toward African Americans.

Procedure

Procedures were exactly the same as in Study 2, with two exceptions. The first is that the additional "impression formation" task (i.e., the "Donald story") was not used in Study 3. In addition, the specific instructions participants received prior to writing about a day in the life of the target were changed such that after getting the photograph of the target, participants were either asked to write about a day in the life of the target "as if you were that person" ("perspective taking") or to simply write about a day in his life without further instruction. Prior to writing, participants were presented with vignettes (which were identical to those used in pretesting) that provided either limited information and simply stated what the target's attitudes were ("no information"), or also provided background information about the target's earlier history with blacks that suggested *why* he felt negatively toward African Americans, followed by the same information about the target's attitudes ("information"; Appendix C gives the full instructions for the no information and information conditions). Following this, participants responded to questions about the essay task and their perceptions of the target, completed several explicit measures (described below), and provided demographic information. They were then told that they were done with the first study, but that they would be debriefed after completing the second study.

The second study was again presented as being related to hemispheric dominance, attention, and rapid categorization, and again used separate informed consent forms. After completing the two filler tasks (which, as in Study 2, were a tracing task with participants' non-dominant hands and a scale measuring need for cognition) participants

completed the same two IATs used in Study 1 in counterbalanced order. After completing the IATs, participants were probed for suspicion and fully debriefed. Although most participants appeared aware that both studies had something to do with race, no participant reported awareness of the connection between the two studies.

Measures

As in Studies 1 and 2, several questions were asked about participants' perceptions of the essay task and the target. Similar to Study 2, a question that asked about participants' similarity to the target was correlated with the four items measuring positive regard for the target (i.e., liking and caring about the target, wanting to spend time with the target, and the target's friendship potential), so all five items were again aggregated ($\alpha = .78$) to form a composite measure of positive regard for the target ("positive regard"). Merging was again measured by two items ("we" and IOS variables), which were aggregated to form the "merging" measure ($r = .58$, $\alpha = .73$). Perception of the target as racist was measured with the same two items as were used in Studies 1 and 2, which were aggregated as a measure of "target racism" ($r = .46$, $\alpha = .63$). Effort spent trying to understand the target was also measured with the same five items used in Studies 1 and 2, which were aggregated to form an "effort" variable ($\alpha = .83$). Last, a single item measure of the essay writing task as "difficult" was again included.

As in Studies 1 and 2, essays were also coded for content by four trained coders who were blind to condition and hypotheses (coder reliability was adequately high; αs ranged from = .75 to 1.0 across different codes). A first code, serving as a manipulation check of whether participants linguistically adopted the target's perspective

(and used in Study 2), was simply a check of whether essays were written in the first person. Several codes were again counts of neutral and negative racial category labels for African Americans, and of African American stereotypes. Again, other codes were subjective, including a code similar to one used in Study 1 but not Study 2, which assessed how stereotypically the target was portrayed ("target stereotypicality;" 0 = "very little or no use of stereotypes," 1 = "some limited use of stereotypes," 2 = "a moderate amount of stereotypes," and 3 = "a substantial amount of stereotypes/target is portrayed as very stereotypical"), as well as codes used in Studies 1 and 2, which measured the extent to which the target was portrayed as a racist, the target's attitudes toward blacks (which could be either positive or negative), and the imagination and creativity of participants' essays ("essay imagination").

As was the case in Study 1 but not Study 2, participants' essays demonstrated substantial variability in neutral references to race, use of negative black stereotypes, target prejudice, and the code representing the target's attitudes toward blacks. Each of these four variables were highly correlated, and were therefore standardized and aggregated to form an overall measure of "essay racism" ($\alpha = .85$).

As in Studies 1 and 2, participants completed a prejudice IAT and a stereotype IAT (because the use of faces in Study 2 did not seem to impact participants' implicit prejudice, the same "name" IAT used in Study 1 was again used here). IAT scores were computed using the *D* statistic, where higher scores indicate greater association of blacks with negative stereotypes or unpleasant words, relative to whites. Participants also completed two feeling thermometers which assessed their feelings about African

Americans and Latinos/Latinas (0 = maximum "coolness" and 100 = maximum "warmth). Scores on the thermometers were reverse-coded so that higher numbers represented greater coolness. Finally, participants again completed scales measuring their internal (α = .69) and external (α = .73) motivation to respond without prejudice.

Results

Descriptive Statistics

Table 6 (below) provides descriptive statistics for all variables which were retained. Possible score ranges for each variable are also provided.

Table 6.
Descriptive Statistics for Study 3 Measures

Variable	α (*r*)	*M (SD)*	Possible Ranges
Positive Regard	.78	2.19 (.94)	1 to 7
Merging	.73 (.58)	1.59 (.77)	1 to 7
Target Racism	.63 (.46)	6.25 (.95)	1 to 7
Effort	.83	4.74 (1.18)	1 to 7
Essay Difficulty	-	3.81 (1.52)	1 to 7
Target Stereotypicality	-	1.55 (.75)	0 to 3
Essay Imagination	-	1.51 (.68)	0 to 3
Essay Racism	.85	.00 (.76)	-
Prejudice IAT	-	.34 (.31)	-
Stereotype IAT	-	.38 (.31)	-
Black Thermometer	-	33.18 (15.76)	-100 to 100
Latino/Latina Thermometer	-	29.06 (18.51)	-100 to 100
Internal Motivation to Respond without Prejudice	.69	7.21 (1.35)	1 to 9
External Motivation to Respond without Prejudice	.73	5.27 (1.65)	1 to 9

Data Examination

Prior to conducting primary analyses, distributional characteristics of all variables were examined. All variables appeared to be normally distributed, with the exception of the merging variable, which evidenced slight positive skew. A square root transformation normalized this variable, and subsequent analyses tested both the untransformed and square root transformed merging variable. None of these analyses led to substantively different conclusions, so results using the untransformed variable are reported.

Primary Analyses

As discussed above in the Introduction, the primary hypotheses for Study 3 were that perspective taking would lead to greater positive regard for and merging with the target (Hypothesis 1), greater expression of racism in participants' essays (Hypothesis 2), and attitudes which reflected more negativity (or less positivity) toward blacks (Hypothesis 3). Furthermore, it was expected that information (which in a pretest was associated with perceptions that the target was less responsible for his attitudes and that his attitudes were more defensible) might lead on its own to greater positive regard for the target and to more negative attitudes toward blacks. These two effects (i.e., of perspective taking and information) could be independent and additive, or they could have an interactive effect, although in either case, the cell which represented perspective taking combined with information was expected to have the greatest positive regard and merging, and the least positive attitudes toward blacks.

To test these hypotheses, each factor was examined in a traditional analysis of variance (ANOVA). Sex was also considered as an additional factor. In no analysis did

sex emerge as a significant predictor, nor did it interact with any other variables, so it is not discussed further.

Manipulation Check. A first test focused on whether participants complied linguistically with the perspective taking instructions (i.e., wrote their essays in the first person). Analyses indicated that 45 of 52 (86.5%) participants in the combined perspective taking conditions and only 5 of 48 (10%) participants in the combined no perspective taking conditions wrote their essays in the first person, $\chi 2$ (1, $N = 100$) = 57.85, $p < .0005$, indicating overall compliance with perspective taking instructions.

Perceptions of the Target and the Essay Task. Analyses were next conducted on the "effort," "difficulty," and "racist target" questions. As expected, no significant effects emerged from these analyses, showing that participants in all conditions put in equal effort in trying to imagine the target, found the essay task equally difficult, and found the target equally racist.

Hypothesis 1 – Potential Mediators. The next tests focused on hypotheses concerning the positive regard and merging variables. A marginally significant effect for perspective taking on positive regard was found, F (1, 96) = 2.88, $p = .09$, partial $\eta^2 = .03$, as was a significant effect of information, F (1, 96) = 8.86, $p = .004$, partial $\eta^2 = .09$. The interaction was not significant ($F < 1$, $p = .75$). As can be seen from an examination of the means, the perspective taking/information cell (PT/INFO; $M = 2.58$, $SE = .18$) had the highest positive regard for the target, followed by the no perspective taking/information (NO PT/INFO; $M = 2.33$, $SE = .19$), perspective taking/no

information (PT/NO INFO; $M = 2.10$, $SE = .18$) and no perspective taking/no information cells (NO PT/NO INFO; $M = 1.74$, $SE = .18$).

For perceiver-target merging (or, considering the actual means here, and as discussed in Study 1, *lack of distancing* from the target), there was a significant effect of perspective taking (F (1, 96) = 3.91, $p = .05$, partial $\eta^2 = .04$), no effect of information ($F = 1.85$, $p = .17$, partial $\eta^2 = .02$), and no interaction ($F < 1$, $p = .87$). Examination of the cell means showed that the PT/INFO cell was highest in merging ($M = 1.85$, $SE = .15$), followed by PT/NO INFO ($M = 1.62$, $SE = .15$), NO PT/INFO ($M = 1.52$, $SE = .16$), and NO PT/NO INFO ($M = 1.34$, $SE = .15$) cells.

Taken together, these two findings help clarify how perspective taking and information differentially (and independently) affected perceptions of the target. That is, information on its own impacted participants' positive regard for the target (i.e., participants *disliked* him less than when information was present), perhaps because the information itself portrayed a person who participants felt somewhat sorry for, considering his history. Information did not affect participants' perception of feeling psychologically merged with the target.

Perspective taking, however, impacted participants' willingness to report not being *completely* distanced from the target, and only had a small and marginally significant impact on positive regard. This suggests that under the right conditions, participants can feel somewhat merged with even a dislikable target, even though they do not like him much more than perceivers who have not taken his perspective.

Next, analyses focused on whether positive regard or merging was related to any of the variables which assessed participants' racial attitudes. Unlike the findings of Studies 1 and 2, analyses revealed a marginal effect of perspective taking on positive regard and a significant effect on merging. Furthermore, information also had a significant effect on positive regard. Therefore, if either positive regard or merging (or both) predicted participants' attitudes then either or both of these variables might be considered further as mediators – that is, if perspective taking or information also had an effect on the same attitudes. If no direct effects of perspective taking or information were found on participants' attitudes, the possibility would still exist that that perspective taking or information would indirectly affect participants' racial attitudes through positive regard or merging.

Analyses showed that greater positive regard (t (98) = 2.86, p = .005, β = .28) and merging (t (98) = 3.78, p < .0005, β = .36) both predicted greater coolness toward blacks, and could thus both be considered further. Neither positive regard nor merging was significantly associated with any other measures of participants' racial attitudes.

Hypothesis 2 – Essay Racism. Similar to Studies 1 and 2, it was hypothesized that perspective taking would be associated with the greatest expression of racism in participants' essays. Other essay codes were also examined, such as the target's stereotypicality and the imaginativeness of the essays. Starting with the latter first, as expected (and also as found in Studies 1 and 2), no differences in essay imagination were found. For target stereotypicality, however, a main effect of information was found,

F (1, 96) = 9.47, p = .003, partial η^2 = .09, showing that when information was present, participants described the target in less stereotypical ways. No effects for perspective taking emerged, nor was there an interaction effect (Fs < 1, ps > .76). An examination of means showed that the PT/INFO condition wrote the least stereotypical essays (M = 1.29, SE = .11), followed by the NO PT/INFO (M = 1.35, SE = .17), PT/NO INFO (M = 1.75, SE = .17) and NO PT/NO INFO (M = 1.78, SE = .13) conditions. Without information about the reasons for the target's attitudes, participants probably used the only information they had available to construct their narratives – stereotypes of what a racist might be like.

For racist content in participants' essays, there was a main effect of perspective taking, F (1, 96) = 14.74, p < .0005, partial η^2 = .13. A main effect of information was also found, F = 6.32, p < .01, partial η^2 = .06. The interaction was marginally significant (F = 2.85, p = .10). An examination of the cell means revealed that participants in the PT/INFO (M = .31, SE = .17) cell wrote essays with the greatest racist content, followed by the PT/NO INFO (M = .19, SE = .10), NO PT/INFO (M = .01, SE = .15), and NO PT/NO INFO (M = -.57, SE = .13) cells.

Thus, as expected (and as found in Studies 1 and 2), perspective takers wrote essays with significantly more racist content than participants who did not take the target's perspective, showing that perspective takers were again able to role-play a racist target, accessing and expressing on behalf of the target their own knowledge of African American stereotypes and other content related to racial attitudes. Interestingly, the effects of information alone also led to somewhat higher essay racism. This may reflect

that when participants learned of the target's history with African Americans, and became aware of the fact that he had experienced many conflicts with members of this racial group, they assumed that he would act in particularly racist ways, which also led them to access and express racist content which was used to describe the target. Taken together with the finding above on stereotypicality, this suggests that perspective takers, particularly when information about why the target held racist attitudes was present, went beyond merely stereotyping the target and tried to construct realistic narratives about a day in his life.

Hypothesis 3 – Participants' Racial Attitudes. Four variables were considered here: the prejudice and stereotype IATs, and the African American and Latino/Latina thermometer ratings. Similar to Studies 1 and 2, no direct effect of condition was found for either IAT. In addition, no direct effect of condition was found for the Latino/Latina Thermometer.

An ANOVA examining participants' expression of coolness toward blacks showed a main effect of information, F (1, 96) = 3.87, p =.05, partial η^2 = .04, but no effect for perspective taking, F (1, 96) = 2.04, p = .15, partial η^2 = .02, although the latter finding was nonetheless in the hypothesized direction with non-significantly greater coolness for perspective takers. The interaction was not significant ($F < 1, p > .70$). Examining the means showed that PT/INFO (M = 32.50, SE = 3.41) had the greatest coolness toward blacks, followed by NO PT/INFO (M = 27.61, SE = 3.62), PT/NO INFO (M = 25.73, SE = 3.41), and NO PT/NO INFO (M = 20.68, SE = 3.48).

Tests of Indirect Effects. The results above show that information directly impacted participants' racial attitudes while perspective taking did not. Furthermore, information (and to a lesser extent, perspective taking) also had an effect on positive regard, which was associated with participants' racial attitudes. Similarly, perspective taking had an effect on merging (i.e., less distancing from the target), which also was associated with participants' racial attitudes.

Because of this, additional tests were conducted examining the main effects of information and perspective-taking condition on participants' coolness toward blacks (the interaction of perspective taking and information never approached significance for positive regard, merging, or coolness), controlling for positive regard and merging. This would allow for two important tests. First, by controlling positive regard and/or merging, a test of whether the main effect of information on attitudes would be reduced to insignificance is possible (i.e., at least partial mediation). Second, even though perspective taking did not have a significant and direct effect on attitudes, the possibility exists that perspective taking influenced attitudes *indirectly* through either positive regard or merging (i.e., that even lacking a direct effect on attitudes, the indirect effect of perspective taking would be significant through one of these variables).

The first of these tests used multiple regression with perspective taking and information entered as effects coded predictors of coolness (i.e., exactly the same as a main effects analysis in ANOVA, minus their interaction), with positive regard and merging also entered as predictors in separate models (i.e., analyses of covariance; ANCOVAs). Multiple regression analyses were used so that beyond tests of significance,

coefficients for each predictor could be examined and reported. A first test reproduced the ANOVA from above, in order to generate standardized coefficients for perspective taking ($\beta = .14$) and information ($\beta = .20$) on coolness toward blacks, which might be compared to those from the ANCOVA analyses.

Next, positive regard – which was significantly predicted by information and marginally predicted by perspective taking – was added to the model. This analysis revealed that when controlling for positive regard, neither perspective taking (t (96) = 1.07, $p = .29$, $\beta = .11$) nor information (t (96) = 1.30, $p = .20$, $\beta = .13$) approached significance in predicting attitudes. The coefficients for each of the condition variables was reduced, while the effect of positive regard was significant, t (96) = 2.17, $p = .03$, $\beta = .22$. This suggests that there was an indirect effect of the condition variables on attitudes through positive regard for the target.

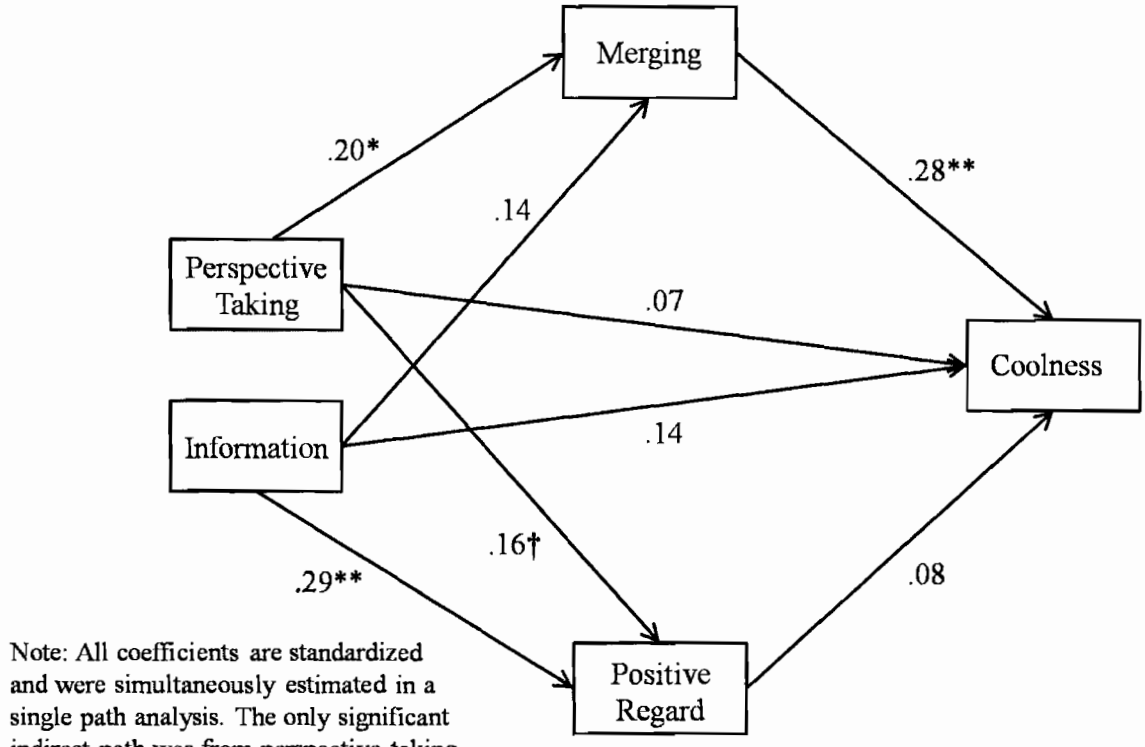
The next test focused on the influence of merging as a covariate. This analysis also revealed that controlling for merging, neither perspective taking (t (96) = .83, $p = .41$, $\beta = .08$) nor information (t (96) = 1.60, $p = .11$, $\beta = .15$) significantly predicted attitudes. Merging, on the other hand, was significant (t (96) = 3.31, $p = .001$, $\beta = .32$), suggesting that in addition to positive regard, there was also an indirect effect of the condition variables through merging with the target.

A final test examined the effects of the condition variables and both covariates in a single model. This was done to see whether positive regard or merging would prove the stronger covariate, and thus most clearly represent the indirect effect of condition on attitudes. Although this analysis could be done in regression, a path analysis would allow

for 1) simultaneous modeling of all effects, where each effect is controlling for all others in the model, and 2) a significance test of any indirect effects. Because of this, a path analysis was modeled using AMOS 6.0 (Arbuckle, 2006). In this model, perspective taking and information main effects (effects codes) were each used to simultaneously predict positive regard, merging, and coolness toward blacks, while positive regard and merging were also predictors of coolness. This analysis revealed (see Figure 3, below) that perspective taking significantly predicted merging ($\beta = .20$, $p < .05$) and marginally predicted positive regard ($\beta = .16$, $p = .08$), but did not directly predict coolness ($\beta = .07$, $p = .44$). Information significantly predicted positive regard ($\beta = .29$, $p = .003$) but did not predict either merging ($\beta = .14$, $p = .17$) or coolness ($\beta = .14$, $p = .17$).

Turning to the effects of the covariates, positive regard did not predict coolness ($\beta = .08$, $p = .41$), but merging did predict coolness ($\beta = .28$, $p = .003$). Most importantly, significance tests of the indirect effects revealed that the only significant indirect path was from perspective taking to coolness ($p = .05$). Because merging was the only variable which significantly predicted coolness, the indirect effect was primarily through merging, and not positive regard. In summary, in a model which simultaneously used perspective taking, information, positive regard, and merging to predict participants' racial attitudes, only merging continued to directly predict these attitudes, and the only significant effect of condition was an indirect effect from perspective taking, through merging, to more negative attitudes toward blacks. These findings are discussed in more detail below, in the Discussion section.

Figure 3
Path Model of Effects of Condition on Coolness toward Blacks.



Note: All coefficients are standardized and were simultaneously estimated in a single path analysis. The only significant indirect path was from perspective taking to attitudes, $p = .05$. † $p < .10$, * $p < .05$, ** $p < .01$

Motivation to Respond without Prejudice. To examine whether condition interacted with internal or external motivation to respond without prejudice, a set of dummy codes that represented a test of the perspective taking/information cell against each of the other cells was first computed. Interaction terms were created by multiplying these dummy codes by the centered moderator variables, and the variables were tested in sets using multiple regression analyses as was done in Studies 1 and 2 (i.e., a first step entered all dummy codes and the centered moderators, a second step entered interactions

of dummy codes with the moderators, and a third step also tested the three-way interaction terms), with significance of the interactions based on changes in R^2.

For the two IAT variables, no significant moderated effects were found, so these variables are not discussed further. For the African American Thermometer, no interaction effects were found, although external motivation to respond without prejudice significantly predicted greater coolness toward African Americans (t (95) = 2.75, p = .007, β = .26) and internal motivation to respond without prejudice significantly predicted less coolness (t (95) = -4.07, p < .0005, β = -.37). For the Latino/Latina thermometer variable, again no interaction effects were found, although internal motivation to respond without prejudice again significantly predicted less coolness toward Latinos/Latinas, t (94) = -4.93, p < .0005, β = -.45. External motivation did not significantly predict coolness toward Latinos and Latinas.

Thus, unlike Studies 1 and 2, motivation to respond without prejudice did not significantly moderate the effects of condition on participants' racial attitudes. This lack of effect is explored below, in the Discussion.

Discussion

Study 3 contributed several important tests of the hypotheses under consideration in this dissertation by showing that taking the perspective of a racist target can indirectly lead participants to express attitudes that are similar to those of the target. First, unlike Studies 1 and 2, perspective taking led to significantly greater merging with (or less *distancing* from) the target and marginally higher positive regard for the target. This was true even though overall, levels of positive regard and merging were fairly low (as in

Studies 1 and 2). More importantly, and consistent with other models of perspective taking (e.g., Cialdini et al., 1997; Galinsky et al., 2005; Maner et al., 2002), perceiver-target merging in Study 3 was an important intermediary variable. That is, even though the direct effect of perspective taking on attitudes was not significant, there was a significant indirect effect of perspective taking, through merging, on attitudes. Also important is that the effects of condition – whether direct or indirect – did not simply reflect a general outgroup bias because attitudes toward Latinos and Latinas were not affected by condition.

Furthermore, when both merging and regard (which was also associated with attitudes toward blacks) were entered together along with condition as predictors of coolness, merging had a unique and significant effect above and beyond that of regard, while the effect of positive regard did not approach significance. That is, a model was tested where perspective taking and information were entered as simultaneous predictors of positive regard, merging, and attitudes, and where positive regard and merging were also predictors of attitudes. In this model, the only significant predictor of attitudes was merging, and the only significant effect of condition on attitudes was the indirect path from perspective taking to attitudes, primarily through merging.

Although there are caveats here, such as the lack of a significant direct effect of perspective taking on attitudes, this finding does provide some further evidence that the effects of perspective taking on behavior can, at least in some cases, be explained in part by a sense of merging with the target. For example, Maner et al. (2002) found that "oneness" mediated the effects of perspective taking on helping, while empathic concern

and personal distress did not. Galinsky et al. (2008) also found that lower self-ratings of intelligence (which was considered evidence of cognitive merging with a cheerleader target) but not liking for the target provided an indirect path for the effects of trait perspective taking on lower analytical task performance. Similarly, when both merging and positive regard were simultaneously examined here, merging continued to have an effect on attitudes, while positive regard did not.

The fact that no effects were found for implicit tests (IATs) in this study could reflect many things. Several researchers have found that explicit and implicit tests do not always converge. For example, Payne et al. (2008) found that the lack of structural fit between measures of implicit and explicit attitudes is one reason that rather low correlations between these measures are found. When structural fit was low (i.e., the methods that implicit and explicit tests used to measure racial attitudes were not similar to one another), the correlations between explicit and implicit measures decreased to the point where they were essentially measuring different things. The authors further detailed how explicit measures involve participants reading a phrase or sentence and using a numerical scale to express their evaluative responses. In the case of a feeling thermometer, which was the explicit measure of racial attitudes in the current study, participants were asked to give a global evaluation in numeric terms for an entire group. The implicit measures used here, on the other hand, were stereotype and prejudice IATs, where participants rapidly classified words that were symbolic referents to group categories, stereotypes, or positive and negative words, and their attitudes were measured

by the relative speed of responding during dual categorization blocks. This certainly demonstrates a lack of structural fit.

Other researchers have argued that cognitive consistency plays a role in the lack of fit between implicit and explicit measures (Gawronski et al., 2008; see also Devine, 1989), that implicit and explicit attitude tests measure essentially different things (e.g., Greenwald & Banaji, 1995), or may measure separate but related constructs, with the true relationship obscured because of low reliability in measurement (e.g., Cunningham et al., 2001). In truth, there is much debate over exactly what implicit measures such as the IAT are actually measuring (e.g., De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2009; Fazio & Olson, 2003), even though a recent meta-analysis of IAT research showed that it has substantial predictive validity across many domains, such as predicting interracial intergroup behavior (Greenwald, Poehlman, Uhlmann, & Banaji, 2009). So it is not particularly surprising that implicit and explicit attitude measures diverged here in terms of capturing the effects of perspective taking and information on attitudes.

Turning to the hypothesis that predicted greater expressions of racism in participants' essays, Study 3 again found that perspective takers wrote essays with more racist content than other participants (replicating an effect found in Studies 1 and 2). Again, as was discussed in Studies 1 and 2, there was no evidence that perspective takers *themselves* became more racist as a result of perspective taking, but it is nonetheless telling that this effect was present. This will be discussed further in the General Discussion.

A few other findings that emerged from Study 3 are worthy of mention. First, although Studies 1 and 2 both found evidence that variables measuring some form of motivation to respond without prejudice were associated with somewhat more negative attitudes toward blacks (i.e., internal motivation to control prejudice was related to more automatic stereotyping in Study 1; external motivation to control prejudice was related to coolness toward blacks relative to whites in Study 2), in Study 3, neither measure of motivation to respond with prejudice was associated with either indirect or direct measures of racial bias. Although no definitive explanation for this discrepancy is known, one possibility is related to the fact that perspective taking led to somewhat greater merging in Study 3 while in Studies 1 and 2, this was not the case. That is, the way that information about the target was presented in Study 1 (i.e., visually, via the "white pride" slogan on the target's shirt) – which proclaimed that his attitudes were extreme and that he was not ashamed of holding them – may have been affectively "charged," leading participants to try and consciously suppress any negative thoughts about blacks after taking his perspective. In Study 2, even though the target's attitudes were less extreme and obvious (and thus, less "charged") than in Study 1, the very similarity and likability of the target might have made perspective takers feel particularly uncomfortable and aware of their own racial attitudes after the essay writing task. While speculative, the fact that perspective taking did not lead to greater merging with the target in Study 1 or 2 provides some support – perspective takers did not want to admit feeling like they were "one" with either target.

An examination of the targets used in Studies 1 and 2 further suggests a scenario by which the interactions between condition and motivation to respond without prejudice may have occurred for different reasons in Studies 1 and 2. In Study 1, immediate dislike for the target may have led to an attempt to suppress stereotypes after taking his perspective, which may have led to rebound. In Study 2, where the target was liked somewhat more and was probably not considered as racist as the Study 1 target, perspective takers might not have had as many thoughts that needed to be suppressed. This could be true even though participants tried to distance themselves from the target after the essay task – perhaps even *because* the target was similar in many ways. For example, it is easy to imagine a participant thinking "This woman is kind of similar to me – my friends come to *me* for advice too, and I like to laugh – but *she* is a racist, and *I* most definitely am *not*!" That is, the target might have been "too close for comfort," inducing a need for participants to reaffirm their own egalitarian values.

In Study 3, however – where perspective taking *did* lead to greater merging (even if overall, levels of merging were still low) – participants might not have felt as strong (or as immediate) a need to distance themselves from the target's attitudes. That is, first of all, the target's attitudes were presented in a relatively neutral way. Participants simply read that he did not like blacks, that he tried to avoid them whenever possible (suggesting that he did not *actively* seek out conflict with them), and that only when he couldn't avoid them did he react with disrespect and hostility. Although this still portrayed him as a racist, it did so in a way that likely provoked less immediate distaste. Furthermore, unlike in Study 2, no inconsistent information about the target was presented. That is, he was not

portrayed as a "generally likable person" who just happened to also not like blacks. Again, this explanation is speculative, but the fact that perspective taking led to greater merging (and to a lesser extent, positive regard) in Study 3 lends some support to the idea that perspective takers were not trying as hard to distance themselves from the target (and presumably, from his attitudes, by suppressing thoughts about prejudice).

One other interesting finding from Study 3 is that information, on its own, led to greater positive regard for the target (or less *dislike* for him). In addition, information provided a direct route to somewhat cooler attitudes toward blacks. Although this effect did not remain significant in a simultaneous model that tested the effects of perspective taking, information, positive regard, and merging on attitudes, the fact that on its own, information influenced participants' attitudes is certainly interesting.

One possible explanation for this is that the information made participants more aware of African American stereotypes concerning aggression. For example, one of the "reasons" provided for why the target did not like African Americans is that he had several run-ins with them when he was younger where he was beat up by them. By being given this information, participants may have been thinking about African Americans in the context of blacks being violent and aggressive, which might have given them "permission" to *think* more negative thoughts about this group, rather than trying to limit their stereotypical thinking. Another possibility is that information made participants feel as if blacks themselves were prejudiced (because the story stated that the target believed that he was attacked because he was white), or made participants feel a form of parallel

empathy with the target (i.e., felt the target's anger over being a "victim" of oppression). In either case, information could have led to less positive sentiment toward blacks.

CHAPTER V

GENERAL DISCUSSION

Does Perspective Taking Transmit Racial Attitudes?

This dissertation set out to examine one central question: Does taking the perspective of a racist target make participants more racist? Of course, this statement is an oversimplification of the issue. Racial attitudes are difficult to change, and any changes that *were* found as a result of perspective taking were likely (and hopefully) very temporary. After three studies where participants learned about a target's attitudes in a variety of ways, and in comparison to a variety of control conditions that were employed, the short answer seems to be "Not directly." Below, the major findings from all three studies will be briefly reviewed, along with a discussion of these effects.

Study 1 Summary. In Study 1, participants learned about the target's attitudes simply by seeing a photograph of this target wearing a "white pride" t-shirt. Some participants took this target's perspective, writing about a day in his life as if they were that person, and control conditions either wrote about the same target without further instruction (except to write their essays in the third person) or wrote about a day in their own lives. To answer an additional question about the effects of perspective taking on behavioral explanations, participants in all conditions of Study 1 were asked to weave three scenarios into their narratives. Two of these scenarios had the potential to be interpreted as consistent with racist attitudes toward blacks, and a third was neutral.

This study provided an important first step in testing the primary question of interest, because prior research upon which this dissertation was in part based (e.g., Galinsky & Moskowitz, 2000) also employed a strategy of having participants "form impressions" of targets using visual information alone. Although Study 1 was rather complicated, and was very likely tiring for participants because they did a large number of tasks, several important effects still emerged, which are discussed below.

First, it should be acknowledged again that condition did not have a direct effect on participants' racial attitudes, which was the primary hypothesis. Furthermore, after taking the perspective of an overtly racist target, participants did not feel more similarity to or positive regard for the target than participants in other conditions. In addition, not only did perspective takers not feel *more* merged with the target than other participants, they actually felt *less* merged (and in any case, not very much merging was found for any participants). To be precise, because of limited variability in merging, this variable was dichotomized to represent whether participants reported any merging at all with the target, and fewer participants than would have been expected by chance expressed any merging in the perspective taking condition, while more participants than would be expected by chance felt merged in a "write about self" condition. That is, the effects for merging were opposite from what would be expected from past research (e.g., Cialdini et al., 1997; Galinsky & Moskowitz, 2000; Galinsky et al., 2000; Maner et al., 2002).

Still, despite the lack of a direct effect of condition on racial attitudes and variables representing similarity, liking, and merging with the target, a few noteworthy and important effects did emerge from Study 1. For example, while condition did not

directly affect participants' racial attitudes, it did affect *some* participants' racial attitudes: participants who were relatively higher in internal motivation to respond without prejudice. Study 1 showed that perspective takers who were relatively high in internal motivation to respond without prejudice ended up having higher scores on a measure of implicit stereotyping (i.e., a stereotype IAT). Although this effect was revealed in an exploratory analysis, it was consistent with prior work that has shown that suppressing stereotypes can lead to post-suppression rebound effects (e.g., Galinsky & Moskowitz, 2000; 2007; Macrae et al., 1994; Wyer, 2007). Although a suppression account cannot be definitely claimed here, there are reasons to believe that suppression is one possible reason for perspective takers' elevated stereotype IAT scores. Further research might investigate this possibility by measuring the extent to which participants are aware of trying to suppress negative thoughts and feelings about blacks after taking the perspective of a racist. Another possibility would be to compare perspective taking to other conditions where participants are asked to suppress stereotypes of blacks after writing about a day in the life of a black target.

Turning to a hypothesis that concerned the expression of racism in participants' essays, Study 1 showed that perspective takers made more neutral references to race and wrote essays with substantially more racist content, where they (for example) integrated a greater number of negative stereotypes about blacks (e.g., hostile, unintelligent) into their essays. This effect, which replicated across all three studies, means that perspective takers, in their roles *as* the target, were thinking about how other people think negatively about blacks and also accessing knowledge about stereotypes of blacks, which was used

to create their narratives. While this racial content might have merely reflected their knowledge of commonly held cultural stereotypes about blacks (e.g., Devine, 1989) and racists, or may have in fact reflected their own beliefs in some cases, it is clear that negative associations of blacks with themes such as hostility and ignorance were made more salient for perspective takers. This is discussed in more detail below, after summaries of Studies 1 and 2.

One final goal of Study 1 was to test whether perspective taking led to explanations of the target's behavior that were more similar to participants' explanations of their own behavior. To test this, all participants included three scenarios in their essays and were later asked to explain the target's behavior during these scenarios. Very little support was found for this hypothesis, beyond the fact that participants who wrote about a hypothetical day in their own lives provided significantly fewer marked beliefs than third-person participants (which might be expected, given prior research, e.g., Malle et al., 2000), with perspective takers between the two but not differing from either.

In all, there was an obvious lack of clear support for the predictions in Study 1, but some hints why. For one, the target was not liked in general by participants. Thus, impression management was very likely an important factor, in that participants probably wanted to distance themselves from the target, regardless of his actions. Further, the relevance of race in the scenarios was confounded with the type of behavior in the scenarios – both of the race relevant scenarios involved a social interaction, while the third scenario involved a straightforward and solitary behavior by the target which

seemed to invite an increase in desire reasons, which are typically not as diagnostic of impression management concerns.

Last, given the clear interpretations of the target's behaviors during the race-relevant scenarios along racial lines (but not the behaviors of the self during the same scenarios in the self-control condition), it seems clear that the very behaviors which were being explained were different in the condition where participants wrote about themselves compared with the two conditions where participants wrote about the target. That is, for participants who wrote about the day in the life of the target, the target's behavior was typically interpreted in ways consistent with a perception of the target as a racist – while people who wrote about a day in their own lives interpreted their own behavior along egalitarian lines, or ignored race as a determining factor. Thus, the behaviors that were being explained were very different in the "write about self" condition compared with the other two conditions, which is a confound that might be addressed in future research.

Study 2 Summary. Because the target in Study 1 was quite overtly racist and generally not well-regarded by participants, Study 2 introduced several changes. First, the target used was female rather than male, in order to make her more similar to the majority of participants in the study. She was also deliberately made to appear more likable, and her attitudes toward blacks were crafted to be much more subtle than in Study 1. Each of these changes was in an attempt to foster positive regard for and merging with the target. In these goals, Study 2 succeeded fully. Participants liked this target more than other targets, found her more similar, and merged with her more.

Despite these changes, no strong effects of condition on merging or attitudes were found. Instead, an interaction between condition and external motivation to respond without prejudice showed that higher levels of external motivation led to greater coolness toward blacks relative to whites for participants in the perspective-taking condition. In addition, this effect was further moderated by internal motivation, such that the effect for external motivation on coolness was stronger when internal motivation was low, but not present when it was high. As was the case in Study 1, the effects of motivation to respond without prejudice were only found for perspective takers, and not for participants in other conditions. Further, they reflected a somewhat similar effect – greater negativity toward blacks among participants who had more desire to *not* appear prejudiced. Perhaps just as important (or even more so) was the *decreased* relative coolness toward blacks for perspective takers low in external motivation to respond without prejudice. Still, the finding in Study 2 was different from what was found in Study 1. That is, external rather than internal motivation was the primary interactant with condition, and the dependent variable (coolness towards blacks) was a more controllable measure of participants' attitudes than in Study 1 (where the measure was an implicit measure of stereotyping).

The finding for external motivation to respond without prejudice in Study 2 also provided additional support for the reasoning that following the essay writing task, even when the target was only "subtly" racist, perspective takers to a greater extent than other participants may have tried to avoid thinking in negative ways about blacks. Although a suppression and rebound account here does not seem as likely, there are parallels between findings from Study 2 and previous work on stereotype suppression which might provide

an interpretative frame with which to understand some of the effects from Studies 1 and 2. For example, Wyer (2007) found that only those participants who 1) previously suppressed stereotypes about a black target (while writing about a day in his life), 2) were high in external motivation, and 3) were low in internal motivation to respond without prejudice showed a stereotype rebound effect on a word-stem completion task. Again, future research might explore this hypothesis in relation to taking the perspective of a racist target.

Study 3 Summary. In Study 3, participants again took the perspective of a racist target and wrote about a day in his life as if they were that person, or simply wrote about a day in his life without further instruction. In addition, participants were either given minimal information about the target outside of describing his attitudes and behavioral tendencies toward blacks (serving as a less extreme version of the target from Study 1), or were given this same information, plus details about his prior history with African Americans.

Information about the target's history was given in an attempt to explain why he held racist attitudes, so that even if participants did not agree with the attitudes, they might feel that these attitudes were less controllable or more defensible, and that he was less responsible for them. If this were true, then participants might be more willing to regard him at least somewhat more positively. In support of these hypotheses, an initial pretest with a separate sample showed that when presented with information about the reasons for the target's attitudes toward blacks, participants thought the target's attitudes were more defensible, that he was less responsible for them, and that they were somewhat

less controllable. In the actual study, participants who had information about the target's reasons for being racist liked him more and showed greater target-congruent effects on their own attitudes toward blacks.

The way that the target's attitudes were presented to participants was also changed in Study 3. Rather than using visual information alone to portray his attitudes, participants were simply told that the target was racist. Furthermore, no inconsistent information was given about the target (i.e., no information was given portraying him as "nice"). These changes were made so that the target would still appear unambiguously racist, but would not be so affectively charged a stimulus. As a result, this might take away participants' need to immediately and completely distance themselves from the target after taking his perspective (i.e., allow them to merge with him to some extent, which was not found in earlier studies). It was also expected that this would lead to more target-congruent attitudes.

As hypothesized, these changes did work to encourage less distancing from the target. That is, although there was no *direct* effect of perspective taking on attitudes, perspective taking did lead to less distancing from the target, which in turn was associated with greater coolness, and the indirect effect of perspective taking on coolness was significant. Specifically, when information (which had a direct effect on attitudes and on positive regard) and perspective taking were used to predict positive regard and merging (which both predicted attitudes), and positive regard and merging were simultaneously used to predict attitudes, the only significant predictor of attitudes which emerged was merging. Moreover, the only significant path from condition to attitudes

was the indirect path from perspective taking to attitudes, through merging. This is consistent with work by other researchers who have found that the effects of perspective taking are mediated by overlap with (and not liking) a target (e.g., Galinsky et al., 2008; Maner et al., 2002).

Study 3 also replicated an effect found in both earlier studies, in that perspective takers wrote essays with more racist content, on average, than participants who did not take the target's perspective. This provides firm evidence that taking a racist target's perspective leads participants to think about what it means to be prejudiced, and to express prejudiced attitudes on behalf of the target and from the target's imagined viewpoint.

Racist Essay Content – What Might It Mean?

Considering the racism in the essays, it might seem easy to write this off as obvious – it is hardly surprising that participants who were explicitly instructed to write about the day in the life of a target "as if they were that person" did just that, expressing attitudes consistent with those the targets might be expected to possess and portraying themselves in a way that the racist targets, given the chance, might be expected to portray themselves. Still, content from participants' essays has been viewed as one piece of strong evidence in support of perspective taking and priming hypotheses, such as the contention that taking the perspective of a group member leads to fewer stereotypes about the group (i.e., the content expressed in the essays is less stereotypic after perspective taking, similar to suppression conditions; e.g., Galinsky & Moskowitz, 2000) or that writing about a target who is a member of a stereotyped group leads to *more* stereotypic

essay content, particularly when written from the target's perspective (e.g., Wheeler et al., 2001). Therefore, the fact that perspective takers consistently wrote essays with more racist content across all three studies is an important fact to consider.

To even write an essay that would be perceived as reflecting the actions of a racist suggests that two conditions were likely met. Consider that the codes representing essay racism across all three studies were a mix of counts of racial labels such as "black," use of negative black stereotypes such as "aggressive," and codes that tapped into evidence that the target was portrayed as a racist who had negative attitudes toward blacks. This suggests that to write essays with racist content, first, knowledge of African American stereotypes needed to be present in participants' minds, as did beliefs about how a racist might feel and act toward blacks. Likely, all participants had access to this information, because, for example, African American stereotypes are widely known (e.g., Correll, Judd, Park, & Wittenbrink, 2002; Devine, 1989). Still, even though *knowledge* of "what a racist is like" was probably equal across conditions, meeting the first condition, perspective takers, more than other participants, met the second condition: expressing that content in their essays.

That they did so to a greater extent than other participants consistently across three studies is informative. That is, participants in other conditions had access to the targets' attitudes and their own knowledge of what a racist would be like, but for perspective takers, this information was accessed and expressed to a greater extent than in other conditions. Even though this reflects participants "following directions" – in that they were instructed to create narratives about a target who was described as racist – the

ways that the targets were portrayed may have also represented a desire to "punish" the targets for his or her attitudes, by making him or her appear so distasteful that it would clearly show participants' own derision and lack of agreement. Also possible, and perhaps more important psychologically, is that even though they did not typically like the targets any less than anyone else (and in Study 3, they liked the target *more*), they may have felt affronted at having to get inside these targets' minds – something that other participants did not have to do – and used the opportunities to make the targets appear as vile as possible.

The fact that other participants did *not* portray the target as racist to the same extent as did perspective takers may simply reflect the fact that they, while disliking the targets, did not have any particular reasons to punish them for their attitudes; after all, in everyday life, people may encounter others whom they do not like. This is a fact of life, and when it happens, mature people simply acknowledge it and move on.

What is most interesting to consider in all this is that by vividly portraying the targets as racists, perspective takers – even though it was only done for the reasons outlined above – were *thinking* like racists. Because they were specifically asked to "try to imagine a day in the life of this person as if you actually <u>were</u> that person, looking at the world through his or her eyes and walking through the world in his or her shoes," they were probably considering thoughts such as "*If I were racist, what would I do in this situation?"* So while the racism expressed in participants essays did *not* have a direct effect on measures of participants' racial attitudes, and is *not* considered evidence of participants' own racism, it is also not inconceivable it could have effects on other

measures that were not collected in these studies. For example, even while it might not reflect any prejudice and might even encourage egalitarian goals, participants who have just taken the perspective of a racist might find an interaction with an African American to be uncomfortable, with the content they had just written in their essays still fresh in their minds – perhaps even evoking guilt at being able to think such thoughts. Of course, this is just speculation, but it is not entirely far-fetched, and might be fruitfully explored in the future.

Association, Dissociation, and Conflict

After considering the findings of all three studies, two additional facts stand out. In the two studies (1 and 2) where positive regard and self-target merging were equally low for perspective takers and matched control conditions, moderated effects were evident whereby higher motivation to respond without prejudice (internal motivation to respond without prejudice – Study 1; external motivation to respond without prejudice – Study 2) led to more negative attitudes (implicit attitudes – Study 1; explicit – Study 2) toward blacks. Conversely, in Study 3, where perspective taking led to higher merging, no moderated effects were found. Instead, perceiver-target merging was found to provide an indirect path through which perspective taking had an effect on attitudes.

Among other consequences of perspective taking, two different forces may come into play when taking the perspective of a racist target and the relative influence of each force may be impacted by who the target is and what is known about him or her. For example, perspective taking will typically lead to a "growing closer" to (associating with, valuing, and merging with) the target. In some cases, however, such as when the target is

disliked, there will also be a strong impetus to dissociate from the target as soon as the perspective taking attempt is over.

One task, then, is to reconcile findings from previous research with the research presented here, because research has found that perspective taking can still work to improve feelings toward the target and the target's group even when the target is quite dislikable. That is, even for quite distasteful targets, association can still take place. For example, Batson et al. (2002) found that perspective takers empathized with a drug dealer and his group, even asking that funds be allocated to a drug treatment program at the expense of other worthy (and funded) groups. And Batson, Polycarpou, et al. (1997) found that perspective takers empathized with a convicted murderer, which improved attitudes toward convicted murderers in general.

In Studies 1 and 2 of this dissertation, participants may have actually been actively trying to *dissociate* with the target following perspective taking, and this may have ended up driving the fact that the very participants who most wanted to avoid appearing prejudiced ended up producing more biased responses (i.e., perspective takers who were high in motivation to respond without prejudice may have been trying to suppress stereotypes about blacks or any type of thoughts which might be considered racist). In Study 3, however, even though no participants – perspective takers included – seemed to actually *like* the target or merge with him to any great extent, perspective taking instruction did influence participants' perceptions of the target so that they disliked him and dissociated with him *less*. What then may have driven the differences between these studies? Furthermore, why did Studies 1 and 2 not find, as in previous research, that

even when a target is dislikable, some connection can be established through perspective taking which can lead to greater valuing and merging with the target, while Study 3 found did find this effect?

One possibility is that when a racist is presented as extreme (e.g., wearing a "White Pride" t-shirt with Nazi overtones, as in Study 1), or when additional information is presented that is conflicting, such as painting the target as "nice" yet still racist (Study 2), perspective taking with a racist may lead to competing goals (of which participants may not even be consciously aware): to grow closer to the target but to also to create distance between the self and the target. In contrast, in Study 3, the target was not perhaps as heinously despicable as in Study 1 – even if he was still unambiguously racist – and thus participants did not as vigorously engage in distancing themselves. In effect, there may have been less urgency to "condemn" his unpopular attitudes than in Study 1, and yet no confusion that he might have some common values with participants, as in Study 2.

Generally, perspective taking instructions such as those used in this dissertation should tend to foster growing closer, because participants are asked to temporarily become the target, to "see the world through his (or her) eyes, and walk a day in his (or her) shoes." Participants complied with this request, as evidenced by the fact that perspective takers generally wrote their essays in the first person, as if they were the target. But considering the relatively low merging and positive regard (and lack of variability, particularly in Study 1) and the fact that in Studies 1 and 2 perspective takers did not differ from other perceivers on these measures, it is also clear that these

participants did not like the targets or want to merge with them any more than anyone else did (or did not want to admit to it). So, on the one hand, they were complying with instructions to get inside of the targets' heads and grow closer. On the other hand, they were making the targets appear as racist as possible, perhaps as one way to show their disapproval for his or her beliefs. In addition, they were also very likely trying to maintain egalitarian goals and distance themselves from the targets and their attitudes.

Another fact to consider is that across all studies, participants' scores on measures of racism were low, and their motivation to be non-prejudiced was high. This is consistent with the well-known finding that even if old-fashioned bigotry is still alive and well, not many people are willing to admit to it (e.g., Fiske, 1998). The fact that racism was low and motivation to be non-prejudiced was high, but participants still wrote essays that were, at times, filled with shockingly racist content (on behalf of the target) supports the suggestion made here: In Studies 1 and 2, perspective taking for a disliked target led to both association – indexed by greater convergence between the target's (inferred) attitudes and the attitudes expressed in the essays – and dissociation, which is implied by the uniformly low levels of positive regard and merging across in perspective taking and control conditions. In this case, it is easy to see how perspective taking could ironically make participants who most want to be egalitarian actually show more prejudiced responses – because these are the participants who are trying hardest to dissociate from the target and his or her racial attitudes by suppressing thoughts about racial bias.

It is possible that following the essay writing task – which was essentially an exercise in stereotype expression where association with the targets was encouraged, but

dissociation from the targets was probably desired – participants were free to pursue their egalitarian goals and to distance themselves further from the targets. They may have immediately done so, reporting the same low levels of liking and merging as participants in the control conditions. Still, negative stereotypes about and negative regard for blacks may have just been brought forth from participants' own minds, and these powerful associations and feelings were still in conscious awareness. So, in an attempt to banish these thoughts from awareness, they used a fairly common strategy that was in line with their motivated goals – they tried to suppress these stereotypes and negative feelings about blacks (e.g., Macrae et al., 1994), because they did not fit with their self-images as nonprejudiced (e.g., Devine, 1989). The greater was their motivation to be nonprejudiced, the greater were their attempts at suppression, and potentially, the greater was the rebound. This line of reasoning seems quite plausible in Study 1. In Study 2, where the target was much more subtly prejudiced, there was less stereotypical content to suppress – even though some perspective takers (i.e., those low in external motivation to respond without prejudice) responded in a way that was consistent with trying to show that they were not prejudiced. For others (i.e., those higher in external motivation), there was an increase that was not inconsistent with the idea of a rebound effect, although (as discussed in Study 2) typically rebound effects after suppression are found on measures more indirect than those used there.

In Study 3, the pattern of results was not consistent with rebound. Instead, there was a direct effect of perspective taking on merging, and of merging on attitudes. Because the information about the target's attitudes was presented in a relatively sterile

way, this might have allowed perspective takers to not be as immediately "turned off" by the target as was the case in Study 1, where he was not only clearly racist, but also quite vocal about it. In addition, in the no information condition, participants were told that the target "tried to avoid them [African Americans] whenever possible," and that only when he couldn't, did he act hostile and disrespectful. So the way that this "minimal" information was presented may not have completely discouraged merging, and probably discouraged it to a lesser extent than in Study 1.

In the information condition, participants were provided with details about a troubled history with blacks in the target's past. He grew up in a neighborhood where whites were the minority. He got in lots of fights with black kids and felt that this was because of his race. In addition, he usually lost these fights, even ending up in the hospital at one point. His sister was sexually assaulted, and she claimed that a black man was the perpetrator.

This information probably had more than one effect. First, on its own, it probably made participants more aware of common cultural stereotypes they were already aware of – that blacks are violent, mean, and sexually aggressive. Furthermore, this information made the target seem less responsible for his attitudes and made his feelings toward blacks seem more defensible and less controllable. Given these facts, even though participants still did not like the target very much and did not claim to merge with him to any large degree, they may have felt somewhat sorry for the target. For those who also took the target's perspective, they may have felt like it was OK to get inside of his head. Perhaps more importantly, after *being* inside of his head, there was somewhat less

urgency to get *out*. That is, it is unlikely that the key difference between Study 3 and Studies 1 and 2 lies in the "association" goal – which was pursued to the same extent as in earlier studies. Instead, it may have resided in the goal to dissociate, which while still strong, may have been less pressing in Study 3 than in earlier studies.

To the degree that participants felt more merged with the target, their attitudes also became more congruent with his and perhaps more acceptable to hold. And because they did not need to distance themselves further from the target, there was no particular need to suppress their attitudes. In view of the fact that thermometer ratings were generally still low (compare the average of 32.50 in the perspective taking/information cell in Study 3 to the overall mean of 31.32 in Study 1), participants probably did not feel as if they were responding in prejudiced ways – but compared to other conditions and indirectly driven by merging, they were.

Distinguishing Perspective Taking from Simply Thinking about a Racist Target

Across all studies, perspective takers – acting as the targets – accessed their knowledge of African American stereotypes and applied them, along with racist ideals, in the quite vividly portrayed days in the lives of racist targets. Participants in the perspective taking conditions, who displayed the highest levels of racist content in their essays, said things like "He knew there was no reason for this sambo to keep him waiting like this. 'Oh gee, maybe it's because I'm a tall white guy with a shaved head,' he thought to himself. Fucking niggers. He stormed out of the shop with fists clenched." (Study 1), "When we got to the coffee shop, there were two lines: one had a white man at the cash register and the other had an African American woman at the cash register. Although the

line for the register with the Black woman was shorter, I chose to go in the line with the white man up front because I knew he wouldn't try to overcharge me for my coffee." (Study 2), and "One thing I'm not proud of though is that I live in a society that accepts blacks. I hate them, and if I could never seen [*sic*] one again I would be a happy man," (Study 3).

A lot of rage was expressed in these essays on behalf of the target, and in order to generate this content (as discussed above), participants must have been fully aware of African American stereotypes and possessed information about what it should feel like to be a racist. According to Devine (1989), once stereotypical associations are activated, it takes a conscious effort to control them and respond in ways that are inconsistent with stereotyping. Extending her work, this set of studies shows that one does not need to directly consider a member of the stereotyped group or a symbolic equivalent to activate stereotypes about the group; it suffices to think about a member of a *different* group who has beliefs about the stereotyped group to activate these associations, presumably through shared neural representations such as "racist = (person who does not like) blacks," and "blacks = (automatically activated black stereotypes)." Thus, thinking about a racist target leads one to think about the reasons *why* this target might not like blacks, which leads one to think about why blacks, in general, might not be liked (i.e., negative stereotypes and affective responses).

This is not to say that participants who did not take the target's perspective did not activate stereotypes and negative associations about blacks. They did – and in Study 3, when provided with information about the target's reasons for being prejudiced, they

even took on some of the target's attitudes. But what is important to consider here is the relative complexity of taking the perspective of the target versus simply thinking about the target. While the latter task still brings up associations between blacks and negative content, there should not be the same type of inherent intrapsychic conflict. In effect, a participant might think "This target is racist. He thinks blacks are stupid and aggressive (stereotypes). He does not like being around them (affective content). But that is not me." Thus, there is a natural dissociation between the participant and the target.

For perspective takers, however, there is likely a quite profound conflict, as discussed above. For example, "The target is racist. He thinks blacks are stupid and aggressive and does not like being around them. What if *I* were this person? Then *I* would not like blacks. But I am *not* this person." When dissociation is lessened by the impact of other factors, there may be less conflict, or the conflict may be resolved in other ways. For example, "The target is a racist, and even though I am not this person, he has good reasons for being so." In each case, perspective taking leads to growing closer to the target in a way that is not reproduced when one simply thinks about a target. When the target is disliked, this will create a conflict that ultimately needs to be resolved, although the paths to resolution may differ based on several considerations, and only in some cases would this lead to a rebound effect.

Fundamental to the present research is the idea here is that even though perspective does not lead to psychological *indistinguishability* (e.g., Neuberg et al., 1997), it does lead to at least a temporary sense of merging. That is, the instructions encourage merging by suggesting that one should put oneself in another's shoes and

consider life from his or her perspective. This "growing closer" has been demonstrated in many ways, such as perspective taking leading to increased empathic concern and personal distress (e.g., Batson, Early, et al., 1997), greater similarity (Goldstein & Cialdini, 2007), greater valuing (e.g., Batson et al., 2007), and greater oneness or merging with the target (e.g., Cialdini et al., 1997; Galinsky et al., 2005).

Thus, across three studies with three different designs, a variety of instructions, and several types of dependent variables, this dissertation has shown that perspective taking provides a powerful method for understanding others, even when the outcome of this effort is not always desirable, or not always desirable for all perspective takers (i.e., those who would like to not respond in prejudiced ways).

Limitations

A few limitations of the studies should be briefly considered. First, participants in all three studies presented here were college students living in a quite liberal town, and many of these were probably from the West and Northwest (where this research was conducted), where the metropolitan areas from which many students come are also typically quite liberal. As discussed above, college students tend to have very egalitarian goals, and also tend to be very idealistic. Thus, the very idealism of these students, their typically low racism and prejudice, and their goals of equality and a rejection of bias (which is also encouraged and fostered in the college environment) may have worked *against* finding the effects of perspective taking on racial attitudes. Conducting the same research on other samples, such as older individuals (who might tend to be less "idealistic" and more "realistic" about the way things are in the world) might actually

make it easier to find the effects that were under investigation. Furthermore, conducting the same research in other places where the community is less intolerant of prejudice of any kind might also serve to make effects more likely to find.

Another consideration is that the targets used in all three studies were imaginary people. In Study 1, only a photograph was provided (similar to much other perspective taking research). In Studies 2 and 3, although more information was presented about the targets, making them somewhat less one-dimensional than in Study 1, the targets were still not at all "real" (even if some participants viewed them as such). Of course, real people are very complex, as is trying to understand them.

Imagine for a moment knowing someone, and thinking that you know this person well, and then hearing them make a derogatory comment about a social group. In an attempt to understand this person, you might try to "get inside their head," wanting to understand why this person would say such a thing. You might think, "This is someone I like!" You might search for any kind of reason that seemed, well…reasonable. Certainly, this type of perspective taking – which is much more like "real world" perspective taking – is not as involved as spending ten minutes imagining yourself as another person and writing an essay about a day in their life. Still, our repeated efforts to try and understand others are a staple of social interactions, and perspective taking might be one way that attitudes are transmitted and from person to person, parents to children, friend to friend. Thus, it is important to consider not only how perspective taking can help us to become better people, but also to investigate when it might be a less than optimal solution to understanding other minds.

Conclusions

The goal of this dissertation was to show that perspective taking, typically thought of as a beneficial activity, can sometimes be a harmful one. While other perspective taking researchers have also outlined sometimes surprising and negative effects of taking someone's perspective, such as favoring the target's group at the expense of other deserving groups (Batson et al., 2002; Batson, Klein, et al., 1995), claiming more than a fair share of resources for the self (Epley et al., 2006), or acting out the negative stereotypes of a target's group (Galinsky et al., 2008), this study suggests some possible drawbacks of perspective taking.

Starting from a large base of research that has shown how perspective taking increases helping (e.g., Batson, 2010) and facilitates social bonding between perceivers, targets, and their groups (e.g. Galinsky et al., 2005), it was hypothesized that the very social bonds that increase helping and social coordination might also decrease social coordination between perceivers and groups about which a target has negative feelings (e.g., Galinsky et al., 2008). In effect, it set out to show that not every opportunity to take someone's perspective should be taken, and that it is not advisable to get inside of just anyone's head.

Racism in this country is still a large problem, and although the debate over whether it has moved underground still rages on, very few researches would claim that it has disappeared. For example, a search on PsycNET for the terms "prejudice" or "racism," even when limited to the last 5 years, brings up 5,196 entries. This indicates that it is still a topic of great interest. As it should, much of this research focuses on when

prejudice or racism occurs, the processes by which it occurs, and most importantly, what can we do to stop it from occurring. The present research thus contributes to a very important topic by delineating one condition under which prejudice and racism might become stronger. Sometimes, perspective taking may be something to avoid.

APPENDIX A

ESSAY WRITING INSTRUCTIONS – STUDIES 1-3

Perspective Taking Essay Writing Instructions – Study 1 [Information Contained within Brackets was Included in Perspective Taking and 3rd Person Control Conditions]

[In this task, we are interested in people's ability to construct life-event details from visual information alone. Please take a moment to look closely at the photograph of the person you selected.]

[When you are ready, please spend about 5-10 minutes writing a short essay about a typical day in the life of the person in the photograph.] While writing this essay, try to imagine a day in the life of this person as if you were that person, looking at the world through his or her eyes and walking through the world in his or her shoes. Write the essay in the 1st person (e.g., After I woke up, a friend called me).

[On the next page, you will be given details about several events. Try to imagine that these events actually took place during the course of the day. While writing your essay, discuss this person's thoughts and feelings in the moments leading up to the events, during the events, and after the events. Also include things that this person might have done or said before, during, or after these events.] Remember to write about this in the 1st person, however. In other words, write about what this person thought, felt, did, or said *as if you were that person, living those experiences.*

Third Person Essay Writing Instructions – Study 1

While writing this essay, try to imagine a day in the life of this person. Write the essay in the 3rd person (e.g., After he woke up, a friend called him).

Remember to write about this in the 3rd person, however. In other words, write about what this person thought, felt, did, or said *as if you were able to observe this person.*

"Self" Essay Writing Instructions – Study 1

In this study, we are interested in how people construct life-event details. When you are ready, we would like you to spend about 5-10 minutes writing a short essay about a typical day in your own life.

On the next page, you will be given details about several events. Try to imagine that these events actually took place during the course of your day. While writing your essay, discuss your thoughts and feelings in the moments leading up to the events, during the events, and after the events. Also include things you might have done or said before, during, or after these events.

Behavioral Scenarios for all Conditions – Study 1 [Information within Brackets was Used in the "Self" Control Condition Only]

During the course of his or her day, the events below involving the person in the photograph (the *target)* took place [As part of your essay, *try to imagine that the events described below took place*]. Please incorporate these events into your essay.

That morning, the *target* was [you were] trying to ask a black (African-American) clerk in a hardware store where the *target* [you] might find the gardening section. The clerk, who was busy explaining a piece of hardware to a black customer, told the *target* [you] that he would be with him [you] in just a moment. After waiting for a short while, the *target* [you] turned around and walked away.

Later that day, the *target* [you] went into a chain restaurant to buy lunch, and while waiting at a counter, overheard the employees talking about a new general manager. There were no other customers around. The employees mentioned that the manager they

were discussing was black, and one of the employees said something about affirmative action and the employees laughed. The *target* [You] joined in the conversation for a while, and after getting the [your] food, left a very large tip and walked out.

That evening after dinner, the *target* [you] grabbed a book off of his or her [your] table and went out for a walk on a path near a river. During the walk, the *target* [you] sat down on a bench under a lamppost for a while. Later, the *target* [you] walked home.

Perspective Taking Essay Writing Instructions – Study 2

The researcher for this study is interested in the way people construct life-event details about a person they have never met. On the next page, you will be given additional details about the person in the photograph.

After studying the photograph the experimenter gave you, and reading the information about the person in the photograph, you should spend about 5-10 minutes writing a short essay about a typical day in the life of that person. While writing this essay, take the perspective of the person in the photograph. Try to imagine what this person has been through and how the person feels. In other words, try to imagine a day in the life of this person as if you actually were that person, looking at the world through his or her eyes and walking through the world in his or her shoes.

Essay Writing Instructions – Study 3 [Information within Brackets was used in the Perspective Taking Conditions Only]

In this task, we are interested in people's ability to construct life-event details of a person they have never met. When you are ready, please spend about 5-10 minutes writing a short essay about a typical day in the life of the person in the photograph. While writing this essay, try to imagine what a day in the life of this person might be like [a day in the life of this person as if you actually were that person, looking at the world through his or her eyes and walking through the world in his or her shoes].

While writing your essay, discuss this person's thoughts and feelings as they might have occurred during the course of his or her day. You can also include things this person might have said or done during the day. In other words, write about what this person thought, felt, did, or said during the course of his or her day [Remember to write about what this person thought, felt, did, or said during the course of his or her day *as if you were that person, living those experiences*].

On the next page, you will be given additional details about this person. You can feel free to use (or not use) the information you are given to try and write about a day in their life.

APPENDIX B
VIGNETTE ABOUT "AMANDA" – A SUBTLY RACIST TARGET

Information within [Brackets] was not Included in the Non-Racist Target Condition and Information within {Curly Brackets} was not Included in the Racist Target Conditions

The person in the photograph is named Amanda. Amanda is a 24-year-old woman. She was raised in a small, conservative town in a southern part of the Midwest, but moved to the west coast a few years ago. Amanda has a junior college degree, and currently works as an administrative assistant in a medium sized manufacturing company. Her job is the reason she moved, after she was recruited at a job fair held at her college.

Amanda still feels deep roots to her home community, and misses her family and old friends. She would have stayed if there were more better paying jobs. It might have been a small town, but she felt that most of the people there looked at the world the same way she does. She also felt safe there because there was very little crime – a person could go to sleep at night with the windows open and not worry.

Amanda considers herself a "regular" sort of person. She doesn't really consider herself a big reader, but she likes popular magazines. She doesn't like to talk politics [, although she was rather vocal in her opposition to Barack Obama during the 2008 presidential election campaign].

For fun, she likes to go out for a drink or two after work with her friends, and also likes to go dancing. She likes to laugh a lot, and her friends think she is funny. Her friends also think of her as someone they can talk to when they have a problem, because she is a good listener and gives good advice. Although she goes out on dates sometimes, she doesn't have a boyfriend. She loves children, though, and dreams of meeting the right man, someone who shares some of her interests and has similar family values, and settling down with him.

Recently, an acquaintance from work called Amanda to try and set her up on a blind date. The acquaintance described the man as her husband's friend from work, and told Amanda that he was smart, interesting, and really fun to hang out with. {She also described him as very attractive, tall and handsome. Amanda told the acquaintance that although the guy sounded nice, she wasn't really very into going on blind dates, because in the past they hadn't worked out.} [Amanda was pretty excited, and asked what he looked like. The acquaintance described him as very attractive, tall and handsome. Amanda was about to agree to being set up when the acquaintance mentioned that the man was African American. Amanda told the acquaintance that although the guy sounded nice, she wasn't really very into going on blind dates, because in the past they hadn't worked out. In truth, Amanda had never felt very comfortable around African Americans, and felt uneasy about dating a black man.]

APPENDIX C

VIGNETTES ABOUT "JON L." – MINIMAL AND BACKGROUND INFORMATION

Minimal Information

The person in the photo, Jon L., is 23 years old, unmarried, and works in a warehouse. Although he now lives in the Pacific Northwest, Jon grew up in Los Angeles, CA.

Jon has a strong dislike for African Americans and believes that they are an inferior race. Jon avoids African Americans whenever he can, and when he can't, his attitude and behavior is typically aggressive and hostile.

Background Information

The person in the photo, Jon L., is 23 years old, unmarried, and works in a warehouse. Although he now lives in the Pacific Northwest, Jon grew up in Los Angeles, CA.

When Jon was growing up, he lived in a neighborhood where many African Americans also lived, and where whites were the minority group. Throughout his school years, Jon got into a lot of fights, and was badly beaten (ending up in the hospital) on several occasions by a group of other kids who didn't like him. The kids who beat him up were African Americans, and Jon always thought that they picked on him because he was white. Right before Jon graduated from high school, his sister was the victim of a sexual attack, and the perpetrator (whom she described as an African American) was never caught. Soon after, Jon's dad got a job in Oregon, and the family moved up there.

Jon has a strong dislike for African Americans and believes that they are an inferior race. Jon avoids African Americans whenever he can, and when he can't, his attitude and behavior is typically aggressive and hostile.

REFERENCES

Aiken, L. S., & West, S. G. (1991). Multiple regression: testing and interpreting interactions. Newbury Park, CA: Sage.

Ames, D. L., Jenkins, A. C., Banaji, M. R., & Mitchell, J. P. (2008). Taking another person's perspective increases self-referential neural processing. *Psychological Science, 19*, 642-644.

Ames, D. R. (2005). Everyday solutions to the problem of other minds: Which tools are used when? In B. F. Malle & S. D. Hodges (Eds.), *Other minds: How humans bridge the divide between self and others* (pp. 158-173). New York: Guilford Press.

Arbuckle, J. L. (2006). Amos (Version 6.0) [Computer Program]. Chicago: SPSS.

Aron, A., Aron, E. N., Tudor, M., & Nelson, G. (1991). Close relationships as including the other in the self. *Journal of Personality and Social Psychology, 60,* 241-253.

Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of Other in the Self scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology, 63*, 596-612.

Baldwin, J. R., Day, L. E., & Hecht, M. L. (2000). The structure(s) of racial attitudes among White college students. *International Journal of Intercultural Relations, 24,* 553-577.

Bargh, J. A. (1994). The Four Horsemen of automaticity. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (pp. 1-40). Hillsdale, NJ: Erlbaum.

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Directs effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology, 71*, 230-244.

Batson, C. D. (1987). Prosocial motivation: Is it ever truly altruistic? In L. Berkowitz (Ed.), *Advances in experimental social psychology: Vol. 20* (pp. 65-122). San Diego, CA: Academic Press.

Batson, C. D. (1991). *The altruism question: Toward a social-psychological answer.* Hillsdale, NJ: Erlbaum.

Batson, C. D. (1997). Self-other merging and the empathy-altruism hypothesis: Reply to Neuberg et al. (1997). *Journal of Personality and Social Psychology, 73*, 517-522.

Batson, C. D. (1998). Altruism and prosocial behavior. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology: Vols. 1 and 2* (4th Ed., pp. 517-522).

Batson, C. D. (2002). Addressing the altruism question experimentally. In S. G. Post, L. G., Underwood, P. Jeffrey, & W. B. Hurlbut (Eds.), *Altruism & altruistic love: Science, philosophy, & religion in dialogue* (pp. 89-105). New York: Oxford University Press.

Batson, C. D. (2010). Empathy-induced altruistic motivation. In M. Mikulincer & P. R. Shaver Batson (Eds.), *Prosocial motives, emotions, and behavior: The better angels of our nature* (pp. 15-34). Washington, DC: American Psychological Association.

Batson, C. D., & Ahmad, N. (2001). Empathy-induced altruism in a prisoner's dilemma II: What if the target of empathy has defected? *European Journal of Social Psychology, 31,* 25-36.

Batson, C. D., Batson, J. G., Griffitt, C. A., Barrientos, S., Brandt, J. R., Sprengelmeyer, P., & Bayly, M. J. (1989). Negative state relief and the empathy-altruism hypothesis. *Journal of Personality and Social Psychology,* 56, 922-933.

Batson, C. D., Batson, J. G., Slingsby, J. K., Harrell, K. L., Peekna, H. M., & Todd, R. M. (1991). Empathic joy and the empathy-altruism hypothesis. *Journal of Personality and Social Psychology, 61*, 413-426.

Batson, C. D., Batson, J. G., Todd, M., Brummett, B. H., Shaw, L. L., Aldeguer, C. M. R. (1995). Empathy and the collective good: Caring for one of the others in a social dilemma. *Journal of Personality and Social Psychology, 68,* 619-631.

Batson, C. D., Chang, J., Orr, R., & Rowland, J. (2002). Empathy, attitudes, and Action: Can feeling for a member of a stigmatized group motivate one to help the group? *Personality and Social Psychology Bulletin, 28,* 1656-1666.

Batson, C. D., Dyck, J. L., Brandt, J. R., Batson, J. G., Powell, A. L., McMaster, M. R., Griffitt, C. (1988). Five studies testing two new egoistic alternatives to the empathy-altruism hypothesis. *Journal of Personality and Social Psychology, 55,* 52-77.

Batson, C. D., Early, S., & Salvarani, G. (1997). Perspective taking: Imagining how another feels versus imagining how you would feel. *Personality and Social Psychology Bulletin, 23*, 751-758.

Batson, C. D., Eklund, J. H., Chermok, V. L., Hoyt, J. L., & Ortiz, B. G. (2007). An additional antecedent of empathic concern: Valuing the welfare of the person in need. *Journal of Personality and Social Psychology, 93,* 65-74.

Batson, C. D., Fultz, J. & Schoenrade, P. A. (1987). Distress and Empathy: Two qualitatively distinct vicarious emotions with different motivational consequences. *Journal of Personality, 55,* 19-39.

Batson, C. D., Klein, T. R., Highberger, L., & Shaw, L. L. (1995). Immorality from empathy-induced altruism: When compassion and justice conflict. *Journal of Personality and Social Psychology, 68,* 1042-1054.

Batson, C. D., Lishner, D. A., Carpenter, A., Dulin, L., Hajusola-Webb, S., Stocks, E. L., et al. (2003). "...As you would have them do unto you": Does imagining yourself in the other's place stimulate moral action? *Personality and Social Psychology Bulletin, 29,* 1190-1201.

Batson, C. D., Lishner, D. A., Cook, J., & Sawyer, S. (2005). Similarity and nurturance: Two possible sources of empathy for strangers. *Basic and Applied Social Psychology, 27*, 15-25.

Batson, C. D., & Moran, T. (1999). Empathy-induced altruism in a prisoner's dilemma. *European Journal of Social Psychology, 29*, 909-924.

Batson, C. D., O'Quin, K., Fultz, J., Vanderplas, M., & Isen, A. M. (1983). Influence of self-reported distress and empathy on egoistic versus altruistic motivation to help. *Journal of Personality and Social Psychology, 46,* 706-718.

Batson, C. D., Polycarpou, M. P., Harmon-Jones, E., Imhoff, H. J., MItchener, E. C., Bednar, L. L., et al. (1997). Empathy and attitudes: Can feeling for a member of a stigmatized group improve feelings toward the group? *Journal of Personality and Social Psychology, 72,* 105-118.

Batson, C. D., Sager, K., Garst, E., Kang, M., Rubchinsky, K., & Dawson, K. (1997). Is empathy-induced helping due to self-other merging? *Journal of Personality and Social Psychology, 73,* 495-509.

Batson, C. D., Turk, C. L., Shaw, L. L., & Klein, T. R. (1995). Information function of empathic emotion: Learning that we value the other's welfare. *Journal of Personality and Social Psychology, 68,* 300-313.

Batson, C. D., & Weeks, J. L. (1996). Mood effects of unsuccessful helping: Another test of the empathy-altruism hypothesis. *Personality and Social Psychology Bulletin, 22,* 148-157.

Berenguer, J. (2007). The effect of empathy in proenvironmental attitudes and behaviors. *Environment and Behavior, 39,* 269-283.

Berenguer, J. (2010). The effect of empathy in environmental moral reasoning. *Environment and Behavior, 42,* 110-134.

Bernstein, W. M., & Davis, M. H. (1982). Perspective-taking, self-consciousness, and accuracy in person perception. *Basic and Applied Social Psychology, 3,* 1-19.

Brehm, S. S., & Aderman, D. (1977). On the relationship between empathy and the actor versus observer hypothesis. *Journal of Research in Personality, 11,* 340-346.

Brigham, J. C. (1993). College students' racial attitudes. *Journal of Applied Social Psychology, 23,* 1933-1967.

Caruso, E. M., Epley, N., & Bazerman, M. H. (2006). The costs and benefits of undoing egocentric responsibility assessments in groups. *Journal of Personality and Social Psychology, 91,* 857-871.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology, 76,* 893-910.

Cialdini, R. B., Brown, S. L., Lewis, B. P., Luce, C., & Neuberg, S. L. (1997). Reinterpreting the empathy-altruism relationship: When one into one equals oneness. *Journal of Personality and Social Psychology, 73,* 481-494.

Clore, G. L., & Jeffery, K. M. (1972). Emotional role playing, attitude change, and attraction toward a disabled person. *Journal of Personality and Social Psychology, 23,* 105-111.

Coke, J. S., Batson, C. D., & McDavis, K. (1978). Empathic mediation of helping: A two-stage model. *Journal of Personality and Social Psychology,* 36, 752-766.

Correll, J., Park, B., Judd, C. M., & Wittenbrink, B. (2002). The police officer's dilemma: Using ethnicity to disambiguate potentially threatening individuals. *Journal of Personality and Social Psychology, 83,* 1314-1329.

Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science, 2,* 163-170.

Dasgupta, N., McGhee, D. E., Greenwald, A. G., & Banaji, M. R. (2000). Automatic preference for white Americans: Eliminating the familiarity explanation. *Journal of Experimental Social Psychology, 36,* 316-328.

Davis, M. H. (1983a). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44,* 113-126.

Davis, M. H. (1983b). The effects of dispositional empathy on emotional reactions and helping: A multidimensional approach. *Journal of Personality, 51,* 167-184.

Davis, M. H. (2005). A "constituent" approach to the study of perspective taking. In B. F. Malle & S. D. Hodges (Eds.), *Other minds: How humans bridge the divide between self and others* (pp. 44-55). New York: Guilford Press.

Davis, M. H., Conklin, L., Smith, A., & Luce, C. (1996). The effect of perspective taking on the cognitive representation of persons: A merging of self and other. *Journal of Personality and Social Psychology, 70,* 713-726.

Davis, M. H., Mitchell, K. V., Hall, J. A., Lothert, J., Snapp, T. & Meyer, M. (1999). Empathy, expectations, and situational preferences: Personality influences on the decision to participate in volunteer helping behaviors. *Journal of Personality, 67,* 469-503.

Decety, J., & Chaminade, T. (2003). When the self represents the other: A new cognitive view on psychological identification. *Consciousness and Cognition, 12,* 577-596.

Decety, J., & Grèzes, J. (2006). The power of simulation: Imagining one's own and other's behavior. *Brain Research, 1079,* 4-14.

Decety, J., & Hodges, S.D. (2006). The social neuroscience of empathy. In: P. A. M. Van Lange (Ed.), Bridging Social Psychology: Benefits of Transdisciplinary Approaches (pp. 103-109). Mahwah: Erlbaum.

Decety, J., & Jackson, P. L. (2006). A social-neuroscience perspective on empathy. *Current Directions in Psychological Science, 15,* 54-58.

Decety, J., & Meyer, M. (2008). From emotion resonance to empathic understanding: A social developmental neuroscience account. *Development and Psychopathology, 20,* 1053-1080.

Decety, J., & Sommerville, J. A. (2003). Shared representations between self and other: A social cognitive neuroscience view. *TRENDS in Cognitive Sciences, 7*, 527-533.

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: A normative analysis and review. *Psychological Bulletin, 135,* 347-368.

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

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,* 835-848.

Dovidio, J. F., ten Verget, M., Stewart, T. L., Gaertner, S. L., Johnson, J. D., Esses, V. M., et al. (2004). Perspective and prejudice: Antecedents and mediating mechanisms. *Personality and Social Psychology Bulletin, 30,* 1537-1549.

Eisenberg, N., & Fabes, R. A. (1990). Empathy: Conceptualization, measurement, and relation to prosocial behavior. *Motivation and Emotion, 14,* 131-149.

Epley, N., Caruso, E. M., & Bazerman, M. H. (2006). When perspective taking increases taking: Reactive egoism in social interaction. *Journal of Personality and Social Psychology*, *91*, 872-889.

Epley, N., Keysar, B., Van Boven, L., & Gilovich, T. (2004). Perspective taking as egocentric anchoring and adjustment. *Journal of Personality and Social Psychology*, 3, *327-339.*

Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology, 54,* 297-327.

Finlay, K. A., & Stephan, W. G. (2000). Improving intergroup relations: The effects of empathy on racial attitudes. *Journal of Applied Social Psychology, 30,* 1720-1737.

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

Frantz, C. M., Cuddy, A. J., Burnett, M., Ray, H., & Hart, A. (2004). A threat in the computer: The race Implicit Association Test as a stereotype threat experience. *Personality and Social Psychology Bulletin, 30,* 16111-1624.

Franzoi, S. L., Davis, M. H., & Young, R. D. (1985). The effects of private self-consciousness and perspective taking on satisfaction in close relationships. *Journal of Personality and Social Psychology, 48,* 1584-1594.

Fultz, J., Batson, C. D., Fortenbach, V. A., McCarthy, P. M., & Varney, L. L. (1986). *Journal of Personality and Social Psychology, 50,* 741-769.

Gaertner, S. L., & Dovidio, J. F. (1986). The aversive form of racism. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 61-89). Orlando, FL: Academic Press.

Galinsky, A. D., & Ku, G. (2004). The effects of perspective-taking on prejudice: The moderating role of self-evaluation. *Personality and Social Psychology Bulletin, 30,* 594-604.

Galinsky, A. D., Ku, G., & Wang, C. S. (2005). Perspective-taking and self-other overlap: Fostering social bonds and facilitating social coordination. *Group Processes & Intergroup Relations, 8,* 109-124.

Galinsky, A. D., Maddux, W. W., Gilin, D., & White, J. B. (2008). Why it pays to get inside the head of your opponent: The differential effects of perspective taking and empathy in negotiations. *Psychological Science, 19,* 378-384.

Galinsky, A. D., & Moskowitz, G. B. (2000). Perspective-taking: Decreasing stereotype expression, stereotype accessibility, and in-group favoritism. *Journal of Personality and Social Psychology, 78,* 708-724.

Galinsky, A. D., & Moskowitz, G. B. (2007). Further ironies of suppression: Stereotype *and* counterstereotype accessibility. *Journal of Experimental Social Psychology, 43,* 833-841.

Galinsky, A. D., & Mussweiler, T. (2001). First offers as anchors: The role of perspective-taking and negotiator focus. *Journal of Personality and Social Psychology, 4,* 657-669.

Galinsky, A. D., Wang, C. S., & Ku, G. (2008). Perspective-takers behave more stereotypically. *Journal of Personality and Social Psychology, 95,* 404-419.

Galper, R. E. (1976). Turning observers into actors: Differential causal attributions as a function of empathy. *Journal of Research in Personality, 10,* 328-335.

Gawronski, B., Peters, K. R., Brochu, P. M., & Strack, F. (2008). Understanding the relations between different forms of racial prejudice: A cognitive consistency perspective. *Personality and Social Psychology Bulletin, 34,* 648-665.

Goldstein, N. J., & Cialdini, R. B. (2007). The spyglass self: A model of vicarious self-perception. *Journal of Personality and Social Psychology, 3*, 402-417.

Gould, R., & Sigall, H. (1977). The effects of empathy and outcome on attribution: An examination of the divergent-perspectives hypothesis. *Journal of Experimental Social Psychology, 13,* 480-491.

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

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,* 1464-1480.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology, 85,* 197-216.

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,* 17-41.

Heider, F. (1958). *The psychology of interpersonal relations.* New York: Wiley.

Hilmert, C. J., Kulik, J. A., & Christenfeld, N. J. S. (2006). Positive and negative opinion modeling: The influence of another's similarity and dissimilarity. *Journal of Personality and Social Psychology, 90*, 440-452.

Hodges, S. D. (2005). Is how much you understand me in your head or mine? In B. F. Malle & S. D. Hodges (Eds.), *Other minds: How humans bridge the divide between self and others* (pp. 298-309). New York: Guilford Press.

Hodges, S. D., & Biswas-Diener, R. (2007). Balancing the empathy expense account: Strategies for regulating empathic response. In T. F. D. Farrow & P. W. R. Woodruff (Eds.), *Empathy in mental illness* (pp. 389-407). Cambridge, UK: University Press.

Hodges, S. D., & Klein, K. J. K. (2001). Regulating the costs of empathy: The price of being human. *The Journal of Socio-Economics, 30,* 437-452.

Hodges, S. D., & Wegner, D. M. (1997). Automatic and controlled empathy. In W. J. Ickes (Ed.), *Empathic accuracy* (pp. 311-339). New York: Guildford Press.

Ickes, W. (1997). *Empathic accuracy.* New York: Guilford Press.

Jackson, P. L., Brunet, E., Meltzoff, A. N., & Decety, J. (2006). Empathy examined through the neural mechanisms involved in imagining how I feel versus how you feel pain. *Neuropsychologia, 44*, 752-761.

Johnson, S. C. (2000) The recognition of mentalistic agents in infancy. *Trends in Cognitive Science, 4*, 22-28.

Jones, E. E., & Nisbett, R. E. (1972). The actor and the observer: Divergent perceptions of the causes of behavior. In E. E. Jones, D. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 79-94). Morristown, NJ: General Learning Press.

Judd, C. M., Park, B., Ryan, C. S., Brauer, M., & Kraus, S. (1995). Stereotyps and ethnocentrism: Diverging interethnic perceptions of African American and White American youth. *Journal of Personality and Social Psychology, 69*, 460-481.

Lamm, C., Batson, C. D., & Decety, J. (2007). The neural substrate of human empathy: Effects of perspective-taking and cognitive appraisal. *Journal of Cognitive Neuroscience, 19*, 42-58.

Landau, J. (Director), & Cameron, J. (Writer/Director/Producer). (2009). *Avatar* [Motion picture]. United States: Twentieth Century-Fox Film Corporation.

Macrae, C. N., Bodenhausen, G. V., Milne, A. B., & Ford., R. L. (1997). On the regulation of recollection: The intentional forgetting of stereotypical memories. *Journal of Personality and Social Psychology, 72,* 709-719.

Macrae, C. N., Bodenhausen, G. V., Milne, A. B., & Jetten, J. (1994). Out of mind but back in sight: Stereotypes on the rebound. *Journal of Personality and Social Psychology, 67*, 808-817.

Malle, B. F. (1998/2009). *F.Ex: Coding scheme for people's folk explanations of behavior.* University of Oregon. Retrieved April 15, 2009 from http://darkwing.uoregon.edu/~interact/fex.html. (Original work published 1998).

Malle, B. F. (1999). How people explain behavior: A new theoretical approach. *Personality and Social Psychology Review, 3,* 23-48.

Malle, B. F. (2004). History: Past research on attribution and behavior explanation. In B. F. Malle (Ed.), *How the mind explains behavior: Folk explanations, meaning, and social interaction* (pp. 5-27). Cambridge, MA: MIT Press.

Malle, B. F. (2005). Three puzzles of mindreading. In B. F. Malle & S. D. Hodges (Eds.), *Other minds: How humans bridge the divide between self and others* (pp. 26-43). New York: Guilford Press.

Malle, B. F. (2006). The actor-observer asymmetry in attribution: A (surprising) meta-analysis. *Psychological Bulletin, 132,* 895-919.

Malle, B. F. (2010). *Storms revisited: Visual perspective and the actor-observer asymmetry in attribution.* Unpublished manuscript.

Malle, B. F., & Hodges, S. D. (2005). *Other minds: How humans bridge the divide between self and others.* New York: Guilford Press.

Malle, B. F., & Knobe, J. (1997). Which behaviors do people explain? A basic actor-observer asymmetry. *Journal of Personality and Social Psychology, 72*, 288-304.

Malle, B. F., Knobe, J., & Nelson, S. E. (2007). Actor-observer asymmetries in explanations of behavior: New answers to an old question. *Journal of Personality and Social Psychology, 93,* 491-514.

Malle, B. F., Knobe, J., O'Laughlin, M. J., Pearce, G. E., Nelson, S. E. (2000). Conceptual structure and social functions of behavior explanations: Beyond person-situation attributions. *Journal of Personality and Social Psychology, 79,* 309-326.

Malle, B. F., & Pearce, G. E. (2001). Attention to behavioral events during interaction: Two actor-observer gaps and three attempts to close them. *Journal of Personality and Social Psychology, 81,* 278-294.

Maner, J. K., Luce, C. L., Neuberg, S. L., Cialdini, R. B., Brown, S., & Sagarin, B. J. (2002). The effects of perspective taking on motivations for helping: Still no evidence for altruism. *Personality and Social Psychology Bulletin, 28*, 1601-1610.

Marx, D. M., & Stapel, D. A. (2006). It depends on your perspective: The role of self-relevance in stereotype-based underperformance. *Journal of Experimental Social Psychology, 42,* 768-775.

Mason, M. F., & Macrae, C. N. (2008). Perspective-taking from a social neuroscience standpoint. *Group Processes & Intergroup Relations, 11*, 215-232.

McConahay, J. B. (1986). Modern racism, ambivalence, and the Modern Racism Scale. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 91-125). San Diego: Academic Press.

McConnell, A. R., & Leibold, J. M. (2001). Relations among the Implicit Association Test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology, 37,* 435-442.

McCullough, M. E., Worthington, E. L., Jr., Rachal, K. C. (1997). Interpersonal forgiving in close relationships, *Journal of Personality and Social Psychology, 73,* 321-336.

Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology, 31,* 838-850.

Michaelson, J. (2009). The meaning of Avatar: Everything is God (A response to Ross Douthat and other naysayers of 'pantheism'). *Huffington Post.* Retrieved February 20, 2010, from http://www.huffingtonpost.com/jay-michaelson/the-meaning-of-avatar-eve_b_400912.html.

Miller, R. S. (1987). Empathic embarrassment: Situation and personal determinants of reactions to the embarrassment of another. *Journal of Personality and Social Psychology, 53,* 1061-1069.

Monteith, M. J. (1993). Self-regulation of prejudiced responses: Implications for progress in prejudice-reduction efforts. *Journal of Personality and Social Psychology, 65,* 469-485.

Monteith, J. M., Spicer, C. V., & Tooman, G. D. (1998). Consequences of stereotype suppression: Stereotypes on AND not on the rebound. *Journal of Experimental Social Psychology, 34,* 355-377.

Moses, L. J. (2005). Executive functioning and children's theories of mind. In B. F. Malle & S. D. Hodges (Eds.), *Other minds: How humans bridge the divide between self and others* (pp. 11-25). New York: Guilford Press.

Neuberg, S. L., Cialdini, R. B., Brown, S. L., Luce, C., Sagarin, B. J., & Lewis, B. P. (1997). Does empathy lead to anything more than superficial helping? Comment on Batson et al. (1997). *Journal of Personality and Social Psychology, 73*, 510-516.

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2006). Website: http://implicit.harvard.edu/.

Parker, S. K., & Axtell, C. M. (2001). Seeing another viewpoint: Antecedents and outcomes of employee perspective taking. *Academy of Management Journal, 44,* 1085-1100.

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,* 16-31.

Pettigrew, T. F., & Meertens, R. W. (1995). Subtle and blatant prejudice in western Europe. *European Journal of Social Psychology, 25*, 57-75.

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology, 75*, 811-832.

Prentice, D. A., & Carranza, E. (2002). What women should be, shouldn't be, are allowed to be, and don't have to be: The contents of prescriptive gender stereotypes. *Psychology of Women Quarterly, 26,* 269-281.

Regan, D. T., & Totten, J. (1975). Empathy and attribution: Turning observers into actors. *Journal of Personality and Social Psychology, 32,* 850-856.

Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: Evidence from 14- to 18-month-olds. *Developmental Psychology, 33*, 12-21.

Richardson, D. R., Hammock, G. S., Smith, S. M., Gardner, W., & Signo, M. (1994). Empathy as a cognitive inhibitor of interpersonal aggression. *Aggressive Behavior, 20,* 275-289.

Ruby, P, & Decety, J. (2004). How would *you* feel versus how do you think *she* would feel? A neuroimaging study of perspective-taking with social emotions. *Journal of Cognitive Neuroscience, 16*, 988-999.

Schultz, P. W. (2000). Empathizing with nature: The effects of perspective taking on concern for environmental issues. *Journal of Social Issues, 56,* 391-406.

Sears, D. O. (1988). Symbolic racism. In P. A. Katz & D. A. Taylor (Eds.), *Eliminating racism: Profiles in controversy* (pp. 53-84). New York: Plenum Press.

Sedikides, C. (1992). Attentional effects on mood are moderated by chronic self-conception valence. *Personality and Social Psychology Bulletin, 18,* 580-584.

Sevillano, V., Aragonès, J. I., & Schultz, P. W. (2007). Perspective taking, Environmental concern, and the moderating role of dispositional empathy. *Environment and Behavior, 39,* 685-705.

Shaw, L. L., Batson, C. D., & Todd, R. M. (1994). Empathy avoidance: Forestalling feeling for another in order to escape the motivational consequences. *Journal of Personality and Social Psychology, 67,* 879-887.

Sidanius, J., Pratto, F., & Bobo, L. (1994). Social dominance orientation and the political psychology of gender: A case of invariance? *Journal of Personality and Social Psychology, 67,* 998-1011.

Smeesters, D., Wheeler, S. C., & Kay, A. C. (2009). The role of interpersonal perceptions in the prime-to-behavior pathway. *Journal of Personality and Social Psychology, 96,* 395-414.

Smith, E. R., Coats, S., & Walling, D. (1999). Overlapping mental representations of self, in-group, and partner: Further response time evidence and a connectionist model. *Personality and Social Psychology Bulletin, 25,* 873-882.

Smith, E. R., & Henry, S. (1996). An in-group becomes part of the self: Response time evidence. *Personality and Social Psychology Bulletin, 22,* 635-642.

Smith, K. D., Keating, J. P., & Stotland, E. (1989). Altruism revisited: The effect of denying feedback on a victim's status to empathic witnesses. *Journal of Personality and Social Psychology, 57,* 641-650.

Storms, M. D. (1973). Videotape and the attribution process: Reversing actors' and observers' points of view. *Journal of Personality and Social Psychology, 27,* 165-175.

Swarns, R. L. (2007). So far, Obama can't take black vote for granted. *New York Times.* Retrieved March 22, 2010, from http://www.nytimes.com/

Trevarthen, C. (1979). Communication and cooperation in early infancy. In M. Bullowa (Ed.), *Before speech: The beginning of human communication* (pp. 321-347). London: Cambridge University Press.

Vescio, T. K., Sechrist, G. B., & Paolucci, M. P. (2003). Perspective taking and prejudice reduction: The mediational role of empathy arousal and situational attributions. *European Journal of Social Psychology, 33,* 455-472.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54,* 1063-1070.

Weiner, B. (1980). A cognitive (attribution)-emotion-action model of motivated behavior: An analysis of judgments of help-giving. *Journal of Personality and Social Psychology, 39,* 186-200.

Weiner, B., Perry, R. P., & Magnusson, J. (1988). An attributional analysis of reactions to stigmas. *Journal of Personality and Social Psychology, 55,* 738-748.

Wheeler, S. C., DeMarree, K. G., & Petty R. E. (2005). The roles of the self in priming-to-behavior effects. In A. Tesser, J. V. Wood, & D. Stapel (Eds.), *On building, defending, and regulating the self: A psychological perspective* (pp. 245-272). New York: Psychology Press.

Wheeler, S. C., DeMarree, K. G., & Petty R. E. (2007). Understanding the role of the self in prime-to-behavior effects: The Active-Self account. *Personality and Social Psychology Review, 11,* 234-261.

Wheeler, S. C., Jarvis, W. B. G., & Petty, R. E. (2001). Think unto others: The self-destructive impact of negative racial stereotypes. *Journal of Experimental Social Psychology, 37,* 173-180.

Wheeler, S. C., & Petty, R. E. (2001). The effects of stereotype activation on behavior: A review of possible mechanisms. *Psychological Bulletin, 127,* 797-826.

Williams, H. M., Parker, S. K., & Turner, N. (2007). Perceived dissimilarity and perspective taking within work teams. *Group & Organization Management, 32*, 569-597.

Wyer, N. A. (2007). Motivational influences on compliance with and consequences of instructions to suppress stereotypes. *Journal of Experimental Social Psychology, 43,* 417-424.

Wyer, N. A., Sherman, J. W., & Stroessner, S. J. (2000). The roles of motivation and ability in controlling the consequences of stereotype suppression. *Personality and Social Psychology Bulletin, 26,* 13-25.

Zahn-Waxler, C., Radke-Yarrow, M., & Brady-Smith, J. (1977). Perspective-taking and prosocial behavior. *Developmental Psychology, 13*, 87-88.