IMPLICIT HOMOPHILY:
EFFECTS OF IMPLICIT OUTGROUP BIAS ON INGROUP AFFILIATION

Drew Sédar Jacoby-Senghor

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF PSYCHOLOGY
Adviser: Stacey Sinclair

June 2014
Abstract

Across six studies, we test whether implicit anti-Black bias shapes White participants’ affiliative responses toward White targets who are comfortable around Blacks as a function of perceived similarity, a process we term *implicit homophily*. Participants with higher implicit anti-Black bias reported more negative initial affiliative responses toward targets with Black friends than did participants with low implicit bias (Studies 1–3). Also, participants with greater implicit anti-Black bias reported less perceived similarity with White targets with Black friends, but not targets believed to be randomly paired with a Black other. Perceived similarity, in turn, statistically mediated the relationship between participants’ implicit bias and affiliative responses to ingroup targets (Study 2). Furthermore, by manipulating participants’ perceived similarity to targets, we found that implicit anti-Black bias predicted their affiliative responses to targets only when they did not already know how similar targets’ outgroup experiences were to their own (Study 3). Providing ecological and behavioral evidence of implicit homophily, women’s implicit anti-Black bias predicted their likelihood of having Facebook friends with Black friends (Study 4). Finally, implicit homophily effects were found to also occur as a function of nonverbal behavioral cues, such that participants’ implicit anti-Black bias predicted affiliation toward targets who were more nonverbally comfortable around Black strangers (Study 5). The subjective comfort implied by targets’ nonverbal behavior was also more important than objective treatment of the Black stranger (Study 6). Across studies, effects persisted above and beyond the effects of implicit pro-White bias (Studies 1, 2, & 4) and explicit racial bias (Studies 1–6). Implications for research on stigma by association, extended contact, affiliation, and network formation are discussed.
# TABLE OF CONTENTS

Abstract...........................................................................................................................................iii

Acknowledgements........................................................................................................................................v

Introduction.........................................................................................................................................................1

Section 1: Implicit homophily via friend choices.........................................................................................11

  Chapter 1: Implicit outgroup bias predicts affiliation between ingroup members...........................................14

  Chapter 2: Perceived similarity mediates implicit homophily...........................................................................21

  Chapter 3: Specifying the type of perceived similarity.....................................................................................32

  Chapter 4: Implicit homophily patterns in friendship networks....................................................................40

Section 2: Implicit homophily via nonverbal behavior.....................................................................................47

  Chapter 5: Nonverbal behavior as social cue of similarity.................................................................................49

  Chapter 6: Inferring subjective experience of social stimuli..............................................................................54

General Discussion.............................................................................................................................................60

References.........................................................................................................................................................67
ACKNOWLEDGEMENTS

To my dissertation committee, Stacey Sinclair, Nicole Shelton, Nick Turk-Browne, Susan Fiske, and Andy Conway, thank you for the time you’ve contributed to bringing this dissertation to fruition. To my advisor, Stacey Sinclair, thank you for your remarkable patience, compassion, and support. And thank you for teaching me how to do research that is worth being done, research of which I can be proud. To my other advisor, Nicole Shelton, thank you for the dedication, guidance, and encouragement you so selflessly provide your students despite being one of the busiest people I know. To both Stacey and Nicole, I am a better scientist and a better person for my time alongside you.

Now for all the people who don’t care about the subject of this dissertation, but helped make it possible anyway. Mom, I left for college wanting to make enough money to pay off your house. I regret to inform you that’s not happening. But I did write this really long paper and you should know that you being amazing at what you do (all the things you do, including being my mom) might be the single biggest reason I’ve managed to make it this far. You’ve been someone to look up to, someone to fall back on, one of my best friends, and a giver of hugs when they were most needed. You've given up more than I’ll know in order to be amazing at those things, but even the small amount that I am aware of fills me to the brim with appreciation for you. Gabi, thank you for everything you’ve done to make sure I have a family and a home to come back to. It’s made all the difference to me. I love you and I’m honored to join you in showing up my father with our advanced degrees. Akai and Kenya, being away so much, sometimes it seemed like the best way to be a good brother to you was to work to be a person of whom you’d be proud. Thank you for being my motivation. Big Mama, you taught me that legends are made when one is unrestrained by the expectations of others yet committed to something larger than
oneself. Thank you for being the living legend you are, and for demanding, above all else, hard work, fortitude, humility, and self-respect.

To my cohort and those few others who have been constants during my time at Princeton: I’m really, really glad I had you, because this was way hard sometimes. Andreana, you have a knack for ignoring my requests, but that’s also probably why I made it to this acknowledgments page. You have been a better friend to me than I could admit to needing. Dan, thank you for being a kindred spirit in the last place I’d expect: My own apartment. Rarely did I appreciate my mind as much as when we used ours together. You, more than most things in my life, helped remind me what I love in what I do and how satisfying the directed implementation of intellect can be. Kyle, yours has reminded me that kindness is as important as anything else I attempt. Thank you for being the kind of person who makes it easy to take goodness for granted. Peter, grad school is tedious and long and pedantic and, sometimes, a bunch of butts. But you made my life a lot more remarkable, having us crossing unlit expanses and calling it adventure. For all our imagined worlds that were better than the ones we lived in, we’ve somehow found ourselves in one. Good on us. Neither of us really needed more family, but you are that to me, and I’m happy about it. Matt, your friendship kept me grounded and maintained my perspective. You are also a completely ridiculous person. You just happen to be one of the best men I know. This fact is totally unexpected. Mike, you are an older brother to me. From my first train ride out of Princeton to this impending departure from it, you’ve been there with your guidance and your caring. I hope you know the extent of my appreciation. Shane, your inclusive enthusiasm is one of the best features about any person I know. It made all our time here, and our time together, better. Also, I drink more now because of you.
Finally, thanks to Annie Tornabene, Carina Chang, Crystal Styron, Eri Andriola, Grace Oh, Ivy Maina, Maggie Hua, Maddy Bernstein, Rich Daker, and Steph Colello, the research assistants who ran my studies because I couldn’t, coded videos they hated, transcribed data, and were really great people about it.

Now onto the science stuff.
INTRODUCTION

It has been well documented that implicit outgroup bias impairs individuals’ intergroup interactions (Dasgupta, 2004; Greenwald, Poehlman, Uhlmann, & Banaji, 2009; Jost et al., 2009). Despite this wealth of research, the impact of outgroup bias on ingroup affiliation may be currently underappreciated. In the present research we propose and test a process called implicit homophily, whereby one’s implicit outgroup bias predicts ingroup affiliation via perceived similarity with another. We hypothesize that because implicit bias predicts one’s responses toward outgroup members, it may also predict one’s perceived similarity to ingroup members seen engaging in outgroup interaction. An individual with high anti-Black implicit bias, for example, may perceive a person who is comfortable around Blacks as being dissimilar to herself. In turn, one’s level of perceived similarity to another person should predict their affiliative responses toward them (Berscheid, 1985). As a result, ingroup members with similar levels of implicit bias, and therefore similar automatic responses to outgroup members, should be more likely to affiliate than ingroup members with dissimilar levels of implicit bias. We test implicit homophily in an interracial context and show that implicit racial bias predicts liking of ingroup members as a function of their associations with members of racial outgroups.

Implicit homophily potentially has important implications insofar as it reveals the social functionality of implicit outgroup bias beyond direct outgroup interaction and evaluation. Implicit homophily also suggests new possible antecedents of perceived similarity, specifically that individuals with common levels of bias against a third party may experience similarity-based attraction, even with limited awareness of the bias itself. The proposed process also relates to
stigma by association (Goffman, 1963) and provides a new perspective on the mechanisms that may cause non-stigmatized individuals to be devalued in the presence of stigmatized others.

Implicit homophily also has practical implications. For instance, it suggests that individuals’ social networks are likely to be disproportionally populated by people with implicit biases similar to own. As a result, people with greater implicit bias not only will be less likely to have direct intergroup contact, but also be less likely to have friends with outgroup friends. Such lack of social network diversity precludes those within it from experiencing indirect contact, what some researchers consider to be one of the most effective and socially viable means of attenuating outgroup biases (Wright, Aron, McLaughlin-Volpe & Ropp, 1997).

**Effect of Implicit Intergroup Bias on Intergroup Interaction**

Despite the decline in outspoken bigotry in recent years (Bobo, 2001; 2009), implicit bias remains a pervasive and persistent form of group-based prejudice that appears to operate independently of explicit attitudes (Nosek et al, 2007; Nosek & Smyth, 2007; Greenwald et al., 2009). Although individuals might lack direct access to and control over the magnitude of their implicit bias (Greenwald, McGhee, & Schwartz, 1998), its ramifications within intergroup interactions are felt both by the individuals involved in the interaction and can be discerned by others. Whites with high, compared to low, implicit racial bias exhibit higher cortisol and catabolic hormone levels when interacting with Black versus White others, factors indicative of a “fight-or-flight” response, (Blascovich, Mendes, Hunter, Lickel, & Kowai-Bell, 2001; Mendes, Blascovich, Lickel, & Hunter, 2002; Page-Gould, Mendoza-Denton, & Tropp, 2008). People higher in implicit bias are also more likely to experience cognitive depletion subsequent to an intergroup interaction, indicating that such interactions require effortful self-regulation (Govorun & Payne, 2006; Park, Glaser, & Knowles, 2008; Richeson et al., 2003; Richeson & Shelton,
2003; Richeson, Trawalter, & Shelton, 2005; Richeson & Trawalter, 2005). Furthermore, these low-level physiological responses seem to inform people’s conscious expectations for intergroup interactions: Whites with higher levels of implicit racial bias anticipate that they will be less comfortable during an interaction with a Black person than do Whites with lower levels of bias (Towles-Schwen & Fazio, 2003).

Importantly, people spontaneously express their discomfort, and observers are able to indirectly discern another person’s level of implicit bias, in a number of ways. For example, a White person’s implicit racial bias predicts observers’ ratings of their nonverbal friendliness (e.g., eye contact, smiling; Fazio Jackson, Dunton, & Williams, 1995), hesitation and errors in speech (Dovidio, Kawakami, & Gaertner, 2002; McConnell & Liebold, 2001), seating distance (Bessenoff & Sherman, 2000; Fazio et al., 1995), eyeblink (Dovidio et al., 1997), and openness, comfort, and conversation length (Dasgupta & Rivera, 2006) during interracial interactions. Implicit outgroup bias also influences the diversity of one’s social network in ways that are identifiable by others. Specifically, implicit outgroup bias is negatively associated with one’s willingness to engage with outgroup members (Ashburn-Nardo, Knowles, & Monteith, 2003; Heider & Skowronski, 2007; Gabriel, Banse, & Hug, 2007; Neumann, Hulsenbeck, & Seibt, 2004, Towles-Schwen & Fazio, 2003) and quantity of outgroup contact (Aberson & Haag, 2007; Aberson, Shoemaker, & Tomolillo, 2004; Turner, Hewstone, & Voci, 2007). Aberson, Shoemaker, and Tomolillo (2004), for example, found that greater implicit racial bias among Whites was related to having fewer close Black friends. In sum, although one’s level of implicit bias may not be directly accessible (but see Hahn, Judd, Hirsh, & Blair, in press for evidence that people may have more awareness of their implicit attitudes than previously believed), there
should be noticeable differences in the intergroup behavior and social milieus of individuals high, as opposed to low, in implicit outgroup bias.

Given the visceral and visible nature of these social experiences, it is likely that perceivers are able to sense whether they have similar or dissimilar levels of comfort around and contact with members of particular outgroups compared to someone else. We argue that this level of perceived similarity forms a basis for people’s affiliative judgments toward ingroup members.

**Effect of Similarity on Affiliation**

Perceived similarity between individuals is a well-documented determinant of interpersonal attraction and affiliation (Berscheid, 1985; Byrne, 1971; Newcomb, 1961; Tesser, 1993). Even relatively superficial forms of similarity are known to positively impact interpersonal connections. These include physical attractiveness (Berscheid, Dion, Walster, & Walster, 1971; Walster, Aronson, Abrahams, & Rottman, 1966), demographic information (Brewer, 1979; Byrne, Clore, & Worchel, 1966; Smith & Henry, 1996), and product preferences (Ziegler & Golbeck, 2006; Golbeck, 2009). Even incidental similarities that participants might not expect would matter, such as shared birth dates or letters in one’s name, have been shown to improve affiliative responses (Burger, Messian, Patel, del Prado, & Anderson, 2004; Jones, Pelham, Carvallo, & Mirenberg, 2004; Miller, Downs, & Prentice, 1998).

Moreover, emerging research from a number of theoretical perspectives suggest that more meaningful and personally relevant forms of similarity, such as that implied by similar intergroup experiences and preferences, are particularly potent sources of affiliative desire. Research on “I-sharing” shows that when we believe we have had the same subjective experience of something as another person, we like them more (Pinel & Long, 2012; Pinel,
Long, Landau, Alexander, & Pyszczynki, 2006). In the domain of close relationships, Murray and colleagues (Murray, Holmes, Bellavia, Griffin, & Dolderman, 2002) show that romantic couples that perceive more similarity in their values and daily experiences are more likely to be satisfied with their relationships. Finally, shared reality theory postulates that perceiving similarity in subjective experiences is a necessary component of social relationship formation and maintenance, and it is thus a primary determinant of relational outcomes (Hardin & Conley 2001; Hardin & Higgins, 1996).

An alternative prediction might be that because of the strong normative pressures supporting egalitarianism toward certain groups, people will prefer to affiliate with others who appear to have positive intergroup experiences over those who are similar on this dimension. However, there is some evidence to support the possibility that people will affiliate with those expressing group preferences: Castelli and colleagues (2001; 2003; 2005; 2009) have shown that rather than distancing themselves from non-egalitarian others, people tend to favor ingroup members who employ group stereotypes more than those who do not. In addition, Conley, Rabinowitz, and Hardin (2010) showed that pairs of participants primed to think about the racially charged topic of the O.J. Simpson trial had more positive interpersonal perceptions and outcomes if they shared consensus on the topic compared to pairs that did not. Together, this work suggests that despite egalitarian norms, people may like those who seem to have subjective experiences of outgroup members that are similar to their own.

We propose that because implicit outgroup bias predicts subjective experiences of intergroup interactions, it should also predict perceivers’ affiliative responses toward ingroup members when observing their intergroup interactions or social networks. We furthermore propose that perceived similarity to the ingroup target forms the link between implicit outgroup
bias and ingroup affiliation. Greater implicit outgroup bias should be negatively related to feelings of similarity toward ingroup members viewed as having positive outgroup contact which should, in turn, affect willingness to affiliate with them.

**Expanding Stigma by Association Theory**

Our implicit homophily hypothesis builds upon stigma by association theory to expand understanding of how contact with outgroup members impacts how one is socially evaluated. According to existing stigma by association research, perceivers devalue non-stigmatized individuals who have been linked to someone from a stigmatized group (Goffman, 1963). For example, a heterosexual man might be evaluated more negatively for associating with a homosexual, as opposed to heterosexual, male other (Neuberg, Smith, Hoffman, Russell, 1994).

The present research contributes to, and extends, this literature in two important ways. First, we propose perceived similarity with a target person as a novel mechanism by which individuals are devalued because of their interpersonal associations with outgroup members. Second, using this new framework, we provide clarification regarding the factors that should moderate the degree to which individuals’ interactions with outgroup members affects ingroup members’ desire to affiliate with them.

Regarding mechanism, the present work represents a new theoretical perspective on how viewing intergroup interactions affects social evaluations and assortment. Across both clinical and social psychology, stigma by association has most commonly been characterized as involving transference of negative affect or perceptions from stigmatized individuals to associated non-stigmatized targets. For example, Goffman (1963) asserts that stigma can “spread from the stigmatized individual to his close connections” (page 30). Stigma researchers have shown that such spread can occur via mere physical proximity (Hebl & Mannix, 2003; Penny &
Haddock, 2007) or because perceivers assume that individuals who fraternize with the stigmatized have undesirable qualities or are otherwise contaminated by their relationship with the stigmatized person (Burke & Sher, 1990; Corrigan & Miller, 2004; Corrigan, Miller, & Watson, 2006; Mehta & Farina, 1988; Rozin, Markwith, & Nemeroff, 1992; Sigelman, Howell, Cornell, Cutright, & Dewey, 1991). Cognitive accounts similarly postulate that negative affect or perceptions spread from the stigmatized person to the non-stigmatized target via stimulus conditioning and misattribution of affect (Baeyens, Hermans, & Eelen, 1993; Jones, Fazio, & Olson, 2009; Walther, 2002; Walther, Nagengast, & Trasselli, 2005). That is, when stigmatized and non-stigmatized individuals are presented together, perceivers begin to associate the negativity of one with the other. In essence, these explanations all suggest that some of the negativity associated with stigmatized individuals transfers to non-stigmatized individuals with whom they are proximal.

In contrast to this general perspective, implicit homophily offers the novel suggestion that a target’s association with stigmatized outgroup members affects social evaluations through interpersonal comparisons between oneself and the target. In this way, association with stigmatized individuals can signal a degree of similarity between non-stigmatized targets and perceivers in terms of beliefs, intergroup anxieties, and/or life experiences, which influences desired affiliation. We therefore argue that not all effects that appear to be stigma by association are exclusively due to transference of negative evaluations from the stigmatized individual to the non-stigmatized target.

One reason why stigma by association researchers may have focused on the transference perspective to such an extent is that this perspective fits well with the types of stigma they have tended to study. Specifically, most of the relevant research on stigma by association has
concerned gay (Neuberg et al., 1994; Sigelman, Howell, Cornell, Cutright, & Dewey, 1991; Swim, Ferguson, & Hyers, 1999), weight-based (Hebl & Mannix, 2003, Pryor, Reeder, & Monroe, 2012), mental and physical illness (Angermeyer, Schulze, & Dietrich, 2003; Blum, 1991; Corrigan & Miller, 2004; Corrigan, Miller, & Watson, 2006; Haber, Roby, & High-George, 2011; Halter, 2008; Mehta & Farina; 1988; Poindexter & Linsk, 1999; Rozin, Markwith, & McCauley, 1994; Rozin et al., 1992; Wight, Aneshensel, Murphy, Miller-Martinez, & Beals, 2006; Wight, Aneshensel, & Wongvipat 2000), and mental disability stigmas (Birenbaum, 1970; 1992; Gray, 2002; Green, 2003; Phelan, 2005). Although social psychologists often study prejudice and stigma in terms of Black/White relations, we are aware of only one peer-reviewed study (Pryor et al., 2012; Study 3) that addresses whether stigma by association occurs within this specific group dynamic.

There are a number of reasons why findings about these more studied stigmas do not necessarily extend to interracial situations. For one, these stigmas could be construed as more transmissible than Black stigma. In the case of gay (Cottrell & Neuberg, 2005; Cottrell, Richards, & Nichols, 2010) and weight-based stigmas (Park, Schaller, Crandall, 2007; Vartanian, 2010), negative evaluations are in part related to feelings of disgust toward these groups. In turn, work on “magical thinking” has shown that people perceive objects that come in contact with disgusting stimuli as acquiring the stimuli’s properties (e.g., Rozin, Millman, & Nemeroff, 1986; see also Nemeroff & Rozin, 2000). In contrast, racial stigma is often based on group-specific social threats (Sears, 1988; Tajfel & Turner, 2004; Cottrell & Neuberg, 2005). In addition, studies on stigma by association with mental illness and disability often measure the impact on family members. In these cases, perceived heritability or assumptions that family members are responsible for the stigmatized individual’s condition may drive effects (e.g, Corrigan et al.,
2006). Such familial concerns are not relevant to ingroup targets seen with racial outgroup members. Finally, with the exception of mental disability, prejudice against Blacks is also deemed less permissible than against these other stigmatized groups (Crandall, Eshleman, & O’Brien, 2002).

Given these reasons, race-based stigma by association may involve different mechanisms than these more commonly studied forms. Furthermore, it is quite possible that stigma transference and implicit homophily can act simultaneously to affect evaluations of a target person. We believe that in order to separate the unique effects of each process, it is important to study associative devaluation with groups other than those typically considered, especially ones that may be less prone to stigma transference. By specifically focusing on Black/White group distinctions in the present work, we intend to begin addressing these concerns.

The suggestion that perceived similarity mediates the relationship between targets’ intergroup contact and perceivers’ affiliation with them has implications for the factors that should moderate the predicted effects. Most notably and as we have already expressed, we posit that implicit outgroup bias should predict the extent to which perceivers wish to affiliate with ingroup targets because it correlates with experiences and behaviors that can form the basis for similarity inferences. Within the stigma by association literature, surprisingly little work has addressed this fundamental question of whether one’s degree of bias toward a stigmatized group predicts the extent to which he or she devalues individuals associated with members of that group (but see Pryor et al., 2012; Sigelman et al., 1991). Establishing this relationship is necessary to determining whether stigma by association is caused by ingroup favoritism or outgroup antipathy, the environmental conditions that produce the phenomenon, when it is most likely to occur, and how it can be ameliorated.
Another important implication of implicit homophily is that the nature of the relationship between the target and outgroup other should matter to the perceiver. Specifically, perceivers’ implicit outgroup bias should only predict their affiliative responses toward an ingroup target when the presence of the outgroup member reveals clues about the level of similarity shared by perceiver and target. Therefore, when an ingroup target and outgroup other are presented simultaneously but do not share a meaningful connection, perceivers’ implicit bias should not predict affiliative responses to targets. Furthermore, we predict that observers make inferences about targets subjective experiences during outgroup interactions and that this inferred information shapes observers’ affiliative responses. Therefore, rather than devaluing all ingroup targets who interact with an outgroup person, implicit outgroup bias should differentially predict responses to people who have positive or negative outgroup interactions.

In contrast, the stigma by association literature is inconclusive regarding whether such effects are moderated by the apparent relationship or quality of interaction between the stigmatized individual and non-stigmatized target. While some studies have shown that the quality of contact between the two does not moderate perceivers’ evaluations (Hebl & Mannix, 2003, Penny & Haddock 2007), others have found differently (Rozin et al., 1992; Sigelman et al., 1991; Pryor et al., 2012). Especially relevant to the current paper, the single previous examination of Black stigma by association that we are aware of (i.e., Pryor et al., 2012) did not manipulate the relationship between the stigmatized person and the target of evaluation. Through our investigation of implicit homophily, the current research provides the first test of whether bias and quality of association together predict associative devaluation in an interracial context.
SECTION 1: IMPLICIT HOMOPHILY VIA FRIEND CHOICES

In this first set of studies, we test the occurrence of implicit homophily in the context of Black/White friendships and begin to illustrate its consequences. We test whether Whites’ implicit racial bias is related to their affiliative responses to White targets associated with a Black friend across two paradigms and using three different measures of implicit racial bias (Studies 1–4). We also examine whether this is a product of ingroup favoritism (i.e., pro-White bias) or anti-Black bias (Studies 1, 2, & 4). In all experiments explicit racial bias is also measured. Thus, its effects can be explored and controlled for.

We also seek direct and indirect support for the contention that implicit homophily occurs via perceptions of similarity. In Study 2 we examine several potential reasons that one’s implicit bias might predict affiliative tendencies toward an ingroup member associated with an outgroup member, including perceived similarity. In Study 2 we also manipulate the apparent relationship between ingroup targets and the people they are presented with to determine whether the effect depends on a meaningful social connection existing between the two. In Study 3 we conduct another test that perceived similarity drives implicit homophily by directly manipulating participants’ perceived similarity to ingroup targets regarding their comfort with racial outgroup members.

In Study 4, we further explore the consequences of this phenomenon by testing whether patterns of data found in the lab are reflected within an actual online friendship network (i.e., Facebook). If Whites who are high in implicit racial bias have fewer friends with Black friends,

---

1 The work described in chapters 1—4 is currently under revision for publication in *The Journal of Personality and Social Psychology.*
as the notion of implicit homophily predicts, this suggests they are missing opportunities for prejudice reduction afforded by indirect contact.

Despite our specific predictions for how ingroup targets seen with outgroup peers are perceived, the effect of implicit bias on ingroup targets with ingroup peers is less clear. On the one hand, one might predict that because implicit bias against Blacks does not directly relate to feelings about Whites, it should not predict White participants’ affiliative responses to White targets seen with White others. Implicit bias against Blacks also might not predict responses toward these White targets because ingroup interactions are too commonplace to be seen as indicative of the target’s subjective intergroup experiences. On the other hand, in contexts where White targets are seen associating with Blacks, Whites associating with Whites may be evaluated comparatively. In this case, implicit bias against Blacks should predict more positive affiliative responses toward Whites with Whites. In Studies 1-3 we show each participant several White targets paired with either Blacks or Whites and are therefore able to test this secondary question.

In addition to testing the effect of implicit racial bias, we also control for and report the effect of explicit racial bias on affiliative responses toward targets in all four studies in the present research. However, we did not have specific predictions for whether explicit bias would significantly predict affiliative responses. As mentioned previously, explicit prejudice is now widely counternormative. Because people with high explicit bias are, by definition, aware of their bias, they may be vigilant against expressing their bias in situations where they feel their responses are being monitored. With certain participant samples, the range of explicit bias may even be constrained. Finally, the role of explicit bias in racial stigma by association is unconfirmed; Pryor and colleagues (2012) found that explicit bias did not predict more negative evaluations of a White man seen having a conversation with a Black man at a work gathering,
whereas implicit bias did. We therefore remained agnostic to the explanatory power of explicit bias while predicting that the effect of implicit bias on affiliative responses would endure above and beyond the effect of outspoken prejudice.
CHAPTER 1: IMPLICIT OUTGROUP BIAS PREDICTS AFFILIATION BETWEEN INGROUP MEMBERS

Study 1 investigated whether implicit anti-Black bias, above and beyond the effect of implicit pro-White bias, predicts Whites’ affiliative responses to White targets seen with Black or White friends. Participants saw side-by-side still photograph pairs of ostensible friends and were asked to rate each in turn. Some of these photograph pairs were of Whites with Blacks while others were of Whites with Whites. We predicted that greater implicit anti-Black bias would relate to more negative ratings of Whites with Black friends.

Method

Participants. White undergraduates (N = 27; 9 male) were recruited for pay or course credit.

Procedure. Participants were recruited to take part in a study about what makes people become friends. Upon arriving, they first completed a sequential subliminal priming task as an unconscious measure of implicit bias (Lowery, Hardin, & Sinclair, 2001; Sinclair, Lowery, Hardin, & Colangelo, 2005). Next they were told that the researchers were interested in the perceived friendliness of various people and that they would be rating pictures of friends. They were then presented with a series of twenty-four side-by-side same-gender headshot photographs with instructions above each reading “You are rating the friend on the LEFT”. The stimuli included Whites with Blacks as well as Whites with Whites. While viewing each pair, participants answered questions assessing their affiliative responses to the target. Upon completion of the ratings task, participants completed a measure of explicit prejudice, basic demographic questions, and then were debriefed and dismissed.
**Materials**

**Sequential subliminal priming measure.** The subliminal priming task (Lowery, Hardin, & Sinclair, 2001; Sinclair, Lowery, Hardin, & Colangelo, 2005) consisted of 128 trials in which participants were asked to fixate on a dot in the middle of the screen. Black and white pictures of 64 Black (32 female) and 64 White (32 female) faces were then presented parafoveally for 17 ms in random order at the four corners of the screen and then covered with a picture of a sunflower as a backward mask. Next, the word good or bad appeared at the center of the screen where the fixation dot had been. Participants then identified which word had appeared by pressing a corresponding key on the keyboard (e.g., the K key for good and D key for bad). The word to which each key corresponded was counterbalanced across participants. Reaction time was recorded from the onset of the word good or bad until participants provided their response.

Mean reaction times to the words good and bad as a function of exposure to Black and White faces were recorded to serve as our measure of automatic racial bias. Reaction times less than 300 ms or greater than 3,000 ms were coded as missing values, and the remaining values were log transformed (Sinclair, Lowery, Hardin, & Colangelo, 2005). To address our hypothesis that implicit anti-Black bias would independently serve as a significant predictor of perceptions of White targets as a function of with whom they are interacting, we calculated one score for implicit anti-Black bias and a second score for implicit pro-White bias. Implicit anti-Black bias, in which higher numbers indicated greater negativity toward Blacks, was calculated by subtracting the response time for the word bad when primed with a Black face from response time to the word good when primed with a Black face. Implicit pro-White bias, in which higher numbers indicated greater positivity toward Whites, was calculated by subtracting response time to the word good when primed with a White face from response time to the word bad when
primed with a White face. Implicit anti-Black and pro-White bias scores were negatively correlated at a marginally significant level ($r = -.343$, $p = .080$)

**Photograph pairs.** While participants were told that the paired photographs were friends, in reality the photographs were taken from the Productive Aging Database (Minear & Park, 2004). Targets were selected from the database to be roughly equivalent in pretest ratings on attractiveness, age, familiarity, and trustworthiness. In all, participants saw 24 photograph pairs presented in random order, sixteen of which were critical trials. In critical trials, targets on the left side of the screen were always White. For eight of these trials, White targets were paired with a Black friend (4 female pairs). For the other eight trials, White targets were paired with a White friend (4 female pairs). The race of a given target’s friend was counterbalanced across participants such that each target was shown with a White and a Black friend. The remaining 8 filler trials were included to disguise the racial nature of the task. In these trials, the targets being rated (4 female pairs) were various non-White individuals paired with friends of various races.

**Affiliative response measure.** Affiliative response to the target was measured using five questions randomly presented one-by-one under each pair of photographs (e.g., “I would expect to get along easily with this person”, “I like this person”, “I would expect to have a smooth interaction with this person”, “I feel warmly toward this person”, “To what extent do you think you would want to become friends with this person?”; 1 = “not at all”/“strongly disagree”, 6 = “extremely”/“strongly agree”; $\alpha = .92$). For our analyses, a participant’s ratings for all photos of targets paired with a Black person were averaged together to obtain a mean affiliative response rating of targets with Black friends. The same procedure was conducted for ratings of White targets paired with a White person.
Explicit prejudice measure. After the ratings task, participants completed the Ambivalent Racism scale (Katz & Hass, 1988; 1 = “strongly disagree”, 8 = “strongly agree”), from which explicit anti-Black ($\alpha = .68$) and pro-Black ($\alpha = .82$) attitude scores were calculated. Explicit anti-Black and Pro-Black scales were not significantly correlated ($r = .03, p = .873$), and so were entered separately in all analyses.

Analytical Plan for Repeated-Measures Designs

We used Judd, Kenny, and McClelland’s (2001) method for testing moderation with repeated-measures designs for all such studies reported in this manuscript (Studies 1—3, 5, & 6). Doing so allowed us to examine whether participants’ implicit anti-Black bias predicts affiliative responses to White targets as a function of the race of the person they were paired with, above and beyond the predictive value of implicit pro-White bias and explicit racial prejudice.

Accordingly, in each of these studies we computed a difference score of affiliative response ratings for targets seen with a White person and targets seen with a Black person and entered this number as the dependent variable into a multiple regression analysis. All between-subjects factors (e.g., implicit anti-Black and pro-White scores and explicit anti- and pro-Black bias) were then entered simultaneously as mean-centered predictors. Thus, we were able to test the interaction of “friend” race and the between-subjects factors of implicit and explicit racial bias.

To probe significant interactions and determine which simple-slopes effects they were driven by, we separated the difference score dependent variable and conducted two individual regression analyses for affiliative responses to Whites with Blacks and Whites with Whites. For all analyses, we continued to include all available measures of implicit and explicit racial bias as predictors in the regression analysis. Finally, to test whether implicit racial bias would lead to a

---

2 For all studies reported in this paper, the interactions of implicit and explicit racial biases did not significantly moderate any reported effects.
devaluation of Whites with Blacks compared to Whites with Whites, and to account for participants relative rating tendencies of each target type, we also included participants’ mean-centered affiliative response to other White targets they rated as a covariate. For example, when testing affiliative response ratings of Whites with Blacks as the dependent variable, ratings of Whites with Whites was entered into the model as a predictor, and vice versa when testing ratings of Whites with Whites as the dependent variable.

**Results**

Using this analytic strategy, a significant interaction of friend race and implicit anti-Black bias emerged ($\beta = .470$, $t(22) = 2.60, p = .016$). Interactions of friend race with explicit anti- ($\beta = .580$, $t(22) = 3.44, p = .002$) and pro-Black ($\beta = -.363$, $t(22) = -2.23, p = .037$) biases were also found. The interaction of friend race and implicit pro-White bias was not significant ($p = .449$).

**Effect of implicit anti-Black bias on affiliative response.** Probing this significant interaction, we first analyzed the effect of implicit anti-Black bias on ratings of Whites with Blacks. We found that as implicit anti-Black bias increased, affiliative responses to White targets with Black friends became more negative ($\beta = -.381$, $t(21) = 3.09, p = .006$). We also found that implicit anti-Black bias also predicted affiliative responses to White targets with White friends, but in reverse: Greater implicit anti-Black bias predicted more positive affiliative responses to targets with White friends ($\beta = .346$, $t(21) = 2.29, p = .033$; see Figure 1).

**Effect of explicit bias on affiliative response.** The regression models described above also produced estimates of the effect of explicit anti- and pro-Black attitudes on responses to White targets. Participants’ affiliative responses to targets with Black friends significantly decreased when participants held more negative explicit attitudes toward Blacks (anti-Black: $\beta = -.365$, $t(21) = -3.05, p = .006$; pro-Black $\beta = .371$, $t(21) = 3.27, p = .004$), while affiliative
Figure 1. Effect of implicit anti-Black bias on affiliative response toward White targets paired with White and Black friend.

responses to targets with White friends significantly increased as anti-Black increased ($\beta = .436$, $t(21) = 3.27, p = .004$) and marginally significantly decreased as pro-Black bias increased ($\beta = -.261$, $t(21) = -1.76, p = .092$).

Discussion

While statistically controlling for explicit racial attitudes, we found that greater implicit anti-Black bias predicted more negative affiliative responses toward Whites with Black friends and more positive responses toward Whites with White friends. Implicit pro-White bias did not predict affiliative responses to either target type. The fact that implicit anti-Black bias predicts responses toward Whites with White friends suggests participants make comparisons of targets across trials. It seems that in the mind of the participants, to associate with a White friend also meant not associating with a Black friend. If the devaluation of Whites with Blacks were solely
caused by the transference of negative affect, as in some accounts of stigma by association, then we should not expect to see implicit anti-Black bias also predicting affiliative responses to Whites with Whites. It appears that some other process, possibly implicit homophily, is at play. However, questions remain as to the reason that participants implicit bias relates to their affiliative tendencies to ingroup targets in intergroup contexts. In Study 2, we tested potential underlying mechanisms.
CHAPTER 2: PERCEIVED SIMILARITY MEDIATES IMPLICIT HOMOPHILY

In Study 2 we directly tested our hypothesis that participants’ perceived similarity with the target mediates the link between participants’ implicit anti-Black bias and their affiliative responses to targets. We simultaneously tested two other potential mediators arising from plausible alternative explanations of the effects based on other theories of interpersonal evaluation. The first alternative explanation is suggested by the previously discussed stigma transference literature. From this perspective, non-stigmatized targets take on the negative associations elicited by the stigmatized person with whom they are linked (Baeyens et al., 1993; Burk & Sher, 1990; Coovert & Reeder, 1990; Corrigan & Miller, 2004; Corrigan, Watson, & Miller, 2006; Hebl & Mannix, 2003; Jones et al., 2009; Mehta & Farina, 1988; Penny & Haddock, 2007; Pryor et al., 2012; Rozin et al., 1992; Sigelman et al. 1991; Walther 2002). Therefore, participants should be expected to attribute the same specific negative evaluations to Whites as the Blacks that they are paired with. As a second alternative explanation, the expectancy violation literature demonstrates that individuals who violate expectations, especially expectations for ingroup behavior, are often evaluated more negatively (Bettancourt et al., 1997; Biernat, Vescio, & Billings, 1999; Jussim, Coleman, & Lerch, 1987). From this perspective, affiliative responses are driven not by inferences of the target’s similarity, but rather their likelihood of acting in a manner that is uncommon for ingroup members. In consideration of these literatures, we pitted perceived similarity, stigma-transference, and expectancy violation against one another as three potential mediators.

A second goal of the study was to test whether the effect of participants’ implicit bias on their affiliative responses is contingent upon a meaningful social relationship between the target
and the paired other. According to our predictions, a White target’s interaction with a Black person provides the perceiver with important information about the target’s beliefs and experiences around Blacks. This information is then used as a basis for perceived similarity to the target. Mere proximity (e.g., Hebl & Mannix, 2003; Penny & Haddock, 2007) and evaluative conditioning accounts (Baeyens et al., 1993; Jones et al., 2009; Walther, 2002; Walther et al., 2005) predict that simply seeing a White target paired together with a Black person would be enough to make the perceiver dislike the target. We therefore manipulated the ostensible relationship between the target and paired other to determine whether the effect would disappear when participants were told the photos had been randomly paired.

Method

Participants. White participants who were both U.S. citizens and U.S. residents were recruited for pay via Amazon Mechanical Turk. Given that interpersonal evaluations can differ as a function of whether the target is within or outside of one’s own generational cohort (North & Fiske, 2013), we selected participants who were within range of the targets’ perceived age to ensure participants could view the target as a potential target of affiliation. According to pretest responses \( N=42 \), targets’ mean perceived age was 23.24 years \( (SD = 2.19) \). Selecting participants who were within three standard deviations of the mean resulted in a sample of 88 (37 male and 1 participant refusing to answer) participants with an age range of 18 to 30 years old \( (M = 24.48 \text{ years}, SD = 3.69) \).

Procedure. For the current study, we largely replicated the design of Study 2 but with four major changes. First, after they were told that they were participating in a study of how

---

3 Forty-two U.S. residents (33 White, 16 male) pretested these faces on Amazon Mechanical Turk on a 1 (“not at all”) to 7 (“extremely”) scale. The aggregate statistics were calculated for the following traits: Attractiveness \( (\text{mean}=4.14, SD=.54) \), familiarity \( (\text{mean}=3.40, SD=.34) \), trustworthiness \( (\text{mean}=4.75, SD=.34) \), and age \( (\text{mean}=23.24, SD=2.19) \).
friendly people come across as being, participants were randomly assigned to one of two conditions: Half of participants were told that the pairs were friends and saw the instructions “You are rating the friend on the LEFT” while rating each pair. The other half of participants was told the pictures had been randomly paired and saw the instructions “You are rating the person on the LEFT” while rating each pair. Immediately following the initial instructions, participants confirmed whether they were rating the person on the left or the right. Second, participants were presented with six photograph pairs instead of 24. Third, while viewing each photograph pair, participants answered questions measuring responses related to our potential mediators as well as affiliative responses to the target. Finally, instead of a subliminal priming measure of implicit bias, participants completed two single-target IATs as measures of implicit racial bias upon completion of the ratings task.

All other design elements of Study 2 were identical to those in Study 1.

Materials

Photograph pairs. Participants saw the same photographs as in Study 1, but this time viewed only six photograph pairs randomly selected for each participant from the larger pool used in Study 1. Of the six trials, four were critical trials and two were filler trials.

Picture ratings. For all trials, participants rated targets on four dimensions. Affiliative response to the target was assessed with two questions (“I would expect to get along easily with this person.”, “To what extent do you think you would want to become friends with this person?”; α = .82). Perceived similarity was assessed with three questions adapted from the Turban and Jones’ (1988) shared worldview scale (e.g., "This person and I are probably similar in terms of our outlook, perspective, and values."; “This person and I probably see things in much the same way.”; α = .94). Stigma-transference was measured by asking participants the
degree to which targets were likely to evoke three specific forms of affect known to be highly associated with Blacks (Cottrell & Neuberg, 2005; e.g., "How likely is this person to make you feel disgusted/angry/anxious?"; $\alpha = .80$). Finally, expectancy violation was assessed with two questions (e.g., “This person is likely to behave in ways that are surprising to me.”, “How unpredictable does this person seem?”; $\alpha = .60$). All questions were answered on a 6-point Likert scale (1 = “not at all”/“strongly disagree”, 6 = “extremely”/“strongly agree”). All questions were presented in random order for each photograph pair.

**Single Target Implicit Association Tests.** Two Single Target Implicit Association Tests (ST-IAT; Wigboldus, Holland, van Knippenberg, 2006) served as the measures of implicit racial bias. One ST-IAT measured White implicit bias while the other measured Black implicit bias. Within the task, participants see a single discrimination block of the evaluative stimuli (i.e., “good” and “bad”). Next, participants complete a block in which the target stimuli (i.e., either White or Black faces) and positive stimuli are indicated using one keyboard key and negative stimuli are indicated using another. The stimuli pairing is then reversed in the final block. By doing so, the ST-IAT provides a score of evaluative associations for a single target category rather than a relative score as in the traditional IAT. Higher scores indicated greater implicit anti-Black or pro-White bias. Implicit anti-Black bias and implicit pro-White bias were significantly negatively correlated ($r = -.22$, $p = .049$).

**Explicit racial bias measure.** We again measured explicit racial bias using the Ambivalent Racism scale (Katz & Hass, 1988; anti-Black $\alpha = .89$; pro-Black $\alpha = .87$).

**Results**

**Preliminary analyses.** Five participants were removed for having an average response latency of less than 1 second on their target ratings or explicit item responses. Eighty participants
(friends condition: n = 42) remained in the final sample (mean age = 24.46 years, SD = 3.73; 35 male).

**Main analyses**

**Interactions of racial bias with manipulations.** As an addition to our general analytic strategy, photo pair relationship and all two-way interactions between photo pair relationship and each type of bias were included as predictors in our regression model. We found a significant three-way interaction of implicit anti-Black bias, photo pair relationship, and race of the person paired with the target ($\beta = .367, t(73) = 2.29, p = .025$). All other two- and three-way interactions involving implicit anti-Black or pro-White biases were not significant ($p$s > .120). Regarding explicit racial bias, we found a significant two-way interaction of explicit pro-Black bias and race of the person paired with the target ($\beta = .367, t(73) = 2.19, p = .032$). No other two- or three-way interactions involving explicit bias were significant ($p$s > .203).

**Effect of implicit racial bias on affiliative response.** Probing the three-way interaction, we separately analyzed participants who were told that the photos were of friends and those told that the photos had been randomly paired. For participants who were told that the photos were of friends, we found a significant interaction of implicit anti-Black bias and the race of the target’s friend ($\beta = .424, t(37) = 2.76, p = .009$). For targets paired with a Black friend, we found that greater implicit anti-Black bias significantly predicted more negative affiliative responses ($\beta = -.347, t(37) = -2.71, p = .010$). Inversely, greater implicit anti-Black bias significantly predicted more positive affiliative responses to targets paired with a White friend, ($\beta = .294, t(37) = 2.10, p = .043$; see Figure 2).
Figure 2. Effect of implicit anti-Black bias on affiliative response toward White targets when participants are told targets were paired with friends versus were randomly paired.
However, for participants who were told that the photos had been randomly paired, the two-way interaction between implicit anti-Black bias and race of the paired stranger was not significant ($\beta = -.124, t(36) = -.76, p = .450$).

**Effect of explicit racial bias on affiliative response.** Probing the significant two-way interaction of pro-Black bias and race of the paired other with the same predictors as above, we found that for targets paired with Blacks the effect of explicit pro-Black bias was not significant ($\beta = -.015, t(77) = -.13, p = .900$). For targets paired with Whites, greater pro-Black bias significantly predicted more positive affiliative responses ($\beta = .263, t(77) = 2.30, p = .024$).

**Effects of implicit anti-Black bias on potential mediators.** Having observed a significant interaction of implicit anti-Black bias and race of the paired other when participants were told the photo pairs were friends, we tested whether this interaction could be explained by any of the three potential mediators we measured (Baron & Kenny, 1986). As a first step, we created difference scores between ratings of targets with White friends and ratings of targets with Black friends for perceived similarity, stigma-transference, and expectancy violation. These difference scores for each mediator were then entered as dependent variables into separate regression analyses with all four types of racial bias as predictors. For perceived similarity we found the significant interaction of implicit anti-Black bias and race of the target’s friend necessary to suggest mediation ($\beta = .438, t(37) = 2.87, p = .007$). Simple slopes for this interaction are found in the next section. The same interaction of implicit anti-Black bias and race of the target’s friend did not emerge as significant for either stigma-transference ($\beta = -.008, t(37) = .04, p = .965$) or expectancy violation ($\beta = .242, t(37) = 1.51, p = .139$). Moreover, there were no statistically significant effects of implicit pro-White bias, or explicit bias on any of the three possible mediators ($ps > .296$).
Mediation analyses. We proceeded to test perceived similarity as a mediator by simultaneously entering implicit anti-Black bias and the perceived similarity difference score as predictors of the difference in affiliative responses to targets with a White friend and targets with a Black friend. In this model, greater differences in perceived similarity with targets with White friends and targets with Black friends significantly predicted greater difference in affiliative responses to targets with White friends and targets with Black friends ($\beta = .734$, $t(36) = 5.84$, $p < .001$). Furthermore, the effect of implicit anti-Black bias diminished to non-significance ($\beta = .103$, $t(36) = .87$, $p = .392$), suggesting that perceived similarity fully mediated the connection between implicit anti-Black bias and affiliative response to targets paired with ostensible friends. A Sobel (1982) test confirmed the mediation (Sobel’s $z = 2.62$, $p = .009$).

When separately testing the same mediation pathway for targets with Black friends and targets with White friends, the pattern of effects seen with affiliative response was replicated: Greater implicit anti-Black bias predicted decreased feelings of similarity to targets with Black friends ($\beta = -.386$, $t(36) = -2.75$, $p = .009$; see Figure 3). Decreased perceived similarity with targets with Black friends significantly predicted reduced affiliative responses to them ($\beta = .737$, $t(35) = 7.32$, $p < .001$), while the effect of implicit anti-Black bias diminished to non-significance ($\beta = -.059$, $t(35) = -.65$, $p = .523$). Inversely, for targets with White friends, greater implicit anti-Black bias predicted marginally significant increases in feelings of similarity ($\beta = .276$, $t(36) = 1.77$, $p = .085$, $R^2 = .232$; see Figure 3). Greater feelings of similarity with targets with White friends also significantly predicted more positive affiliative responses of them ($\beta = .800$, $t(35) = 9.04$, $t(35) = 9.04$, $p < .001$), and the effect of anti-Black bias again diminished to non-significance ($\beta = .077$, $t(35) = .94$, $p = .353$). Sobel tests confirmed the significant mediation for

---

4 This pattern remained significant when stigma-transference and expectancy violation difference scores were included as simultaneous predictors in the analysis ($\beta = .767$, $t(34) = 6.87$, $p < .001$).
Figure 3. Effect of implicit anti-Black bias on perceived similarity with White targets when participants were told targets were paired with friends versus were randomly paired.
Figure 4. Correlational model with perceived similarity to the target mediating the relationship between implicit anti-Black bias and affiliative response toward Whites with a Black friend.

targets with Black friends and marginally significant mediation for targets with White friends (Black friends: Sobel’s $z = 2.58, p = .010$; White friends: Sobel’s $z = 1.74, p = .082$; see Figure 4).

Discussion

The results provide support for the hypothesis that participants’ perceived similarity to White targets viewed with Black or White friends mediates the relationship between implicit anti-Black bias and affiliative responses. We found that as implicit anti-Black bias increased, perceived similarity to White targets with Black friends decreased, and that perceived similarity in turn predicted more negative affiliative responses to the target. For White targets paired with Whites, we found that greater implicit anti-Black bias predicted greater perceived similarity, and greater perceived similarity predicted more positive affiliative responses. No evidence was found
for either stigma-transference or expectancy violation as significant mediators. Perceived similarity also remained a significant mediator with the inclusion of these two potential alternative mediators in the statistical models.

Participants’ implicit anti-Black bias only significantly predicted perceived similarity and affiliative responses to the target when they were told that target was friends with the paired other, but not when they were told they were viewing randomly paired photos. In contrast to a stigma-transference account, it appears that when participants cannot use the pairing of target and other to infer the subjective experience of the target, the presence of the other ceases to influence their affiliative responses.
CHAPTER 3: SPECIFYING THE TYPE OF PERCEIVED SIMILARITY

After establishing that perceived similarity statistically mediates the effect of implicit anti-Black bias on affiliative responses toward ingroup members, in Study 3 we sought to confirm this finding by directly manipulating this mediator. In addition, we aimed to clarify the specific type of similarity that is inferred from the race of a target’s friend. While there are myriad dimensions on which people may be similar, we have hypothesized that a targets’ friend choices predict how similar participants perceive them to be regarding their comfort interacting with Blacks. If this assumption is true, then directly providing participants with information about how similar a target is regarding her experiences around Blacks should diminish the extent to which implicit anti-Black bias predicts participants’ affiliative responses. However, if participants were provided similarity information on an unrelated but nonetheless important dimension (e.g., personality; see Byrne, Griffitt, & Stefaniak, 1967; Izard, 1960), the effect of implicit anti-Black bias on affiliative responses should persist.

In the current study, we also asked participants about their interest in having a future online interaction with each target. This question was added to test whether participants’ implicit anti-Black bias predicts their stated intentions for interaction with targets as a function of their friend’s race.

Method

Participants. We once again recruited White participants who were U.S. citizens, U.S. residents, and 18 to 30 years old via Amazon Mechanical Turk, resulting in a sample of 78 participants (mean age = 24.53 years, SD = 3.10; 45 male).
**Procedure.** The current study utilized a similar design as Studies 1 & 2. However, participants were sorted into two new conditions. In one condition, participants were given information about how similar their experiences with racial outgroup members were to those of each target (i.e., *experiential similarity* condition). In the other condition, participants were given information about how similar their personalities were to each target (i.e., *personality similarity* condition). The following changes were made to produce these conditions: Before the ratings task, participants completed two surveys, which ostensibly assessed their “social experiences” and personality, and then filled out demographic information. After completing these surveys, we told all participants that they would see pairs of friends and rate one person in each pair. We then told them “We have also collected [social experience/personality] questionnaire responses for each person you will see. Their scores on will be compared to your own. You may use this information at your discretion.” Before rating each target, we provided participants with false feedback about their percent match on one of the questionnaires and fake demographic information for each target. Each participant saw a different match percentage for each target, however the match percentages presented were identical across participants. Identical to Study 2, we presented participants with a total of six photograph pairs.

After rating the targets, participants completed implicit and explicit measures of bias. Due to null effects of implicit pro-White bias in the previous studies, and in an attempt to minimize the experiment’s duration for our online participants, we only measured implicit anti-Black bias in the present study.

**Materials**

**Preliminary Surveys.** At the beginning of the experiment, participants filled out two surveys. One was entitled the “Social Experiences Questionnaire” and ostensibly assessed how
the participant feels in different social situations. In actuality, the questionnaire included a subset of six questions regarding participants’ comfort around members of other ethnic backgrounds (e.g., “I feel uncomfortable about interacting with Black people”; “I feel uncomfortable around people from different backgrounds”, “My friends tend to be similar to me”; $\alpha = .70$). These items were rated alongside six filler questions intended to disguise the racial focus of the experiment (e.g., “I enjoy meeting new people”, “I feel uncomfortable about interacting with elderly people”). The other survey, entitled “Personality Questionnaire”, included a selection of twelve items from the Big Five personality inventory (John & Srivastava, 1999). The order of these two surveys was counterbalanced between participants. After completing these surveys, participants entered some demographic information. Participant race was requested at the end of the study.

Target information. Participants were presented with falsified information about each target before providing their ratings. The information remained on the screen for at least five seconds before participants were given the option to advance to rating the target. Depending on the condition, the information included either targets’ social experiences or personality match with the participant. The match percentage for each target was randomly selected without replacement from a predetermined set, which was the same regardless of condition. The average of the set was 65%. According to a pretest, Mechanical Turk participants’ median and average expected match with a random stranger was 50%, making our match percentage set slightly above average.

Participants also saw fake information about the targets’ home state and date of birth in order to disguise the purpose of the study. This fake demographic information was also randomly selected without replacement from a predetermined set (home state set: Wyoming, Washington D.C., Delaware, Rhode Island, North Dakota, and Vermont; birth date set: 5/30/1988,
11/28/1987, 2/29/1988, 1/1/1989, 12/24/1987, and 1/17/1989). States were chosen according to their low populations and birth dates were among the least common for Americans, thereby lowering the likelihood that targets’ demographics would match those of our participants. Birth dates were also chosen such that the ostensible ages of the targets would be similar to those of our participant sample.

*Photograph pairs.* The set and presentation of photo pairs was identical to Study 2.

*Picture ratings.* For all trials, participants rated three of the five affiliative response questions from Study 1 ($\alpha = .93$). These three items were presented in random order for each photograph pair. After the affiliative response items, participants answered one new question assessing their explicit behavioral intention toward each target (i.e., “We are considering additional studies in which people actually interact online with the person they are rating. How willing would you be to have an actual interaction with this person?”). This item was included to test whether participants' differential affiliative responses to targets with Black friends produced conscious preference to avoid interaction with them.

*Single Target Implicit Association Test.* We measured implicit anti-Black bias using a Single Target Implicit Association Test (ST-IAT; Wigboldus, Holland, van Knippenberg, 2006).

*Explicit racial bias measure.* We measured explicit racial bias using the anti-Black subscale of the Ambivalent Racism scale (Katz & Hass, 1988; $\alpha = .94$).

**Results**

**Preliminary analyses.** Six participants were removed for having an average response latency of less than 1 second on their target ratings and three more were removed for having implicit anti-Black scores that were more than 3 standard deviations from the mean. Sixty-nine
participants (experiential similarity condition: n = 36) remained in the final sample (mean age = 24.61 years, SD = 3.12; 40 male).

Main analyses

Effect of racial bias on affiliative response. The analytic strategy was identical to Study 2. In all analyses, we additionally included participants’ average rating on the six social experience questionnaire items addressing their comfort around ethnic outgroup members as a mean-centered covariate. Inclusion of this variable allowed us to control for differences between people who scored relatively low or high in comfort.

Two significant interactions emerged. First, we found a significant interaction of implicit anti-Black bias and race of the target’s friend ($\beta = .368$, $t(62) = 2.02$, $p = .047$). This was qualified by the predicted three-way interaction of implicit anti-Black bias, race of the target’s friend, and condition ($\beta = -.440$, $t(62) = -2.40$, $p = .019$). A marginal three-way interaction of explicit anti-Black bias, race of the target’s friends, and condition was also found ($\beta = -.449$, $t(62) = -1.98$, $p = .052$). All other two- and three-way interactions were not significant ($ps > .163$).

We first probed the interaction of implicit anti-Black bias and friend race by separately examining ratings of targets with Black friends and targets with White friends. With ratings of targets with Blacks as the dependent variable, the main effect of implicit anti-Black bias on liking was not significant ($\beta = -.231$, $t(62) = -1.66$, $p = .103$). However, a main effect of implicit anti-Black bias was found for ratings of targets with Whites ($\beta = .308$, $t(62) = 2.34$, $p = .022$).

Probing the predicted three-way interaction of implicit anti-Black bias, friend race, and condition, we separately analyzed participants who had been told that they have similar social experiences as the targets and those that had been told that they have similar personalities to the
targets. For participants in the experiential match condition, the interaction of implicit anti-Black bias and friend race was not significant \( \beta = -.250, t(32) = -1.54, p = .157 \). However, this interaction was significant for participants in the personality match condition \( \beta = .380, t(29) = 2.23, p = .034 \). Probing this interaction, we found the same pattern of effects as in previous studies: For targets paired with a Black friend, we found that greater implicit anti-Black bias marginally significantly predicted more negative affiliative responses \( \beta = -.229, t(28) = -1.82, p = .079 \). Inversely, greater implicit anti-Black bias significantly predicted more positive affiliative responses to targets paired with a White friend, \( \beta = .274, t(28) = 2.46, p = .021 \); see Figure 5).

Probing the predicted three-way interaction of explicit anti-Black bias, friend race, and condition, we again separately analyzed participants in each condition. However, the two-way interactions of explicit anti-Black bias and friend race were not significant in either condition (experiential similarity: \( \beta = -.312, t(32) = -1.45, p = .156 \); personality similarity: \( \beta = .246, t(32) = 1.35, p = .186 \)).

*Effect of racial bias on behavioral intention.* No significant two- or three-way interactions of racial bias, friend race, or condition were found for how willing participants were to have a future online interaction with the targets \( ps > .106 \).

**Discussion**

Study 3 provided further support for our hypothesis that perceived similarity, and in particular similarity regarding experiences with racial outgroup members, mediates the effect of implicit outgroup bias on affiliative responses toward ingroup members. We posited that directly providing participants with experiential similarity information should decrease the need for them to infer such information as a function of the race of a target’s friend, but that providing them
Figure 5. Effect of implicit anti-Black bias on affiliative response toward White targets when participants are provided personality similarity information versus experiential similarity information.
with irrelevant similarity information should not decrease the importance of friend race. In line with this prediction, when participants were told that their personality was similar to the people they were rating, implicit anti-Black bias predicted affiliative responses toward White targets in the same manner as in our previous two studies: as implicit anti-Black bias increased, affiliative responses became more negative toward Whites with a Black friend and became more positive toward Whites with a White friend. However, when participants were led to believe that they have similar experiences of intergroup contact as their targets, the race of a target’s friend did not affect participants’ affiliative responses toward them.

Secondarily, we found that implicit anti-Black bias did not significantly predict explicit behavioral intentions toward targets as a function of their friend’s race. Specifically, participants with different levels of implicit bias were equally willing to have a future online interaction with a target, regardless of their friend’s race. While this finding was not in line with our predictions, there are several potential explanations for this null effect. For instance, one issue might be that our participant sample was reluctant to interact with strangers online. Indeed, participants’ affiliative responses to targets ($M = 3.90, SD = 0.80$) were significantly greater than their willingness to interact with them ($M = 3.46, SD = 1.39$; $t(68) = 3.51, p = .001$). Another possibility is that using the average of several behavioral intention questions, instead of only one item, would yield different results. Regardless of the explanation, whether these behavioral intentions within the context of this study mirror real behavior in forming relationship ties remains an open question.
In Study 4 we examined whether the effect of implicit bias on affiliation was observable in participants' online social networks. Specifically, we tested whether participants’ implicit anti-Black bias predicts the likelihood that their friends on Facebook have Black friends. At the time of data collection, each profile displayed pictures of the friends with whom individuals interacted most frequently on Facebook (“Which friends appear in the left column of my profile?,” 2011). Given that interaction frequency over this widely used social networking site is positively correlated with self-reported closeness and trust (Kahanda & Neville, 2009; Xiang, Neville, & Rogati, 2010), Facebook provided us with a unique opportunity to measure the extent to which implicit anti-Black bias predicts individuals’ friendship choices. Furthermore, the racial composition of individuals’ friendship networks was visible not only to us, but also to other users of the site, including potential and existing friends. In addition, contact over Facebook strongly corresponds to individuals’ real-world social networks, with 48% of their real-world contacts also connected to them via this site (Bakshy, Karrer, & Adamic, 2009). As such, using Facebook allowed us to examine predictions consistent with implicit homophily using a very different and behaviorally rich paradigm, illustrating the robustness of this phenomenon.

In this study, we are, in effect, determining whether implicit bias is related to the amount of indirect outgroup contact to which a person is exposed. Having friends with outgroup friends (i.e., indirect contact) has been shown to be an effective means of prejudice reduction (Wright et al., 1997; for a recent meta-analysis see Zhou, 2013), particularly amongst individuals who do not have much direct contact (Eller, Abrams, Gomez, 2012). Thus, whether exposure to such experiences is subject to the vicissitudes of implicit bias is of both theoretical and practical
import. In fact, while we are aware of research examining moderators of the effectiveness of indirect contact for prejudice reduction (Munniksma, Stark, Verkuyten, Flasche, & Veenstra, 2013; Sharp, Voci, & Hewstone, 2011; Tausch, Hewstone, Schmid, Hughes, & Cairns, 2011), we are not aware of prior research examining moderators of exposure to relationships that afford indirect contact.

The importance of perceived similarity observed in Studies 2 & 3 also presents interesting implications for the expected effect of implicit anti-Black bias on actual friendship ties given known differences in the role similarity plays in women’s and men’s relationships. While gender did not significantly moderate effects on initial reactions to targets in the previous three lab studies, women typically demonstrate more similarity, closeness, and interpersonal sharing in their real relationships than do men (Clark & Ayers, 1992; Kuttler, La Greca, & Prinstein, 1999; McNelles & Connolly, 1999; Parker & de Vries, 1993; Sheets & Lugar, 2005; Verkuyten & Masson, 1996). Therefore, two possibilities for significant results arose: 1) As in previous studies, no differences would be found according to gender or 2) because perceived similarity appears to drive the observed effects, implicit bias might predict women’s actual relationship ties more strongly than men’s.

Method

Participants. White undergraduates (N = 143; 70 male) were recruited for pay ($8-20) or course credit.

Procedure. A White female experimenter greeted participants upon their arrival in the lab. Participants first completed a subliminal prime lexical decision task designed to measure implicit anti-Black and pro-White biases (Dovidio et al., 1997; 2002) and then completed explicit racial bias measures. After all participants’ data were collected, a research assistant
accessed publically available information on each person’s Facebook profile. During the time in which profiles were accessed for this study, each profile displayed 10 “friends” (i.e., first-order friends) in a panel on the left side of the profile, with greater numbers of interactions over Facebook increasing the likelihood of a friend being displayed (“Which friends appear in the left column of my profile?,” 2011). Each friend’s profile was then accessed and the number of Black friends he or she had (i.e., second-order friends) was recorded. If the race of the profile holder was not clearly distinguishable, then race was not coded.

**Materials**

**Implicit racial bias measure.** For this task (Dovidio et al., 1997; 2002), participants were subliminally presented with the image of a Black or White person, following which response times to positive (e.g., “good”) and negative (e.g., “bad”) words were measured. Specifically, participants were presented with subliminal face primes for ~17 milliseconds that were then masked with either a “P” inside an oval (signifying “person”) or an “H” inside a rectangle (signifying “house”). The mask was replaced with a word and participants were asked to respond “yes” or “no” as to whether it was a person word (when the preceding mask was a “P”) or a house word (when the mask was a “H”). Participants responded by pressing either “z” or “m” on the keyboard and the meaning of each key (i.e. “yes” or “no”) was counterbalanced across participants. The face primes consisted of 2” X 1.75” black and white photographs of Black and White men and women. Each of the four racial/gender categories was represented by 4 separate photos (i.e., 4 black women, 4 white men, etc.)

Lexical decision task data were analyzed according to the procedures of Dovidio and colleagues (2002). Errors and latencies greater than three standard deviations from each participant’s mean latency were excluded from analyses. These trials accounted for 2.8% of
critical trials. The remaining latencies were log transformed. Implicit bias scores were calculated using the following steps: first, the mean latency for trials in which negative words followed Black faces was subtracted from the mean latency for positive words following Black faces to create the anti-Black bias score; second, the mean latency for trials in which negative words followed White faces was subtracted from the mean latency for negative words following White faces for the pro-White score. Higher scores indicated greater implicit anti-Black bias and implicit pro-White bias, respectively.

Explicit attitude measures. The Attitudes Toward Blacks questionnaire (ATB; Brigham, 1993) and Internal Motivation to Respond Without Prejudice scale (Plant & Devine, 1998) were used as measures of explicit racial attitudes. Each was rated on 7-point Likert scale with higher numbers indicating more explicit prejudice and greater motivation, respectively.

Results

Preliminary analyses. Fifty-three participants did not have a publically accessible Facebook profile and 5 had error rates of greater than 10% on the implicit bias measure, resulting in 85 participants included in the analyses (39 male). There were no statistically significant differences between participants with viewable profiles versus those without on measures of implicit ($p = .95$) or explicit bias ($p = .83$).

Main analyses. To test whether implicit anti-Black bias predicts participants’ likelihood of indirect Black contact, we focused on two dependent variables that correspond to ways in which extended contact has been operationalized in extant research: 1) The number of first-order friends with Black friends (e.g., Eller et al., 2012; Gomez, Tropp, Fernandez, 2011; Paolini, Hewstone, Cairns, & Voci, 2004; Wright et al., 1997) and 2) the number of Black second-order
friends (Eller, Abrams, & Zimmermann, 2011). On average, participants had 1.0 first-order friend with a Black friend ($SD = 0.98$) and 1.6 Black second-order friends ($SD = 2.1$).

We entered these dependent variables separately into multiple regression analyses with mean-centered implicit anti-Black and pro-White scores, explicit racial bias (i.e., ATB), IMS, gender, and the two-way interactions between each of the four racial attitude types and gender as predictors.

First-Order Friends with Black Friends. With the number of first-order friends with Black friends as the dependent variable, we found a significant interaction of implicit anti-Black bias and gender ($\beta = -.368, t(73) = -2.45, p = .017$). There were no other significant effects ($ps > .176$). Testing women and men separately with the four racial bias types as predictors revealed that, for women, greater implicit anti-Black bias predicted fewer first-order friends with Black friends ($\beta = -.374, t(39) = -2.47, p = .018$; see Figure 6). For men, no significant effect of implicit anti-Black bias was found ($\beta = .157, t(34) = .88, p = .387$).

Black Second-Order Friends. With number of Black second-order friends as the dependent variable, we again found a significant interaction of implicit anti-Black bias and gender ($\beta = -.437, t(73) = -2.02, p = .006$). Testing women and men separately with the four racial bias types as predictors revealed that, for women, greater implicit anti-Black bias predicted fewer Black second-order friends ($\beta = -.364, t(39) = -2.38, p = .022$). For men, no significant effect of implicit anti-Black bias was found ($\beta = .265, t(34) = 1.62, p = .115$). A significant main effect of explicit bias also emerged ($\beta = -.570, t(73) = -3.21, p = .002$), such that, collapsed across gender, participants with greater explicit bias had fewer Black second-order friends. There were no other significant main effects or interactions ($ps > .103$).

---

5 All reported significant effects for this study remain so when controlling for participants’ number of Black first-order friends.
Discussion

The results of Study 4 provide further evidence of implicit homophily. Above and beyond the effects of explicit racial bias, implicit anti-Black bias of female participants predicted the composition of their Facebook friends’ friends. That is, greater implicit anti-Black bias predicted fewer friends with Blacks friends and fewer Black friends of friends within one’s Facebook community. The corresponding relationship did not hold for men. This difference is presumably because women’s friendships tend to be based more on interpersonal similarity than do men’s (Clark & Ayers, 1992; Kuttler et al., 1999; McNelles & Connolly, 1999; Parker & de Vries, 1993; Sheets & Lugar, 2005; Verkuyten & Masson, 1996). This finding suggest that women who are higher in implicit anti-Black bias are less likely to have indirect interracial contact, an experience known to remediate racial bias (Wright et al., 1997).
However, there may be other ways that implicit homophily obstructs the potential benefits of indirect contact for both men and women. In particular, future work should explore whether implicit homophily affects relational closeness, a factor which previous research has shown moderates indirect contact prejudice reduction. For example, Tausch and colleagues (2011) found that outgroup trust was improved for participants who experienced indirect contact through ingroup contacts rated as high, but not low, in closeness. Therefore, if implicit anti-Black bias predicts closeness to ingroup contacts with Black friends, implicit homophily might jeopardize the efficacy of indirect contact as a means for prejudice reduction for both genders.

As a whole, the evidence presented in Chapter 1 support the hypothesis that, when viewing cross-group friendships, implicit outgroup bias predicts ingroup affiliation via perceived similarity in comfort with the outgroup.
SECTION 2: IMPLICIT HOMOPHILY VIA NONVERBAL BEHAVIOR

The final two studies further test 1) that implicit homophily occurs as a function of inferred similarity regarding experiences of outgroup contact and 2) that social cues other than friendship ties can serve as basis for such inferences. Although outgroup friendships are one way to indicate one’s comfort with outgroup members, myriad other means exist, including support for particular political policies (Sears, 1988), willingness to live in the same neighborhoods as outgroup members (Bobo & Zubrinsky, 1996), and criticism of forms of popular culture that are strongly associated with a given outgroup (Reyna, Brandt, & Viki, 2009). As mentioned previously, one especially common and direct indication of one’s comfort level with outgroup members, and one that has consistently been implicated as a correlate of implicit bias (see Dasgupta, 2004), is nonverbal behavior. Nonverbal behavior toward outgroup members during interactions might therefore be another social cue via which implicit homophily occurs.

Investigating the role of nonverbal behaviors in implicit homophily also allowed us to test the relative importance of positive versus negative outgroup encounters. According to the implicit homophily hypothesis, observers make inferences about targets subjective experiences during outgroup interactions and that this inferred information shapes observers’ affiliative responses. Therefore, rather than devaluing all ingroup targets who interact with an outgroup person, implicit outgroup bias should differentially predict responses to people who have positive or negative outgroup interactions.

In Studies 5 and 6, we test whether implicit anti-Black bias predicts ingroup affiliation as a function of ingroup targets’ nonverbal comfort around Blacks. In Study 6, we test the assumption that participants use targets’ nonverbal behavior to make inferences about their
subjective experiences by manipulating participants’ perceptions about what information different nonverbal behaviors convey.
CHAPTER 5: NONVERBAL BEHAVIOR AS SOCIAL CUE OF SIMILARITY

Study 5 investigated whether implicit anti-Black bias predicts Whites’ affiliative responses to White targets seen exhibiting either nonverbal comfort or discomfort around Black or White strangers. Participants saw a series of short interactions between strangers and rated a White target in each. We predicted that greater implicit anti-Black bias would relate to more negative ratings of Whites who were nonverbally comfortable around Blacks.

Method

Participants. We once again recruited White U.S. participants, and 18 to 30 years old via Amazon Mechanical Turk, resulting in a sample of 170 participants (mean age = 24.90 years, SD = 3.24; 83 male).

Procedure. Participants were informed that they would watch eight silent videos of interactions between previously unacquainted people and report their impression of one of the two people in the interaction. All participants were told that they would rate the person on the right side of the screen, which in all cases was White. After each video, participants rated their affiliative response to the specified target. After rating all eight videos, they then completed a measure of implicit anti-Black bias, a measure of explicit racial bias, and basic demographic questions.

Materials

Interaction videos. Videos were converted from a past study of interracial interactions (Mallet, Wilson & Gilbert, 2008). Each participant saw a total of eight silent 20-second video clips of White targets interacting with a partner: Two nonverbally comfortable targets with a Black partner, two uncomfortable targets with a Black partner, two comfortable targets with a
White partner, and two uncomfortable targets with a White partner. Three coders (one male) rated targets’ nonverbal comfort on three dimensions (i.e., comfort, friendliness, and rapport; 1 = “very low” to 7 = “very high”; inter-rater reliability $\alpha = .91$). Average nonverbal comfort ratings for high and low comfort exemplars did not differ by partner race (comfortable: $M_{Black} = 5.3, M_{White} = 5.3$; uncomfortable: $M_{Black} = 2.8, M_{White} = 2.7$).

**Affiliative response measure.** To assess affiliative responses to the targets, participants responded to the same three items as in Study 3 ($\alpha = .91$).

**Single Target Implicit Association Test.** We measured implicit anti-Black bias using a Single Target Implicit Association Test (ST-IAT; Wigboldus, Holland, van Knippenberg, 2006).

**Explicit racial bias measure.** We measured explicit racial bias using the anti-Black subscale of the Ambivalent Racism scale (Katz & Hass, 1988; $\alpha = .88$).

**Results**

**Preliminary analyses.** Fourteen participants were removed for having an average response latency of less than 1 second on their target ratings. One hundred fifty-six participants remained in the final sample (mean age = 24.84 years, $SD = 3.17$; 80 male).

**Main analyses**

**Effect of racial bias on affiliative response.** For testing three-way interactions of racial bias, partner race, and nonverbal comfort, the two within-subjects variables, nonverbal comfort (uncomfortable vs. comfortable) and partner race (Black vs. White), were accounted for in a similar manner as in previous studies: First, two difference scores were calculated, one for targets with White interaction partners and one for targets with Black partners, by subtracting average affiliative response ratings for uncomfortable targets from those of comfortable targets. Second, this difference score for targets with Black partners was subtracted from the difference
score for targets with White partners. The resulting score was entered as the dependent variable in our model.

The three-way interaction of implicit anti-Black bias, nonverbal comfort, and partner race was significant ($\beta = -.174$, $t(153) = -2.19$, $p = .030$). However, the three-way interaction of explicit anti-Black bias, nonverbal comfort, and partner race was not significant ($p = .164$).

To probe the significant three-way interaction of implicit anti-Black bias, nonverbal comfort, and partner race, by entering nonverbal comfort difference scores for Black and White targets as dependent variables in two separate linear regression models. For each regression model, we added as a covariate the difference score for targets with partners of the race that were not included in the dependent variable. With the nonverbal difference score for targets with Black partners as the dependent variable, we found a significant interaction of implicit anti-Black bias and nonverbal comfort ($\beta = .163$, $t(152) = 2.29$, $p = .024$). With the nonverbal difference score for targets with White partners as the dependent variable, this same interaction was not significant ($p = .234$).

We probed the interaction of implicit anti-Black bias and nonverbal comfort for Blacks by separately examining ratings of targets who were nonverbally comfortable versus uncomfortable with Blacks. For each regression model, ratings of all other target types not included in the dependent variable were added as covariates. For targets exhibiting nonverbal comfort with Black partners, implicit anti-Black bias was found to significantly predict affiliative response ($\beta = -.142$, $t(149) = -2.30$, $p = .023$), such that greater implicit bias predicted more negative ratings. For targets exhibiting nonverbal discomfort with Black partners, the relationship between implicit bias and affiliative response was not significant ($p = .441$).

**Discussion**
The findings for Study 5 conceptually replicated those of Studies 1 – 4. Controlling for explicit racial bias, greater implicit anti-Black bias predicted more negative affiliative responses to targets who appeared comfortable with Black interaction partners. Interestingly, implicit bias did not predict responses to targets uncomfortable with a Black partner. This pair of effects suggests that observers’ evaluations of ingroup members are especially influenced by ingroup members’ comfortable, as opposed to uncomfortable, outgroup contact. Previous research has shown that stereotypes about Whites suggest they are biased against Blacks (Alicke & Govorun, 2005; Judd, Park, Yzerbyt, Gordijn, & Muller, 2005) and that people tend to expect and employ bias against outgroups when interacting with ingroup members (Clark & Kashima, 2007; Ruscher, Cralley, & O’Farrell, 2005). Implicit homophily might therefore occur as a function of positive outgroup contact because such contact is more unexpected and therefore more salient when inferring ingroup targets’ subjective experiences.

Unlike previous studies in which race of a target’s friend was manipulated, we found that implicit anti-Black bias also did not predict affiliative responses toward targets with White interaction partners. One possible cause for this change in effects between paradigms is that the current study provided a broader range of social scenarios with both positive and negative inter- and intra-group interactions. In doing so, the comparative nature of the first studies (i.e., Whites with friendship ties with either Blacks or Whites) may have been attenuated in the current experiment. Targets paired with a White partner may no longer have been particularly relevant to participants’ feelings about outgroup contact. Another possibility is that the comparison group changed between the paradigms. In the first studies, the alternative to targets having a Black friend was to have a White friend. In the current study, the comparison group may have been targets who were uncomfortable around Blacks. The significant difference between how implicit
bias predicted responses to targets with differing comfort levels around Blacks hints at this possibility. Regardless, identifying moderators of the relationship between implicit outgroup bias and evaluations of ingroup targets viewed with ingroup others warrants further investigation.
CHAPTER 6: INFERRING SUBJECTIVE EXPERIENCE OF SOCIAL STIMULI

Study 6 tests the crucial assumption of the implicit homophily hypothesis that observers infer targets’ subjective experiences with outgroup contact from the social cues they exhibit. Although Studies 2 & 3 provided evidence that participants make assumptions of similarity or dissimilarity based on targets’ outgroup contact, the current study provides a direct test of an alternative hypothesis: Participants with higher implicit outgroup bias simply may not prefer positive treatment of outgroup members. According to this perspective, it is the behavior itself (befriending or kindly treating outgroup members) that observers like or dislike, or feel is similar or dissimilar to their own. The implicit homophily hypothesis suggests that it is not objective behavior per se that is important, but rather the message about the targets’ subjective experiences that those behaviors imply. To test the relative importance of participants’ approval of outward treatment and their inferences of targets’ subjective experiences, in Study 6 we directly manipulated whether targets’ nonverbal behavior matched or was different than their ostensible internal states.

Method

Participants. We once again recruited White U.S. participants, and 18 to 30 years old via Amazon Mechanical Turk, resulting in a sample of 92 participants (mean age = 24.95 years, SD = 3.08; 54 male).

Procedure. The procedure was identical to Study 5 with one exception: After explaining that they would be viewing silent videos and evaluating one of the people in each, participants were given additional information to ostensibly help them with their task. In the control condition, participants were told that “research shows that in initial interactions people's behavior
usually corresponds to their inner states. For example, people who feel the most comfortable during a first-time interaction usually appear friendly. But, people who feel the most uncomfortable usually appear unfriendly.” In the opposite meaning condition, participants were instead told that “research shows that in initial interactions people's behavior usually does not correspond to their inner states. For example, people who feel the most comfortable during a first-time interaction usually appear unfriendly. But, people who feel the most uncomfortable usually appear friendly.” After this manipulation, the current study proceeded identically to Study 5. Affiliative response (α = .91) and explicit racial bias scales (α = .89) once again exhibited adequate reliability.

**Results**

**Preliminary analyses.** Five participants were removed for having an average response latency of less than 1 second on their target ratings. Eighty-seven participants remained in the final sample (mean age = 24.90 years, SD = 3.15; 51 male).

**Main analyses**

*Effect of racial bias on affiliative response.* The condition variable was accounted for in an identical manner to Studies 2 & 3. The two within-subjects variables, nonverbal comfort (uncomfortable vs. comfortable) and partner race (Black vs. White), were accounted for in an identical manner as in Study 5.

A significant three-way interaction of implicit anti-Black bias, nonverbal comfort, and partner race was found (β = -.535, t(81) = -2.29, p = .025). This effect was qualified by the predicted significant four-way interaction of condition, implicit anti-Black bias, nonverbal comfort, and partner race (β = .488, t(81) = 2.37, p = .020). All other three- and four-way interactions were not significant (ps > .262).
We probed the four-way interaction of condition, implicit anti-Black bias, nonverbal comfort, and partner race by separately analyzing ratings of targets paired with Blacks or Whites. To do so, we entered nonverbal comfort difference scores for Black and White targets as dependent variables in two separate linear regression models. For each regression model, we added as a covariate the difference score for targets with partners of the race that were not included in the dependent variable.

**Effect of implicit anti-Black bias on ratings of targets with Black partners.** With the nonverbal difference score for targets with Black partners as the dependent variable, we found a significant interaction of implicit anti-Black bias and nonverbal comfort ($\beta = .464, t(81) = 2.09, p = .040$). This interaction was qualified by the predicted significant interaction of condition, implicit anti-Black bias, and nonverbal comfort ($\beta = -.483, t(81) = -2.56, p = .012$). The two-way interaction of condition and nonverbal comfort was not significant ($p = .292$).

We probed the three-way interaction of condition, implicit anti-Black bias, and nonverbal comfort by separately examining ratings of targets who were nonverbally comfortable versus uncomfortable with Blacks. Ratings of all other target types not included in the dependent variable were added as covariates in the regression model. For targets exhibiting nonverbal comfort with Black partners, we found a significant main effect of implicit anti-Black bias on affiliative response ($\beta = -.443, t(79) = -2.51, p = .014$) such that greater bias predicted diminished affiliative response to the target. This main effect was qualified by the predicted interaction of condition and implicit anti-Black bias ($\beta = .351, t(79) = 2.28, p = .025$). The main effect of condition was not significant ($p = .126$). For targets exhibiting nonverbal discomfort with Black partners, we also found the predicted significant interaction of condition and implicit anti-Black
bias ($\beta = -.472, t(79) = -2.41, p = .018$). The main effects of condition and implicit anti-Black bias were not significant ($ps > .304$).

For ratings of targets who were comfortable around Blacks, participants in the control condition replicated the effect found in Study 5: Implicit anti-Black bias significantly predicted affiliative responses to targets who were comfortable around Blacks ($\beta = -.352, t(33) = -3.00, p = .005$; see Figure 7, panel A), such that as implicit bias increased, affiliative responses diminished. However, when participants were told that apparent nonverbal discomfort indicated greater internal comfort and apparent nonverbal comfort indicated greater internal discomfort, the effect of implicit bias on affiliative response to targets who were nonverbally comfortable around Blacks was not significant ($\beta = .025, t(42) = 0.19, p = .854$). Instead, with participants in this opposite meaning condition, we found the predicted inverted result: Greater implicit anti-Black bias predicted diminished affiliative responses to targets who were nonverbally uncomfortable around Blacks ($\beta = -.386, t(42) = -2.55, p = .014$; see Figure 7, panel B). In contrast, for participants in the control condition, implicit anti-Black bias did not significantly predict affiliative responses to targets who were uncomfortable around Blacks ($\beta = .210, t(33) = 1.42, p = .262$).

**Effect of implicit anti-Black bias on ratings of targets with White partners.** To finish probing the significant four-way interaction of racial bias, partner race, nonverbal comfort, and condition, we turned our focus to ratings of targets paired with White friends. With the nonverbal difference score for ratings of targets with White partners as the dependent variable, the interaction of implicit anti-Black bias and nonverbal comfort ($\beta = -.395, t(81) = -1.72, p = .089$) was only marginally significant. The three-way interaction of condition, implicit anti-Black bias,
Figure 7. Effect of implicit anti-Black bias on affiliative response toward nonverbally comfortable White targets in the control condition (panel A) and nonverbally uncomfortable White targets when the meaning of the nonverbal behaviors have been inverted (panel B).
and nonverbal comfort and two-way interaction of condition and nonverbal comfort were not significant (ps > .112).

We investigated the marginally significant interaction of implicit anti-Black bias and nonverbal comfort by separately examining ratings of targets who were nonverbally comfortable versus uncomfortable with Whites. Condition was dropped as a predictor due to the lack of significant higher-level interactions. Ratings of all other target types not included in the dependent variable were added as covariates. We found that the effect of implicit anti-Black bias was not significant for either targets comfortable ($\beta = -0.040$, $t(81) = -0.39$, $p = .695$) or uncomfortable ($\beta = 0.049$, $t(81) = 0.55$, $p = .582$) around Whites.

**Discussion**

We found evidence that implicit homophily occurs as a function of observers’ inferences about targets subjective experiences of outgroup contact. When participants were told that external, nonverbal comfort related to internal, subjective discomfort, the effect found in Study 5 was replicated: As implicit anti-Black bias increased, affiliative responses to White targets nonverbally comfortable with a Black interaction partner became more negative. However, when participants were told that people who appear nonverbally uncomfortable are actually relatively more subjectively comfortable than those who appear nonverbally comfortable, then the effect switched: As implicit anti-Black bias increased, responses to targets nonverbally uncomfortable with a Black partner became more negative. Participants cared less about the objective treatment of the Black interaction partner and more about the subjective experience of the target that the behavior implied.
GENERAL DISCUSSION

We found consistent support for implicit homophily across six studies, three different paradigms, and using three different measures of implicit racial bias. As predicted, Whites’ implicit racial bias was related to their affiliative responses to White targets who were comfortable around Blacks, as manipulated by friendship status (Studies 1 – 4) and nonverbal comfort (Studies 5 & 6). This relationship was not driven by explicit prejudice (Studies 1 – 6). Nor was it driven by ingroup favoritism; implicit anti-Black bias, not pro-White bias, predicted affiliation toward ingroup targets as a function of the ethnicity of their friends (Studies 1, 2, & 4). In addition, several pieces of evidence substantiate our claim that this effect occurs due to feelings of similarity with the target, rather than mere transfer of negative affect as predicted by stigma by association theory. In Study 2, perceived similarity, but not affect or expectancy violation, mediated the relationship between implicit racial bias and affiliative responses. Also in this study, implicit anti-Black bias predicted affiliative responses when targets were in the presence of an ostensible Black friend but not when in the presence of a Black stranger. Presumably, target relationship status moderated the relationship between implicit bias and affiliative responses because Black friends provide information regarding targets’ comfort around Blacks and, thus, target and perceiver similarity on this dimension. In contrast, simply being proximal to a Black stranger does not provide information with which perceivers can judge their similarity to the target. Study 3 further tested this mediator by directly manipulating it and provided clarification of the specific type of perceived similarity involved. When given direct information about how similar they were to targets regarding comfort with outgroup members, participants’ implicit anti-Black bias no longer predicted their affiliative responses. However,
when participants were instead given personality similarity information, implicit anti-Black bias again predicted affiliative response, suggesting inferences of targets’ comfort around Blacks continued to impact their evaluations. In Study 4, implicit anti-Black bias predicted the number of Black friends of friends women had on Facebook, but not those of men. We interpret this gender difference as due to differences in the basis of women’s and men’s friendships; women’s friendships tend to be based on interpersonal similarity in than men’s (Clark & Ayers, 1992; Kuttler et al., 1999; McNelles & Connolly, 1999; Parker & de Vries, 1993; Sheets & Lugar, 2005; Verkuyten & Masson, 1996). Finally, in Study 6, participants’ evaluations were influenced more by targets’ apparent subjective experience of their Blacks interaction partners than their actual nonverbal treatment of them. This finding suggests that the inferences people make about others’ outgroup comfort is what drives implicit homophily, rather than the presence or specific quality of outgroup contact.

The additional finding that implicit anti-Black bias predicts affiliative responses toward Whites who appear to have White friends also coincides with the postulated role of similarity in implicit homophily. In Studies 1 – 3, greater implicit anti-Black bias predicted increased affiliative response to a White person who appeared to be friends with a fellow White person. Similar to liking of Whites associated with Blacks, this relationship was marginally mediated by perceived similarity and moderated by the apparent relationship between the people being observed (Study 2). That is, implicit anti-Black bias predicted affiliative responses toward Whites who were friends with Whites but not those thought to be randomly paired with a White stranger. It seems that when both intraracial and interracial interactions are observed in quick succession, the choice to befriend ingroup members connotes the decision not to befriend outgroup members. Information about such relative group preference provides a basis upon
which individuals of varying degrees of implicit racial bias can ascertain their similarity with a
target and spur corresponding affiliation. A stigma transference account has a more difficult time
explaining this finding because fellow Whites are not subject to a stigma that can spread to the
target. However, given that implicit anti-Black bias did not predict affiliation toward Whites with
Whites in Studies 5 & 6, questions remain as the exact conditions that moderate this effect.

We did not take a stance on the degree to which explicit racial bias would similarly
predict homophily. However, it is interesting to note that we did not find consistent results with
explicit bias. Explicit bias predicted affiliative responses toward White targets with a Black
person only in Studies 1 & 4. Though there are reasonable explanations for the explicit bias
findings, we were unable to identify one that fully accounts for the pattern across studies. One
possibility is that implicit measures of racial bias are more valid predictors of behavior than
explicit measures (Greenwald et al., 2009), particularly when the behavior is relatively subtle,
such as spontaneously considering people’s friends when evaluating them (Toosi, Babbitt,
Ambady, & Sommers, 2012). Nonetheless, given the myriad of potential explanations, future
research specifically seeking to understand the relative importance of implicit and explicit bias in
the effect of one’s social milieu on liking of that person is warranted.

The articulation of implicit homophily has many intriguing practical and theoretical
implications. In terms of theory, it offers a potential alternate explanation of some findings
assumed to stem from affective or evaluative transfer brought about by stigma by association.
For example, perhaps evaluations of a White target who seems comfortable casually chatting
with an African American person (Pryor et al., 2012) are not driven by the automatic transfer of
negative assumptions about the Black person to the White one. Rather, White individuals high in
implicit racial bias may have assumed that the target person did not share their discomfort
around Blacks and this perceived dissimilarity accounted for their evaluations of the target. It would be interesting to identify when perceptions of similarity versus stigma spread drive evaluations of the friends and allies of members of groups more commonly examined within the stigma by association literature. Do people distance themselves from associates of heavyweight individuals due to the spread of negative feelings and assumptions or because of perceived similarity in comfort around heavyweight individuals? Maybe the relevant mechanism is governed by the nature of the relationship, with the genetic ties of family provoking the spread of negative assumptions but the choice connoted by friendship eliciting thoughts about the likelihood one would make a similar choice.

Whether implicit homophily or stigma transference effects predominate might furthermore be determined by the absolute affective valence of the stigma. As mentioned earlier, the social cognitive account of stigma transferences describes it as driven by affective stimuli, and, similarly, social and clinical psychological research on stigma by association has largely studied social stimuli that are strongly associated with disgust and danger (e.g., gays and people with mental illness or physical illness). While implicit Black bias predicts discomfort, the negative affective valence of the bias may not be great enough among most people to produce evaluative conditioning effects. This is not to say that implicit homophily and stigma transference are mutually exclusive phenomena. Not only is it theoretically plausible that both can occur simultaneously, it is also likely that they do so but with varying degrees of relative importance for different stimuli in different contexts.

Implicit homophily also provides a broader perspective on the effect of one’s social network on others’ affiliative responses and evaluations. From this perspective the individuals with whom one fraternizes does not affect evaluations only when one’s associates are members
of stigmatized groups. Rather, any interpersonal relationships that signal attitudes, beliefs, or experiences observers perceive as a meaningful basis of similarity can shape evaluations of targets. From this perspective, the effects found within this paper should generalize to non-White perceivers and non-Black targets. For example, Black’s automatic reactions to Whites should predict their liking of Black targets seen engaging in positive contact with Whites. Furthermore, this effect should hypothetically be found for any attitude object about which one has automatic bias. For example, one’s implicit bias toward a gender, political party, product brand, particular social dominance orientation, or beliefs in a just world should predict affiliative responses toward others as a function of targets’ cues that they experience these things in similar or dissimilar ways to the observer.

As a final comment on the theoretical implications of his work, the current research was greatly influenced by and is in dialog with the work of Pryor and colleagues (2012). Although both programs of research share an interest in how an ingroup target’s association with stigmatized others can lead to devaluation of the target, several distinctions can be made between the two programs: First, Pryor and colleagues were centrally interested in the avenues by which negative evaluations spread between people. In contrast, the current work focuses on how implicit responses to social stimuli may form the basis for similarity judgments and thereby affiliation. As such, the current work speaks less to specific social cognitive mechanisms by which social evaluations occur and speaks more to the fundamental processes driving relational ties and social network formation. Second, Pryor and colleagues address the contextually specific roles of deliberative versus automatic bias in producing stigma by association. In doing so, they argue that, on the one hand, when clear relationships exist between the non-stigmatized target and the stigmatized other, both explicit and implicit biases should predict target evaluations. On
the other hand, they argue that implicit bias alone should continue to predict evaluations of non-stigmatized targets with purely coincidental associations to stigmatized others. However, by manipulating relationship status, and also by demonstrating that our effects are driven by perceptions of similarity in outgroup contact experience, we provide evidence that explicit bias may not always predict evaluations of people with clearly defined relationships and that implicit bias may not always predict evaluations of people with simple associations. Instead, implicit bias relates to perceivers’ subjective experiences of outgroup contact and their inferences of similarity on this dimension to targets. Lastly, Pryor and colleagues looked at how targets’ associations with overweight and Black people affected evaluations of them and found different patterns of effects for each stigma type. However, across studies, stigma type was confounded with their deliberative/automatic bias distinction, such that the reason for the different effects could be due either to the specific nature of the stigma or their theoretical argument. To introduce an alternative explanation of their effects informed by our implicit homophily hypothesis, these different patterns of effects might be the result of different predominant mechanisms for each stigma type. As previously mentioned, if weight-based stigma is more strongly affective in nature than Black stigma, associations with overweight people might be more likely to produce stigma transference, while racial bias might be more likely to lead to implicit homophily.

Regarding practical implications of this work, implicit homophily suggests a means by which those individuals who could benefit most from indirect outgroup contact actually avoid it. Knowing that one’s friends have outgroup friends has been shown to reduce prejudice without the anxiety and fraught interactions that may characterize direct intergroup contact (Wright et al., 1997). Indirect contact has also been found to be particularly effective for those who do not have a great deal of direct intergroup contact themselves (Eller et al., 2012). These advantages are
especially important given the tendency for people high in implicit bias to avoid and have difficulty with direct outgroup contact. However, the current findings, particularly those of Study 4, suggest that individuals’ implicit racial bias negatively relates to the degree to which they are attracted to relationships that afford indirect contact.

This new problem of when and how social networks are being shaped, even at this secondary level, to decrease the likelihood of outgroup interactions and experiences deserves continued focus in the future. In light of the correlational design of Study 4, we especially look forward to future experimental research examining the precursors to selecting relationships that facilitate indirect contact. Research articulating the manner in which an ingroup contacts’ outgroup friends should be revealed when prejudice reduction is the goal would also be useful. Consideration of the present findings and the indirect contact literature in concert suggests that becoming friends with an ingroup member before revealing one’s relationships with outgroup members is a more effective means of changing attitudes than such revelation prior to the solidification of the ingroup friendship.

The notion of implicit homophily also joins a variety of findings trying to understand how and why individuals end up in social networks that are demographically and attitudinally homogenous (Bishop, 2009). With respect to interracial relations, social network diversity is constrained by a variety of forces. At a sociological level it is hampered by extensive residential segregation. For example, despite increasing demographic diversity in the United States, it is estimated that the average White person lives in a neighborhood that is approximately 75% White (Rugh & Massey, 2013). Furthermore, even when people have opportunities for outgroup contact, psychological factors differentially attract individuals to others, thereby fostering further clustering. That is, people are less likely to seek out and maintain interactions with outgroup
members (Shelton & Richeson, 2005), something that is particularly true for those who are explicitly or implicitly biased (Towles-Schwen & Fazio, 2003). The present research adds to this work by suggesting that biased individuals are also less likely to have ingroup friends known to affiliate with outgroup members, creating an additional layer of homogeneity. Finally, once individuals are connected within a social network, norms may emerge that reify and exacerbate bias. Individuals have been shown to favor those who stereotype or otherwise disadvantage outgroup members (Castelli et al., 2009; Castelli, Tomelleri, & Zogmaister, 2008), which likely encourages subsequent bias. Further empirical verification of the causal factors of attitudinally homogeneous social networks, and experimental research delineating the ramifications of such networks for attitude strength and change, as well as for outgroup network members’ well-being, are rich avenues for future inquiry.

In conclusion, our focus in this research has been on how automatic, unconscious biases and social signaling come together to determine social network composition. Our findings suggest that significant barriers exist to fighting network homogeneity and, therefore, to reducing extant levels of implicit prejudice. But our work is also hopeful. Compared to viewing attitudinal and racial homogeneity as something intractably fundamental to social networks, in highlighting the importance of perceived similarity with another for shaping these problematic outcomes, this analysis uncovers a social psychological lever for change.
REFERENCES


Byrne, D., Griffitt, W., & Stefaniak, D. (1967). Attraction and similarity of personality characteristics. *Journal of Personality and Social Psychology, 5*(1), 82.


