Emotion

Constructing Bias: Conceptualization Breaks the Link Between Implicit Bias and Fear of Black Americans

Kent M. Lee, Kristen A. Lindquist, and B. Keith Payne

CITATION

Lee, K. M., Lindquist, K. A., & Payne, B. K. (2017, July 6). Constructing Bias: Conceptualization Breaks the Link Between Implicit Bias and Fear of Black Americans. *Emotion*. Advance online publication. http://dx.doi.org/10.1037/emo0000347

Constructing Bias: Conceptualization Breaks the Link Between Implicit Bias and Fear of Black Americans

Kent M. Lee, Kristen A. Lindquist, and B. Keith Payne
University of North Carolina at Chapel Hill

Negative affect toward outgroup members has long been known to predict discriminatory behavior. However, psychological constructionist theories of emotion suggest that negative affect may not always reflect antipathy for outgroup members. Rather, the subjective experience depends on how negative affect is conceptualized as specific discrete emotions (e.g., fear vs. sympathy). Our current research integrates theories of implicit bias with psychological constructionist theories of emotion to understand the implications of negative affect toward outgroup members. Across 3 studies, we find evidence that conceptualization of negative affect toward Black Americans as sympathy, rather than fear, mitigates the relationship between negative affect and fear of Black Americans on self-report and perceptual measures, and reduces racial bias on a psychophysiological measure. These studies provide evidence that conceptualization of negative affect can shape reactions to outgroup members. We discuss the implications of these findings and ground them in theories of implicit bias, social cognition, and affective science.

Keywords: implicit attitudes, emotion, social cognition, prejudice

When Jonathan Ferrell knocked on the door of a Charlotte, NC home, he was disheveled and calling for help. It was 2:36 AM, and something was clearly wrong. The homeowner in this position might have felt sympathy for Ferrell and offered to help, or she might have felt afraid and sought help for herself. In fact, the homeowner called 911 and reported that a Black man was breaking into her house. It was a fateful decision. When the police arrived, one of the officers also interpreted Ferrell as a threat and killed the unarmed man with 10 gunshots (Leland, 2015).

The experience of specific emotions has important consequences for how we treat outgroup members, even in less dramatic situations. For example, White Americans who feel resentment toward Black Americans are more likely to oppose government policies aimed to help Black Americans such as affirmative action (Kinder & Sears, 1981; Tuch & Hughes, 2011). By contrast, White Americans who feel sympathy for the continued inequality faced by Black Americans are more likely to support policies like affirmative action (Hutchings, 2009). More broadly, it is well-known that discrete emotional experiences such as fear, disgust, anger, resentment, or sympathy play an important role in shaping intergroup behavior (e.g., Cuddy, Fiske, & Glick. 2007; Dasgupta, DeSteno, Williams, & Hunsinger, 2009; DeSteno, Dasgupta, Bartlett, & Cajdric, 2004; Ray, Mackie, Smith, & Terman, 2012; see Smith & Mackie, 2015 for a review). In one study, experiencing anger, but not sadness, increased discrimination toward outgroup members (DeSteno et al., 2004). Thus, reducing emotions associated with antisocial behavior (e.g., discrimination) and increasing emotions associated with prosocial behavior (e.g., helping) may be an important part of reducing intergroup conflict.

In this research, we investigate the hypothesis that experiences of discrete emotions toward outgroup members depend on how people subjectively interpret the meaning of their affective responses. By *affect*, we mean a general sense of positivity or negativity accompanied by some degree of arousal (Barrett & Russell, 1998; Lang et al., 1998; Larsen & Diener, 1992; Russell, 1980; Watson & Tellegen, 1985). We assume that there are individual differences in the tendency for White Americans to experience negative affect toward Black Americans because of preexisting stereotypes, prior experiences, or uncertainty surrounding outgroup members. Much previous research suggests that affective responses can be subjectively interpreted in a variety of different ways, which gives rise to different experiences of discrete emotions (e.g., Lindquist & Barrett, 2008; Oosterwijk, Topper, Rotteveel, & Fischer, 2010). We predict that how affective responses are experienced as discrete emotions will dictate whether individuals subsequently behave in an antisocial versus prosocial manner. In particular, we focus on the tendency to conceptualize negative affect toward Black Americans as fear versus sympathy. We find evidence that when individuals conceptualize their negative affect as sympathy, an emotion associated with prosocial outcomes, this prevents negative affect from being experienced as fear, an emotion associated with antisocial outcomes.

Affect, Discrete Emotions, and Discrimination

For decades, social psychology has known that negative affect is a common reaction to certain outgroup members. For instance, many social cognition studies demonstrate participants' tendency to make automatic, spontaneous, and unbidden evaluations of outgroup members along a continuum of affective valence ranging from positive to negative (e.g., Devine, Plant, Amodio, Harmon-

Kent M. Lee, Kristen A. Lindquist, and B. Keith Payne, Department of Psychology and Neuroscience, University of North Carolina at Chapel Hill.

Correspondence concerning this article should be addressed to Kent M. Lee, Department of Psychology and Neuroscience, University of North Carolina at Chapel Hill, Campus Box 3270, Chapel Hill, NC 27514. E-mail: kentl@email.unc.edu

Jones, & Vance, 2002; Fazio, Jackson, Dunton, & Williams, 1995; Wittenbrink, Judd, & Park, 1997). A common finding is that participants demonstrate negative affect to outgroup members on implicit measures of affect, even if they do not report negative evaluations on explicit measures. *Implicit measures* are defined as measures that do not rely on introspection, which in turn make it difficult for participants to control or modify their responses (e.g., in socially desirable ways; Fazio & Olson, 2003; Greenwald & Banaji, 1995). Negative affect toward Black Americans on implicit measures predicts behaving less warmly to Black individuals during social interactions (Dovidio, Kawakami, & Gaertner, 2002; Fazio et al., 1995; McConnell & Leibold, 2001), making more racially biased hiring decisions (Ziegert & Hanges, 2005), and greater readiness to perceive anger on faces of Black Americans (Hugenberg & Bodenhausen, 2003). Thus, it is often assumed that negative affect toward outgroups motivate unfavorable evaluations and antisocial behaviors (e.g., discrimination) toward members of those outgroups (e.g., Devine, 1989; Fazio, 1990, 2007; Fazio et al., 1995).

However, other evidence suggests that not all negative affect captured by implicit measures is likely to promote discriminatory behavior. Implicit measures may capture negative affect associated with some prosocial emotions (e.g., sympathy, compassion). For instance, in a series of studies, Uhlmann, Brescoll, and Paluck (2006) found evidence that negative affect on an implicit association test can reflect sympathy for outgroup members. In one study, Uhlmann and colleagues found that participants were equally likely to associate Black Americans with negative words related to prejudice (e.g., stupid, lazy, violent) as with negative words related to the history of oppression faced by Black Americans (e.g., oppressed, brutalized, mistreated). In another study, Uhlmann and colleagues found that participants made more negative associations with a fictional group when the group was described as having faced a history of oppression, than when no history of oppression was described. Thus, the presence of negative affect toward an outgroup may not in and of itself reflect emotions associated with antisocial outcomes (e.g., fear or resentment) toward that outgroup. Additionally, negative affect may not necessarily predict discriminatory behavior. Instead, negative affect may reflect prosocial emotions (e.g., sympathy) that include negative components. In particular, one possibility is that although sympathy is often perceived as a positive emotion, it is subjectively experienced as negative in the moment. Research on compassion, a related concept, supports this interpretation (Condon & Feldman Barrett, 2013). Sympathy and compassion are valued emotions because of their prosocial implications but their subjective experience may be more negative than positive.

Consistent with the idea that not all negative affect predicts discriminatory behavior, many theories of intergroup interactions hypothesize that certain discrete emotions (e.g., fear, disgust, or anger) promote discrimination, whereas others promote prosocial behaviors (e.g., guilt; e.g., Cottrell & Neuberg, 2005; Cuddy et al., 2007; Dasgupta et al., 2009; DeSteno et al., 2004; Fiske, Cuddy, Glick, & Xu, 2002; Mackie, Maimer, & Smith, 2009; Mackie, Smith, & Ray, 2008; Ray et al., 2012; Seger, Banerji, Park, Smith, & Mackie, 2017; Smith, 1993, 1999). For example, fear may be associated with a desire to avoid an outgroup (Dumont, Yzerbyt, Wigboldus, & Gordijn, 2003), whereas anger may be associated with exclusion or aggression against an outgroup (Mackie, Devos,

& Smith, 2000). In contrast, feelings of guilt about past transgressions by a majority group toward an outgroup may be associated with support for issuing formal apologies (McGarty et al., 2005) and making reparations to that outgroup (Leach, Iyer, & Pederson, 2006; Schmitt, Behner, Montada, & Muller-Fohrbrodt, 2000; Swim & Miller, 1999). Guilt regarding past aggression toward an outgroup also predicts decreased support for future acts of aggression toward that outgroup (e.g., Maitner, Mackie, & Smith, 2007). Thus, experiences of certain discrete emotions may transform general negative affect into more or less antisocial versus prosocial behavior.

At first blush, it may seem difficult to reconcile evidence that general negative affect tends to predict discriminatory behavior, but that some specific discrete emotions predict prosocial behavior. However, the tension between these seemingly disparate findings is dissolved if affective reactions are merely a building block of the discrete emotional experiences that are grounded in a situational context. For example, encounters with an outgroup member might provoke fear when taking place in the middle of the night, but not in the light of day. This hypothesis is consistent with psychological constructionist models of emotions.

The Psychological Construction of Emotion

According to psychological constructionist views on emotion (Barrett, 2006, 2009, 2013; Clore & Ortony, 2013; Cunningham, Dunfield, & Stillman, 2013; Lindquist, 2013; Russell, 2003) how a perceiver makes meaning of his or her affective state as an instance of a discrete emotion (e.g., fear, sympathy, anger, sadness) alters the experience of that affective state. In particular, affective experiences are made meaningful with regard to the specific context in which they occur (i.e., a particular situation or setting, e.g., a deserted street or a supermarket). A perceiver's interpretation of his or her affective state in a specific context ultimately shapes the emotional experiences (e.g., as a negative emotion with antisocial or prosocial implications) and subsequent behaviors (e.g., discriminatory vs. prosocial behaviors) that may arise from that affective state.

The psychological constructionist perspective stands in contrast to basic emotion views (e.g., Ekman & Cordaro, 2011; Izard, 2007; Levenson, 2011; Panksepp & Watt, 2011), which hypothesize that individuals are born with a set of universal discrete emotions that cannot be reduced to more basic building blocks. The psychological constructionist perspective also stands in contrast to causal appraisal models of emotion (e.g., Roseman, 2011), which hypothesize that discrete emotions are triggered by specific cognitive appraisals of the stimulus and situation. Causal appraisal models of emotions argue that how one makes meaning of a stimulus (e.g., an outgroup member) reflexively triggers a specific emotional reaction (e.g., fear, sympathy). In contrast, psychological constructionist views argue that emotions emerge from the context-sensitive combination of basic affective reactions, called *core affect*, and stored knowledge about specific emotion categories, called *concept knowledge* (Barrett, 2006; Clore & Ortony, 2013; Cunningham et al., 2013; Lindquist, 2013; Russell, 2003).

Core Affect

Core affect is the mental representation of the body's current state and is commonly experienced as having some degree of

positivity or negativity (i.e., affective valence) and high or low arousal (Barrett, 2006, 2009; Lindquist, 2013; Russell, 2003, 2005, 2009; Russell & Barrett, 1999). Importantly, core affect is continuously present and continuously changing in relation to real or imagined events occurring around a perceiver. These events may be external (e.g., encountering a wild animal) or internal (e.g., changes in hormone levels; imagining a traumatic experience) to the perceiver's body (Barrett, 2006, 2009; Lindquist, 2013; Mac-Cormack & Lindquist, 2017; Russell, 2003, 2005, 2009). Core affective associations may be learned via prior experience or cultural norms and become relatively automatic. For instance, a person may associate negative affect with outgroup members due to prior experiences with those groups (e.g., past negative interactions with Black Americans), cultural norms (e.g., learned negative stereotypes about Black Americans), associated concepts (e.g., associations between Black Americans and other negative concepts such as poverty), or due to neophobia (e.g., discomfort due to relatively few interactions with Black Americans).

Researchers hypothesize that core affect forms the basis of not only emotional experiences, but also of other evaluative states such as attitudes, prejudice, and decision making (Barrett & Bliss-Moreau, 2009; Bechara, Damasio, Tranel, & Damasio, 1997; Cabanac, 2002; Cunningham & Zelazo, 2007; Damasio, 1994; Wundt, 1897/1998). Critically, psychological constructionist models hypothesize that core affect itself is not sufficient to experience an emotion toward an outgroup member. Core affect is transformed into a discrete emotional experience (e.g., anger, disgust, fear, sympathy) when it is made meaningful as an instance of a discrete emotion category in a given context by drawing on emotion concept knowledge.

Concept Knowledge. *Emotion concept knowledge* is a second fundamental building block of emotional experiences. Emotion concept knowledge is a person's rich cache of the sensory and visceromotor experiences that accompany emotion concepts such as "anger," "disgust," "fear," "sympathy," and so forth. Emotion concept knowledge encodes the physiological, behavioral, phenomenological, and contextual features that describe an emotion category such as fear across the myriad contexts in which that emotion has occurred for that person in the past (e.g., fear of snakes, fear of heights, fear of an intruder, fear of outgroup members; Barrett, 2013; Lindquist, 2013; Wilson-Mendenhall, Barrett, Simmons, & Barsalou, 2011). Emotion concept knowledge is thus said to be situated, insofar as knowing what fear is involves accessing knowledge about experiences of fear across prior contexts in which fear was experienced.

The process of making meaning of core affect using concept knowledge is called *conceptualization* (Wilson-Mendenhall et al., 2011). Conceptualization is not unique to the creation of emotional experiences, but is involved in other conscious experiences as well, including visual processing. For instance, visual recognition of a focal figure (e.g., a priest) is often facilitated by backgrounds matching the focal figure (e.g., a cathedral) and hindered by backgrounds that do not match the focal figure (e.g., a football field; Davenport & Potter, 2004). Behavioral evidence suggests that individuals do in fact experience fear (i.e., experience the world as more threatening) when they make meaning of negative core affect using concept knowledge of "fear" (Lindquist & Barrett, 2008). Participants in the study first received a vignette that primed either fear or anger concept knowledge before they under-

went either a negative or neutral mood induction. Finally, participants completed a measure of risk perception as an implicit measure of fear. Participants primed with fear demonstrated the greatest risk perception, but only if they also experienced the negative affect induction, rather than the neutral induction. Thus, participants' core affective state (e.g., negative affect) and accessible concept knowledge (e.g., concept knowledge of fear) caused a particular emotional experience (e.g., an instance of fear) to emerge from a negative core affective experience.

The Psychological Construction of Discrimination-Related Emotions

There is some existing evidence that conceptualization can alter the effects of implicit affect toward social groups. In one study, researchers measured implicit affective reactions to gay couples, and then manipulated participants' conceptualization of their affective reactions (Cooley, Payne, & Phillips, 2014). One group was told that affect toward gay couples reflected their intentional evaluations of gay couples. Another group of participants was told that affect arose unintentionally. Participants who held highly negative affective reactions toward gay couples reported more explicit homophobia toward gay couples, but only when they were encouraged to believe that their affect arose intentionally. In another study, researchers manipulated whether subjects interpreted their implicit affective responses as reflecting their own attitudes or not (Cooley, Payne, Loersch, & Lei, 2015). Participants high in negative affect toward gay couples endorsed more explicit homophobia, but only if they also interpreted their affective responses as their own attitudes. These studies are consistent with our hypothesis that the meaning of negative affect toward outgroups is determined by how this affect is conceptualized in a given context. Past research has not considered how the same affective response may give rise to different interpersonal discrete emotions, a question we examine in this research.

The Present Studies

In three studies, we tested our hypothesis that conceptualization of negative affect as sympathy rather than fear can shape emotional experiences toward Black Americans. In a pilot study, we first ruled out the possibility that participants' implicitly measured negative affect reflected a single discrete emotion instead of general negative affect. In Study 1, we examined whether individual differences in the tendency to conceptualize negative affect as fear versus sympathy would predict whether participants would endorse fear toward Black Americans. In Study 2, we experimentally investigated whether conceptualization of negative affect as sympathy rather than fear would reduce self-reported fear and increase self-reported sympathy for Black Americans. Finally, in Study 3 we examined how conceptualization of negative affect as fear or sympathy influenced participants' perceptual and psychophysiological reactions to faces of Black and White individuals. Across all studies, we selected a sample that would ensure sufficient power $(1 - \beta \geq .80)$ to detect a medium-sized effect.

Pilot Study

In our pilot study, we sought to establish that participants' implicitly measured negative affect toward Black Americans re-

flected general negative affect toward Black Americans, rather than a specific negative emotion (e.g., fear). Results from a power analysis suggested that a sample size of 121 participants would provide sufficient power ($1 - \beta \geq .80$) to detect a medium-sized effect in a random effects model for multiple linear regression. We collected data from 144 online participants from Amazon Mechanical Turk. Because we were interested in racial attitudes toward Black Americans, we only included participants who self-identified as non-Black ($N = 135$) for analysis. Further, we excluded one participant who indicated the ability to read Chinese characters, which renders the stimuli in our implicit measure less ambiguous. The remaining 134 participants (49% female) ranged in age from 19 to 86 years ($M = 37.80$). Of the participants included in our analyses 82.8% identified as White, 6.0% as Hispanic or Latino, 7.5% as Asian, 3% as "other," and .7% did not specify a race or ethnicity.

We measured participants' negative affect using the affect misattribution procedure (AMP; Payne et al., 2005). The AMP is a sequential priming task in which participants see a real-life image (e.g., a Black or White face) followed by a Chinese pictograph. The prime serves to elicit an affective response that the participant may misattribute to the Chinese pictograph. Any misattributed response provides an indirect measure of participants' affective responses to the primes (see Method of Study 1 for full procedural details; see also Figure 1). Next, we asked participants to rate the extent to which they felt 10 positive and negative emotions (angry, afraid, disgusted, guilty, sympathy, enthusiastic, happy, hopeful, proud) toward Black Americans and White Americans. Emotion ratings for each group were presented separately and in random order. Participants indicated the degree to which they felt each emotion on a 1 (*never true*) to 5 (*always true*) scale. Finally, participants answered general demographic questions and a suspicion check, which asked participants to guess the purpose of the study. No participants were able to guess the hypothesis of the study.

Rather than predicting ratings of any single discrete emotion (e.g., fear), we hypothesized that implicit negative affect toward Black Americans on the AMP would be related to greater explicit ratings of multiple negative emotions and fewer explicit ratings of multiple positive emotions toward Black Americans. First, we computed an overall AMP score by subtracting the proportion of "unpleasant" responses following White primes from the proportion of "unpleasant" responses following Black primes. Thus, a higher score on the AMP reflects greater negative affect toward Black Americans relative to White Americans. The AMP demonstrated acceptable reliability (Cronbach's $\alpha = .65$). We also created composite scores for each emotion. First, we reverse scored participants' ratings for the positive emotions (enthusiastic, happy, hopeful, proud, sympathy). We reverse scored sympathy because it was negatively correlated with implicitly measured negative affect on the AMP. Next, we subtracted each emotion rating for White Americans from the emotion rating for Black Americans. Higher scores reflect more of each negative emotion and less of each positive emotion felt toward Black Americans relative to White Americans. We then averaged the scores for each emotion together to create a score reflecting negative affect.

When the emotions were averaged together, the composite was significantly associated with implicit bias ($r = .31$, $p < 001$). In a simultaneous regression, no single explicit emotion uniquely predicted negative affect toward Black Americans ($ps > .05$; see Table 1). However, the variance explained for the whole block was significant ($r = .39$, $p = .03$). This suggests that the variance shared in common across the emotions (i.e., valence) was significantly associated with implicit negative affect. Thus, the pilot study provides evidence that implicitly measured negative affect toward Black Americans is related to general negative affect toward Black Americans relative to White Americans, but that association is not driven by a single discrete negative emotion. This pattern is consistent with our assumption that implicit affect toward outgroups serves as an ambiguous signal that is available to be conceptualized as different discrete emotions based on the context. We tested whether such conceptualizations are influential in Study 1.

Study 1

According to the psychological constructionist view, how perceivers conceptualize their negative affect (e.g., as an instance of fear vs. sympathy) shapes the emotions that arise from this affect. Therefore, we predicted that participants' preexisting conceptual knowledge about their affect would predict how they make meaning of their negative affect toward Black Americans. In Study 1, we thus measured participants' preexisting tendencies to conceptualize their negative affect as specific discrete emotions and assessed whether this knowledge moderated the effect of negative affect toward Black Americans on self-reported fear and sympathy. First, we measured participants' negative affect using the AMP. Afterward, we asked participants to rate the degree to which they perceived their affect as an instance of fear and/or sympathy toward Black Americans. Finally, we had participants rate their agreement with statements reflecting explicit fear or sympathy toward Black Americans.

We predicted that individual differences in participants' tendency to conceptualize their affect as fear versus sympathy would moderate the relationship between negative affect on the AMP and self-reported fear and sympathy toward Black Americans. Specifically, we predicted that participants high in negative affect will
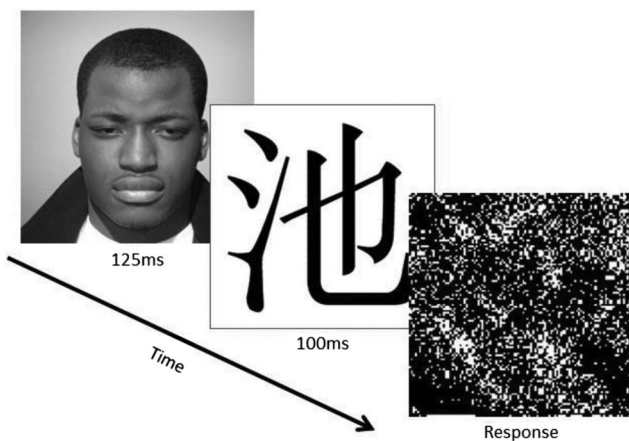

Figure 1. Diagram of a typical trial on the affect misattribution procedure (AMP). Images drawn from the AMP race database (Payne, Brown-Iannuzzi, & Panter, 2010). Reprinted with permission.

Table 1

Unstandardized Regression Coefficients Predicting Negative Affect on the AMP from Emotion Ratings in the Pilot Study

Emotion	Negative affect on AMP
Afraid	.21
Angry	−.11
Disgusted	.19
Guilty	.07
Sad	−.08
Sympathy	−.13
Enthusiastic	−.13
Happy	−.20
Hopeful	−.02
Proud	.15

Note. AMP = affect misattribution procedure.

endorse more fear, but only if they tended to conceptualize their affect as fear rather than as sympathy. We also predicted that participants high in negative affect will endorse more sympathy, but only if they tended to conceptualize their affect as sympathy rather than fear.

Method

Participants. Results from a power analysis determined that a sample size of 126 participants would provide sufficient power (1 − β ≥ .80) to detect a medium-sized two-way interaction. We collected data from 291 online participants from Amazon Mechanical Turk. We included only participants who identified as non-Black in the study (*N* = 272). Of these participants, 13 participants did not finish the study and nine participants indicated that they could speak Chinese. Further, 12 participants did not follow instructions and pressed the same key on all trials of the AMP so their responses were not valid. The remaining 238 participants ranged in age from 18 to 81 (*M* = 38.87, *SD* = 12.25), and 49.2% identified as female and 50.8% identified as male. Of these participants, 88.7% identified as White, 5.9% identified as Asian, 2.5% identified as Hispanic/Latino, .8% identified as Native American/Pacific Islander, 2.1% identified as "other."

AMP. The AMP is a sequential priming task in which participants first see a real-life image followed by a Chinese pictograph across trials. Participants are instructed to judge whether they find the Chinese pictograph pleasant by pressing one key (e.g., *P*) or unpleasant by pressing a different key (e.g., *Q*). Participants are also instructed to do their best not to be influenced by the real-life image in their judgments. The logic of the AMP is that participants will have an affective reaction to the real-life image and that this affect will sometimes be misattributed to the Chinese pictograph. The pictograph should have ambiguous value to those participants who do not read Chinese. Across trials participants' misattributed reaction to the Chinese character reflects their affect to the real-life image. We selected the AMP as our measure of negative affect because it has been used in previous studies investigating the role of interpretation on the consequences of negative affect (Cooley et al., 2014; Cooley et al., 2015). Research suggests that the AMP is capable of measuring both affective and semantic associations (see Blaison et al., 2012; Gawronski & Ye, 2014). Additionally, the

AMP has been used to assess affect in over 150 studies and has high reliability and predictive validity (Payne & Lundberg, 2014).

Participants completed 80 trials of the AMP in which they viewed a prime for 125 ms followed by a Chinese character for 100 ms, and, finally, a visual mask that remained on screen until participants responded. The primes included 20 photographs of Black individuals and 20 photographs of White individuals with neutral facial expressions. We instructed participants to judge the Chinese characters as either pleasant or unpleasant. Further, we warned the participants that the primes ("the real-life pictures") might influence their judgment of the Chinese characters and so they should do their best not to be influenced by the photographs.

Procedure. First, participants completed the AMP. Afterward, we asked participants about the degree to which they perceived their affect as indicative of fear and of sympathy to measure participants' preexisting tendencies to conceptualize their negative affect toward Black Americans as fear versus sympathy. We drew on prior research demonstrating that participants' preexisting perceptions about the meaning of their affect toward gay couples can influence explicit homophobia (Cooley et al., 2014; Cooley et al., 2015). Participants were then asked to consider the "gut feelings" they experienced while completing the AMP, and rated their perceptions of the degree to which their negative affect reflected fear ("My gut feelings toward Blacks reflect fear," "My gut feelings toward Blacks reflect anxiety") and sympathy ("My gut feelings toward Blacks reflect sympathy," "My gut feelings about Blacks reflect compassion"). There were four items in total. The fear and sympathy items were intermixed and the items were presented in random order. Participants responded using a 1 (*strongly disagree*) to 5 (*strongly agree*) scale.

Afterward, participants completed our dependent measure. This measure included three statements regarding fear of Black Americans ("Blacks are threatening," "Blacks are scary," "Blacks are frightening") and three statements regarding sympathy for Black Americans ("Blacks continue to face oppression," "Blacks are unfairly looked down upon," "Blacks are unfairly harassed"). These statements were presented in random order and the fear and sympathy statements were intermixed. Participants responded by rating how true each statement was on a 1 (*never true*) to 5 (*always true*) scale. After completing the dependent measure, participants provided demographic information and were asked to guess the hypothesis of the study before being debriefed. No participants correctly guessed the hypothesis of the study.

Results and Discussion

We predicted that conceptualization of negative affect as an instance of fear versus sympathy would moderate the relationship between negative affect and self-reported fear toward Black Americans. We first computed separate scores for self-reported fear and self-reported sympathy toward Black Americans by taking the average of the relevant items. Both the fear (Cronbach's α = .92) and sympathy (Cronbach's α = .87) items demonstrated good reliability. Additionally, the AMP demonstrated acceptable reliability (Cronbach's α = .69). We standardized all variables using *z* scores prior to analyses. We then conducted a regression analysis predicting self-reported fear of Black Americans from negative affect, emotion condition, and their interaction.

To assess whether the tendency to conceptualize negative affect as fear or sympathy predicted self-reported fear and sympathy toward Black Americans, we used the aforementioned variables in two regression analyses predicting explicit fear of Black Americans and explicit sympathy for Black Americans from negative affect, participants' conceptualization of their negative affect, and their interaction. We standardized all variables using z scores prior to analyses. When predicting explicit fear, we did not find a significant main effect of negative affect ($b = .02$), $t(235) = .33$, $p = .74$. However, as predicted, we did find that participants who conceptualized their negative affect as more indicative of fear than sympathy endorsed more fear of Black Americans ($b = .49$), $t(235) = 8.01$, $p < .001$. Importantly, the predicted interaction was significant ($b = .10$), $t(234) = 2.04$, $p = .04$ (see Figure 2). Simple slopes analysis revealed that participants who conceptualized their negative affect as more indicative of fear than sympathy endorsed more fear of Black Americans, and this effect was stronger for participants who showed high negative affect toward Black Americans ($b = .59$), $t(234) = 7.63$, $p < .001$, compared to participants who showed low negative affect ($b = .39$), $t(234) = 5.07$, $p < .001$. Thus, participants' negative affect toward Black Americans moderated the relationship between their conceptualizations about the meaning of their negative affect and self-reported fear.

Another way to examine the interaction between negative affect toward Black Americans and conceptualization of this affect as emotions is to examine conceptualization of affect as a moderator of the relationship between negative affect and self-reported fear toward Black Americans. When predicting self-reported fear from negative affect, we found a marginal trend in which participants who showed high negative affect toward Black Americans tended to endorse more fear of Black

Americans than participants who showed low negative affect, but only if they also conceptualized their affect as more indicative of fear than sympathy ($b = .12$), $t(234) = 1.70$, $p = .09$. When participants conceptualized their affect as more indicative of sympathy than fear, participants high and low negative in affect toward Black Americans did not differ in their self-reported fear ($b = -.08$), $t(234) = 1.13$, $p = .26$.

Although our primary hypothesis concerned fear, we also examined effects on explicit sympathy ratings. Participants who conceptualized their negative affect more as sympathy than as fear tended to endorse more sympathy for Black Americans ($b = -.33$), $t(235) = 5.08$, $p < .001$. We did not find a significant main effect of negative affect ($b = -.07$), $t(235) = 1.16$, $p = .25$, but the interaction between negative affect and conceptualizing affect as sympathy was significant ($b = .17$), $t(234) = 3.32$, $p = .001$. Simple slopes analyses revealed that participants who showed low negative affect toward Black Americans tended to endorse more sympathy than participants showing high negative affect, but only if they also tended to conceptualize this negative affect more as sympathy than as fear ($b = -.24$), $t(234) = 2.53$, $p = .01$. However, participants high and low in negative affect toward Black Americans did not differ in their self-reported sympathy if they perceived their negative affect more as fear than as sympathy ($b = .10$), $t(234) = 1.41$, $p = .16$.

These results provide evidence that participants' existing tendencies to conceptualize their negative affect as fear rather than as sympathy moderated the relationship between negative affect toward Black Americans and explicitly self-reported fear of Black Americans. Further, our findings provide initial support for the hypothesis that conceptualization moderates the relationship between negative affect toward an outgroup and different antisocial or prosocial negative emotions.
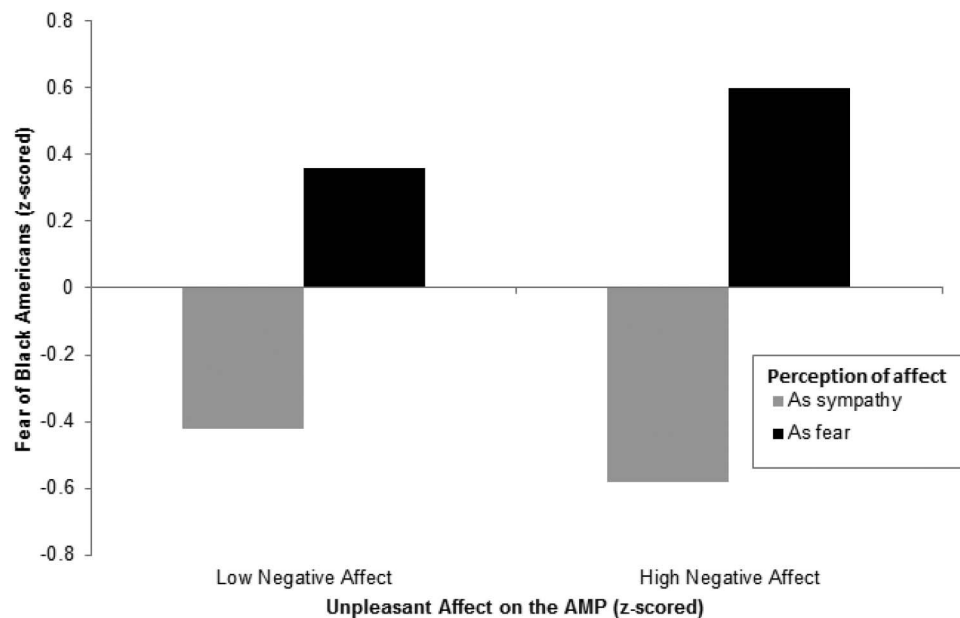
Figure 2. Perception of negative affect as fear moderates the relationship between negative affect on the affect misattribution procedure (AMP) and self-reported fear for Black Americans, Study 1.

Study 2

In Study 2, we experimentally tested our hypothesis that conceptualization of negative affect as fear versus sympathy would shape whether participants experienced fear or sympathy as a product of negative affect toward Black Americans. After we measured participants' negative affect toward Black Americans using the AMP, we manipulated participants' tendency to conceptualize their negative affect as either an instance of fear or sympathy. Participants then provided ratings of agreement with statements reflecting self-reported fear or sympathy toward Black Americans. We predicted that participants high in negative affect toward Black Americans would endorse feeling more fear and less sympathy toward Black Americans when they had conceptualized their negative affect as fear, but not sympathy in the fear condition. Conversely, we predicted that participants high in negative affect would endorse more sympathy and less fear when they had conceptualized their negative affect as sympathy, but not fear in the sympathy condition.

Method

Participants. We collected data from 202 online participants from Amazon Mechanical Turk. Our sample size was informed by the results of the power analysis reported in Study 1. However, we included only participants who self-identified as non-Black ($N =$ 189) in our analyses. Of these participants, nine did not complete the study, four indicated they could speak Chinese, one participant's responses on the AMP were not correctly recorded due to a computer error, and one participant correctly guessed the hypothesis of the study. Additionally, two participants did not follow instructions and only responded by pressing the same key on all trials of the AMP so their data were not valid. The remaining 172 participants ranged in age from 18 to 72 years ($M = 35.65$, $SD = 12.86$) and 51.2% identified as female and 48.8% as male. Of these participants, 88.4% identified as White, 4.7% as Asian, 4.1% as Hispanic/Latino, 1.2% Native American/Pacific Islander and 1.7% as "other."

Procedure. First, participants completed 40 trials of the AMP to measure negative affect toward Black Americans. Following the AMP, we manipulated participants' conceptualization of their negative affect. We drew our manipulation from previous studies demonstrating that participants' conceptualization of their negative affective states can influence their experienced emotions (Lindquist & Barrett, 2008). We also drew on more recent evidence showing that participants' inferences about the meaning of their affect toward gay couples can influence self-reported homophobia (Cooley et al., 2014; Cooley et al., 2015). In one condition, we suggested to the participants that any negative affect experienced toward Black Americans during the implicit measure reflected fear of Black Americans. In the second condition, we suggested that negative affect reflected sympathy toward Black Americans. Following the suggestion, we asked participants to generate two or three reasons why this might be the case. This manipulation relies on the confirmation bias, in which searching for evidence consistent with a hypothesis increases belief in it (Nickerson, 1998). The instructions we used are as follows:

Research has demonstrated that White Americans commonly hold negative associations with Blacks compared to Whites (i.e., Whites often are quicker to associate "Bad" with a picture of a Black face compared to a White face). Empirical evidence suggests that these negative associations are often due to feelings of fear [sympathy] because of beliefs about the status of Blacks in American society. Please generate two or three reasons why any negative feelings you may have felt toward Black faces might reflect fear of [sympathy for] Black Americans. Type the reasons you generate in the box below.

After generating two to three reasons why their negative affect toward Black Americans may reflect fear (sympathy), participants completed our dependent measures examining explicit fear and sympathy for Black Americans. These measures were the same as in Study 1. Next, participants completed the internal and external motivation to respond without prejudice scales (Plant & Devine, 1998). The internal motivation to conceal prejudice scale contains five items reflecting internal, personal reasons for responding without prejudice (e.g., "Being nonprejudiced to Black people is important to my self-concept"). The external motivation to conceal prejudice scale contains five items reflecting external, normative reasons for responding without prejudice (e.g., "I try to act non-prejudiced toward Black people because of pressure from others"). In the scale, internal and external motivation to conceal prejudice items are mixed together and participants respond with a 1 (*strongly disagree*) to 9 (*strongly agree*) scale. We included this scale to rule out the possibility that our manipulations would influence self-reported emotions because they affected motivations to control prejudice. Afterward, participants provided demographic information. Finally, we asked participants to guess the purpose of the study to gauge suspicion before debriefing them. We excluded one participant for correctly guessing the hypothesis of the study.

Results and Discussion

We predicted that conceptualization of negative affect as an instance of fear versus sympathy would moderate the relationship between negative affect and self-reported fear toward Black Americans. We first computed separate scores for self-reported fear and self-reported sympathy toward Black Americans by taking the average of the relevant items. Both the fear (Cronbach's $\alpha = .92$) and sympathy (Cronbach's $\alpha = .87$) items demonstrated good reliability. Additionally, the AMP demonstrated acceptable reliability (Cronbach's $\alpha = .69$). We standardized all variables using z scores prior to analyses. We then conducted a regression analysis predicting self-reported fear of Black Americans from negative affect, emotion condition, and their interaction.

We found that participants with more negative affect toward Black Americans endorsed more fear of Black Americans ($b = .23$), $t(169) = 3.02$, $p = .003$. However, we did not find a main effect of emotion condition ($b = .01$), $t(169) = .15$, $p = .88$. Importantly, the predicted interaction between negative affect and emotion condition was significant ($b = .17$), $t(168) = 2.20$, $p = .03$ (see Figure 3). The significant interaction demonstrated that conceptualization of negative affect moderated the relationship between negative affect and self-reported emotions toward Black Americans. Simple slopes analyses revealed that participants high in negative affect toward Black Americans reported more fear than participants low in negative affect, but only in the fear condition
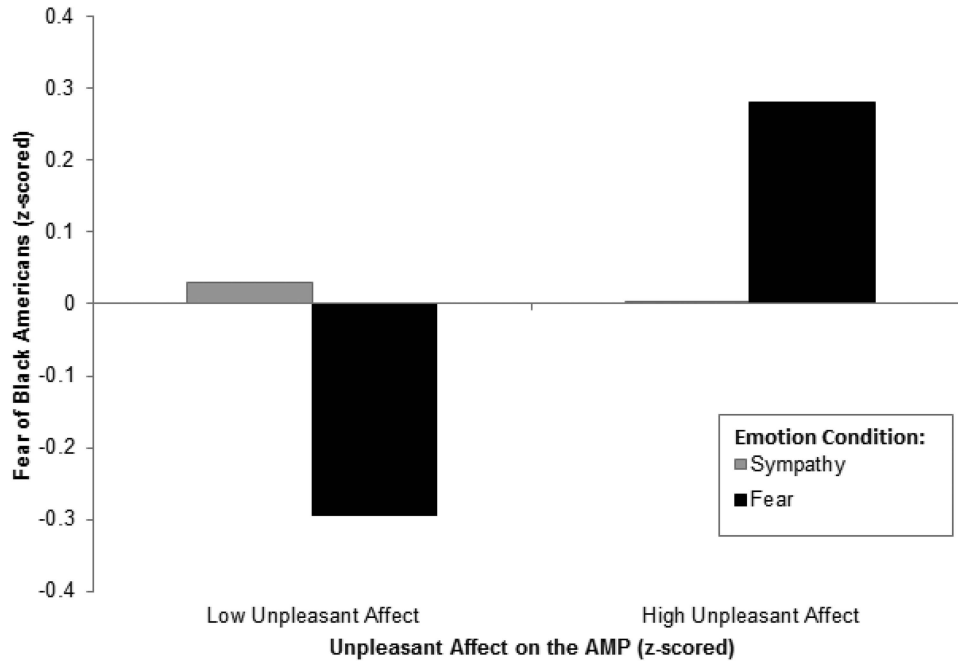
Figure 3. Emotion condition moderates the relationship between negative affect on the affect misattribution procedure (AMP) and self-reported fear of Black Americans, Study 2.

($b = .37$), $t(80) = 3.92$, $p < .001$. In the sympathy condition, participants high and low in negative affect did not differ in their self-reported fear ($b = .04$), $t(88) = .04$, $p = .72$. Thus, when participants were encouraged to conceptualize their negative affect as sympathy, negative affect was unrelated to fear of Black Americans. Conversely, when participants were encouraged to conceptualize their negative affect as fear, negative affect predicted fear of Black Americans.

To account for the possibility that our self-report findings may have been due to demand characteristics, we computed scores for internal and external motivation to conceal prejudice by taking the average of the relevant items and then examined whether our manipulation changed participants' motivations to conceal prejudice. External motivation to conceal prejudice, in particular, examines normative motivations for participants to respond without prejudice and thus can be considered an index of social desirability. Internal motivation to conceal prejudice did not significantly differ between participants in the fear condition ($M = 6.93$, $SD = 1.87$) and participants in the sympathy condition ($M = 7.11$, $SD = 1.46$), $t(153.78) = .70$, $p = .48$. Further, external motivation to conceal prejudice also did not significantly differ between participants in the fear condition ($M = 4.62$, $SD = 2.16$) and participants in the sympathy condition ($M = 4.36$, $SD = 2.20$), $t(170) = .78$, $p = .44$. These results demonstrate that emotion condition did not change participants' motivations to conceal prejudice.

To further rule out the possibility of demand characteristics, we next reran the regression analysis predicting explicit fear from negative affect, emotion condition, and their interaction while also including internal and external motivation to conceal prejudice as covariates. Including external motivation to conceal prejudice in the regression equation allowed us to account for the possibility that demand characteristics influenced participants' self-reported

fear for Black Americans. We standardized all variables using z scores prior to analyses. Even after controlling for internal and external motivations to conceal prejudice, we still find the predicted interaction between negative affect and emotion condition ($b = .16$), $t(166) = 2.24$, $p = .03$. We also find the predicted simple effects where participants high in negative affect tend to endorse more fear of Black Americans than participants low in negative affect, in the fear condition ($b = .26$), $t(78) = 2.98$, $p < .01$, but not in the sympathy condition ($b = -.02$), $t(86) = .15$, $p = .88$. These analyses, along with the correlational findings of Study 1, provide evidence against the alternative explanation that our findings for explicit fear of Black Americans were due to demand characteristics.

Next, we examined whether conceptualization would similarly moderate the relationship between negative affect and self-reported sympathy for Black Americans. As with self-reported fear, we conducted a regression analysis predicting sympathy for Black Americans from negative affect toward Black Americans, emotion condition, and their interaction. We standardized all variables using z-scores prior to analyses. We also included internal motivation and external motivation to respond without prejudice as covariates. Participants higher in negative affect toward Black Americans endorsed less sympathy for Black Americans ($b = -.27$), $t(169) = 3.71$, $p < .001$. However, we did not find a significant main effect of emotion condition ($b = -.02$), $t(169) = .34$, $p = .74$, nor did we find a significant interaction between negative affect and emotion condition ($b = -.01$), $t(168) = .18$, $p = .86$. Controlling for internal and external motivation to conceal prejudice did not change the finding for the interaction ($b = .02$), $t(166) = .32$, $p = 75$. Thus, we did not find a parallel effect in which conceptualization moderated participants' relationship

between negative affect toward Black Americans and sympathy for Black Americans.

Taken together, our findings suggest that conceptualizing negative affect as sympathy reduced fear toward Black Americans, but did not result in negative affect giving rise to sympathy for Black Americans. One interpretation of this finding is that participants found it difficult to conceptualize their negative affect as an emotion that may also be associated with positive affect. Research on compassion (Condon & Feldman Barrett, 2013), a related concept to sympathy, suggests that sympathy might be associated with both positivity and negativity, depending on the context. In particular, compassion is viewed as a positive emotion, but is often experienced as negative in the moment. It may be that conceptualizing negative affect as sympathy can prevent fear from arising from this negative affect, but not necessarily give rise to an experience of sympathy from negative affect because participants associate this emotion with positive affect. Our finding from the pilot study that negative affect was positively correlated with reports of fear, but negatively correlated with reports of sympathy toward Black Americans is consistent with this account.

Our findings provide experimental evidence that conceptualization of negative affect as fear versus sympathy can shape the emotions which arise from negative affect toward an outgroup. Our findings also suggest that negative affect toward an outgroup does not necessarily result in negative emotions associated with antisocial outcomes. Rather, the meaning of negative affect may be malleable and may depend on the emotion concept knowledge applied by the experiencer in making meaning of his or her affective state.

Study 3

One concern is that the prior studies may examine how participants label their affect, but do not necessarily examine their actual emotional experiences toward outgroup members. Study 1 may have captured participants' preexisting tendencies to label their affect as fear or sympathy. Similarly, in Study 2, our manipulation may have changed how participants labeled their negative affect in the moment, but did not necessarily influence their emotional experiences toward Black Americans. To rule out this possibility, in Study 3 we built upon the findings of Studies 1 and 2 by examining the impact of conceptualization of negative affect as instances of discrete emotions on dependent variables that did not require self-report: visual perception and psychophysiological responses.

Studies of visual perception suggest that concept knowledge shapes low-level visual perception and allows a perceiver to make meaning of sensory inputs (e.g., Aminoff, Schacter, & Bar, 2008; Bar, Aminoff, & Schacter, 2008; Davenport & Potter, 2004; Goldstone, 1995; Hansen, Olkkonen, Walter, & Gegenfurtner, 2006). In particular, studies of social perception show that social category information (e.g., whether a face is Black or White) influences how participants see others' facial expressions as neutral versus threatening. For example, when viewing an animated face shift from a neutral to angry expression, participants tend to perceive anger on the face sooner if the face is Black rather than White (Hugenberg & Bodenhausen, 2003; see also Shapiro et al., 2009). White participants also tend to categorize racially ambiguous faces as "Black" when these faces are angry, rather than neutral or happy

(Hugenberg & Bodenhausen, 2003; Hutchings & Haddock, 2008). Critically, experiences of fear magnify the overperception of anger on Black faces compared with White faces. For example, participants evaluated neutral Black faces as angry when primed with a self-protection goal, rather than a goal to search for a mate (Maner et al., 2005). We thus assessed participants' perceptual judgments of Black versus White faces as angry as a visual measure of participants' fear. Unlike explicit self-reported fear, which we used in Studies 1 and 2, this visual measure was a more implicit measure of participants' fear. We predicted that conceptualization of negative affect as an instance of sympathy, rather than fear, would moderate the links between negative affect and the perception of Black faces as aggressive, which we indexed as perceptions of anger on Black faces.

We also included skin conductance as an implicit measure of participants' affective responses. Skin conductance is an index of the arousal dimension of core affect (Lang, Greenwald, Bradley, & Hamm, 1993). Although arousal is often considered orthogonal to valence (Russell, 1980), many individuals experience arousal and valence as linked, such that increases in arousal are accompanied by increases in valence (Kuppens, Tuerlinckx, Russell, & Barrett, 2013). In line with this work, a recent study linked increased arousal with negative affect toward Black Americans (Terbeck et al., 2012). In the study, White participants demonstrated less negative affect toward Black Americans on an implicit measure after taking a drug that inhibits arousal by blunting autonomic nervous system activity.

Further, older studies have shown that participants sometimes demonstrate increased skin conductance responses to Black versus White Americans (Rankin & Campbell, 1955; Vidulich & Krevanick, 1966). In fact, more recent studies suggest that the observed differences in skin conductance responses to Black versus White individuals may reflect readiness to associate outgroup members with threat (Navarrete et al., 2009; Olsson, Ebert, Banaji, & Phelps, 2005). Although there were not differences in skin conductance responses to ingroup versus outgroup faces during classical conditioning (learning to associate ingroup vs. outgroup faces with a shock), participants later showed reduced extinction in their skin conductance response to the outgroup faces (Olsson et al., 2005). These findings suggest that participants continue to associate outgroup faces, but not ingroup faces, with threat, even when they no longer predict shocks. Inspired by this previous research linking skin conductance responses to outgroup threat, we predicted that conceptualization would moderate the relationship between negative affect and skin conductance responses to Black Americans. Specifically, we predicted participants who are high in negative affect would show increased skin conductance to Black Americans, but only if they were manipulated to conceptualize their affect as fear rather than sympathy.

Method

Participants. We recruited 140 White introductory psychology students to participate in a laboratory study in exchange for partial course credit. Our sample size was informed by the results of the power analysis reported in Study 1. Our participants were 62.9% female and 37.1% male. Their ages ranged from 18 to 45 ($M = 19.30$, $SD = 2.80$).

Skin conductance. We measured skin conductance responses from the distal phalanx of the third and fourth fingers of the participants' nondominant hands. Skin conductance was sampled at 500 Hz using 1 5/8 in diameter Ag/AgCl disposable electrodes (Mindware Technologies LTD., Gahanna, OH). Skin conductance was recorded using Mindware's BioNex 8-Slot Chassis and processed and analyzed using Mindware's BioLab (Version 3.0.13). Before analysis, skin conductance data were cleaned for artifacts. We calculated skin conductance responses by comparing the peak skin conductance level to the trough prior to the onset of the response. As is common, we considered valid only those responses that surpassed $> .02$ μSiemens (μS) and began 1 s to 3 s after stimulus onset (Dawson, Schell, & Filion, 2007; Prokasy & Kumpfer, 1973; Stern, Ray, & Quigley, 2000).

Procedure. First, skin conductance sensors were attached to the participants. Next, participants completed the AMP. In the AMP, we used 60 stimuli (30 Black and 30 White individuals with neutral expressions). Afterward, we administered our emotion manipulation, which was the same as in the Study 2. Following the manipulation, participants completed the two dependent measures. First, participants saw the same real-life images of Black and White individuals presented in the AMP while their skin conductance responses were measured. Each image was presented for 5 s with a random intertrial interval ranging from 10 s to 14 s for a total of 60 trials.

Second, participants completed our visual measure of fear, which was an anger judgment task involving posed angry expressions on Black and White faces. Although we were primarily interested in how participants perceived the faces of Black individuals, we included faces of White individuals as a comparison group. We took a subset of five Black and five White faces from the real-life images we used in the AMP and morphed them to varying levels of anger using the software FaceGen (Singular Inversions Inc, Toronto, Ontario, Canada). The software package we used allowed us to morph the expression on the faces to depict varying degrees of signal strength for the angry expression (see Figure 4). Using this software, we created morphs that ranged in signal strength of expressed anger from 0% to 50% anger, varying by increments of 5%. In total, there were 110 faces (5 Black and 5 White individuals with 11 morphs each), which were randomly presented such that the face of the same individual was not presented twice in a row. Participants judged the level of anger on each face using a 1 (*not at all angry*) to 9 (*extremely angry*) scale. Afterward, participants answered general demographic questions.
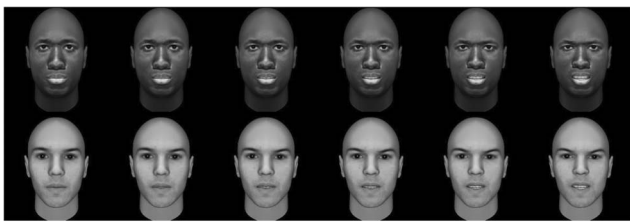

Figure 4. Example of morphed Black and White faces ranging from 0% to 50% signal strength in Study 3. Images adapted from the affect misattribution procedure (AMP) race database (Payne et al., 2010). Published with permission.

Results and Discussion

Anger perception. First, we examined whether conceptualization would moderate the relationship between negative affect and perception of anger on Black faces. We standardized all variables using z scores prior to analyses. We then computed separate scores for participants' judgments to morphed Black and morphed White faces by taking the average of participants' ratings of anger on the morphed Black and morphed White faces, respectively. Higher ratings reflect greater perceived anger.

We conducted a regression analysis predicting judgments of anger of Black faces from negative affect toward Black Americans, emotion condition, and their interaction. We found that participants high in negative affect rated the morphed Black faces as angrier compared with participants low in negative affect ($b = .29$), $t(137) = 3.51$, $p = .001$. However, we did not find a main effect of emotion condition ($b = -.02$), $t(137) = -.24$, $p = .84$. Importantly, the predicted interaction was significant ($b = .18$), $t(136) = 2.22$, $p = .03$ (see Figure 5). Simple slopes analyses revealed that in the fear condition, participants high in negative affect toward Black Americans perceived more anger on the Black faces than participants low in negative affect ($b = .46$), $t(69) = 4.28$, $p < .001$. However, in the sympathy condition, participants high and low in negative affect did not differ in their ratings of perceived anger ($b = .10$), $t(68) = .84$, $p = .40$. These results demonstrate that conceptualization of negative affect as an instance of sympathy, rather than fear reduced the relationship between negative affect and perceived of anger on the morphed Black faces.

We did not predict, but found, that participants high in negative affect toward Black Americans also tended to rate the morphed White faces as angrier than participants low in negative affect ($b = .18$), $t(137) = 2.19$, $p = .03$, demonstrating generalization of negative affect to the perception of all faces as angry. These findings might reveal individual differences in reactivity to social threat in general. However, we did not find a main effect of emotion condition ($b = -.06$), $t(137) = -.74$, $p = .46$, nor a significant interaction of negative affect and emotion condition ($b = .14$), $t(136) = 1.69$, $p = .09$. Thus, conceptualization moderated the relationship between negative affect toward Black Americans and perceptions of anger on the faces of Black Americans, and to a lesser extent, White Americans.

Skin conductance responses. We were also interested in whether skin conductance in response to Black faces would depend on how participants conceptualized their affective reactions. We computed a difference score by subtracting the average skin conductance response magnitude to White faces from the average skin conductance response magnitude to Black faces. The mean skin conductance magnitudes for each emotion condition are shown in Figure 6.

On average, participants showed larger racial bias in skin conductance responses in the fear condition compared to the sympathy condition ($b = .18$), $t(113) = 1.96$, $p = .05$. However, this effect was not moderated by participants' AMP scores ($p > .05$), suggesting that the fear condition was sufficient to increase arousal in response to black faces, both for participants low and high in negative affect. Our results provide partial support of our hypothesis. Participants demonstrated less autonomic nervous system activity when encouraged to conceptualize their negative affect as
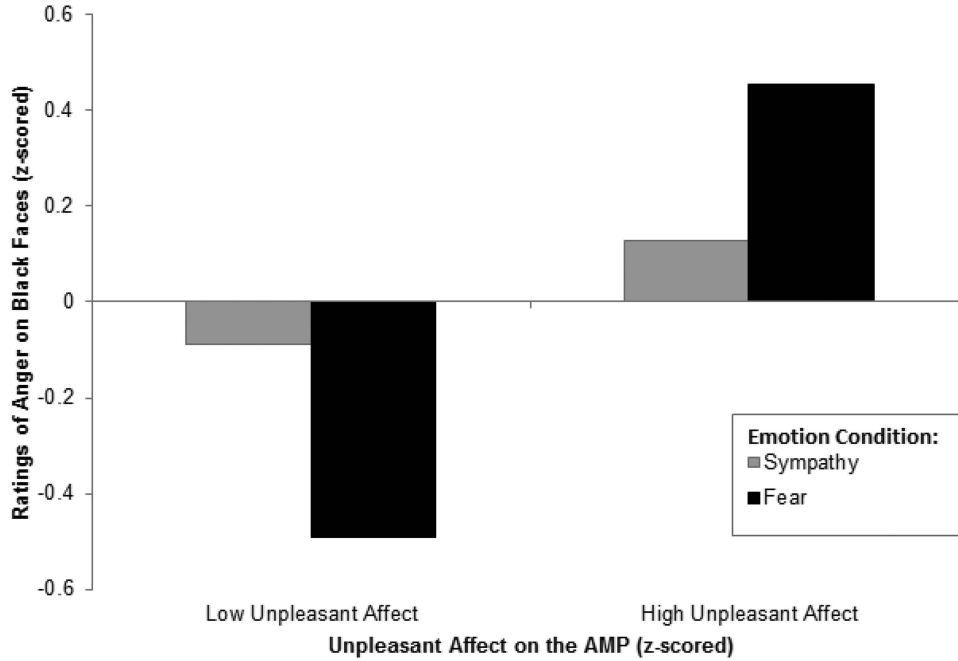
Figure 5. Emotion condition moderates the relationship between negative affect on the affect misattribution procedure (AMP) and perception of anger on the faces of Black Americans, Study 3.

sympathy, rather than fear. These findings suggest that conceptualization changed participants' emotional reactions to Black Americans, regardless of whether the preexisting reactions were relatively weak or strong.

The finding that implicit negative affect did not predict skin conductance is surprising, as a previous study has linked physiological arousal to negative affect toward Black Americans (Terbeck et al., 2012). However, other studies have found autonomic nervous system activity and affective valence to be orthogonal (e.g., Lang et al., 1993). Thus, one reason we did not find that individual differences in affect moderated the effect of conceptualization on autonomic nervous system activity may have been
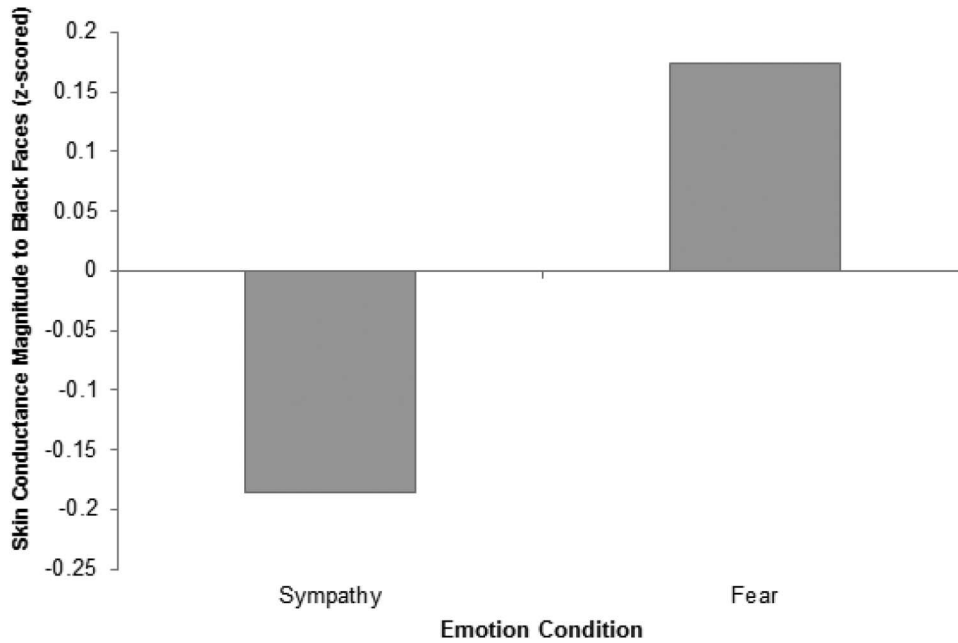

Figure 6. Sympathy reduces skin conductance magnitude to Black faces, Study 3.

because no relationship was present between our implicit measure of affect and our measure of arousal.

Summary. Study 3 provides additional evidence that conceptualization of negative affect as an instance of a specific discrete emotion shapes subsequent emotions toward outgroup members. Our results in Study 3 revealed that conceptualization moderated the relationship between negative affect and perception of anger on Black faces. Further, we find evidence that conceptualization influenced participants' autonomic nervous system activity. Importantly, our findings in Study 3 provide evidence that conceptualization changed participants' emotional experiences. Thus, Study 3 also rules out the alternative explanation that our manipulation solely changed how participants labeled their negative affect toward Black Americans in Study 2. Overall, our results provide partial support for the hypothesis that conceptualization of negative affect toward Black Americans can shape emotions toward outgroup members. Specifically, conceptualization determines whether White participants perceive and respond to Black Americans as threats.

General Discussion

Across three studies, we provide evidence that emotion concept knowledge and conceptualization can shape how negative affect toward Black Americans is experienced. In our pilot study, we demonstrated that our implicit measure captured general negative affect rather than any specific negative emotion associated with Black Americans. In Study 1, we provide evidence that individual differences in participants' tendencies to conceptualize their negative affect as fear versus sympathy predict whether antisocial emotions are likely to arise from their affect. In Study 2, we manipulated whether participants were encouraged to conceptualize their negative affect toward Black Americans as fear versus sympathy and found that conceptualization of negative affect determined whether participants experienced fear toward Black Americans. In Study 3 we extended the findings of Studies 1 and 2 by demonstrating that conceptualization of negative affect moderated the relationship between negative affect and perceptions of anger on Black faces. In Study 3, we also demonstrated that that manipulating participants' tendency to conceptualize their negative affect as sympathy decreased psychophysiological arousal in response to faces of Black Americans.

Our findings are unlikely to be explained by social desirability effects. Although such motivations can easily affect explicit expressions of prejudice, we found no evidence that they account for the effects of our conceptualization variables on reports of fear. In Study 2 our manipulation of conceptualizations as fear versus sympathy could potentially influence social desirability. However, we found no effects on internal or external motivations to control prejudice. In Study 1, we measured naturally occurring variability in perceptions of affect as fear versus sympathy. This correlational design is less likely than our conceptualization manipulation to create demand effects or social desirability bias, and we found the same pattern of results. Finally, Study 3 used a perceptual measure of fear, and skin conductance, which are less likely than self-reports to be influenced by socially desirable responding. Across multiple methods we found that conceptualization broke the link between negative implicit affect and fear of Black Americans.

Consistent with a psychological constructionist framework of emotion, the findings of Studies 1 through 3 suggest that concept knowledge about discrete emotions shape whether more or less prosocial discrete negative emotions emerge from negative affect. In particular, whether emotion concept knowledge about sympathy or fear was salient determined whether negative affect toward Black Americans gave rise to the perception of Black Americans as threats (e.g., fear). In accord with previous studies of emotion (e.g., Lindquist & Barrett, 2008; Oosterwijk et al., 2010), our findings suggest an important role of concept knowledge in determining emotions, in this case, emotions that arise from negative affect toward outgroup members.

Furthermore, our findings reconcile evidence that negative affect often predicts discriminatory behavior (e.g., Devine et al., 2002; Fazio et al., 1995; Wittenbrink et al., 1997) with evidence that some discrete negative emotions predict prosocial behavior (e.g., Leach et al., 2006; McGarty et al., 2005; Schmitt et al., 2000; Swim & Miller, 1999). Implicit measures may sometimes capture negative affect that a perceiver conceptualizes as a negative emotion associated with antisocial outcomes (e.g., anger, fear, disgust, contempt). This negative affect would thus predict discriminatory behavior. Conversely, a perceiver may conceptualize negative affect as a prosocial negative emotion (e.g., sympathy, guilt, shame, compassion), which explains how certain negative emotions tend to predict prosocial behavior. Thus, the emotion concept knowledge that individuals apply to their negative affect may determine whether negative affect results in discriminatory behavior.

Process-Based Accounts of Implicit Bias

Our findings have important implications for process-based accounts of implicit bias. Early theories of implicit bias assumed that implicit attitudes reflected relatively fixed associative representations in memory (Fazio, 1990; Wilson, Lindsey, & Schooler, 2000). In contrast, more recent theories have focused on dynamic processes. For example, the associative propositional evaluation model (Gawronski & Bodenhausen, 2006, 2011) assumes that associative processes determine which concepts are made accessible in a given context. That activated content serves as input to propositional reasoning that assesses whether the content (e.g., "Black people are bad") is true or false. Intentional behavior is based on accessible information, only if it is judged to be true. The present findings are consistent with dynamic processing accounts rather than accounts of attitudes as fixed representations. However, our findings suggest that propositional reasoning evaluates content beyond a simple true/false dichotomy. Propositional reasoning may be used to evaluate the meaning or significance of content activated by associational processes. For example, propositional reasoning may be used to make sense of negative affect as fear after conceptual knowledge about fear becomes more accessible in a given context. Thus, our findings suggest that dynamic process models should account for the role of propositional reasoning beyond a simple true/false dichotomy in order to capture how individuals might construct complex responses from simpler processes.

Given that our data are consistent with dynamic process models, one fruitful future direction may be an integration of different dynamic process models in psychology. For example, the situated

inference model of priming is a dynamic model that assumes people construct actions from simpler cognitions (Loersch & Payne, 2011). The model argues that priming affects behavior when primes activate information that is then interpreted in light of situated affordances. For example, priming stereotypes (e.g., "Blacks are dangerous") is likely to influence behavior (e.g., avoidance) provided that the stereotype is applicable to the current situation (e.g., encountering a Black American in a dark alley). Making a particular subset of conceptual knowledge more accessible (e.g., criminality and Black Americans), increases the likelihood that this conceptual knowledge is applied in subsequent situations (e.g., encounters with Black Americans). Interestingly, these dynamic accounts of implicit bias have much in common with psychological constructionist models of emotion. For example, perceivers may apply primed conceptual knowledge (e.g., fear) to form a conceptualization about the meaning of negative affect (Barrett, Wilson-Mendenhall, & Barsalou, 2015) toward outgroup members in a relevant context.

By viewing complex psychological states as the situated combination of simpler psychological processes, both a situated inference model (Loersch & Payne, 2011) and psychological constructionist approaches (e.g., Barrett, 2006, 2009; Lindquist, 2013) can explain important context-based effects on implicit measures. For example, participants demonstrated less negative affect toward Black individuals on an implicit measure when they were presented in front of a church interior, rather than a graffiti-covered street corner (Wittenbrink, Judd, & Park, 2001). In another study, participants demonstrated less negative affect toward Black individuals if the Black individuals wore suits rather than prison uniforms (Barden, Maddux, Petty, & Brewer, 2004). These studies suggest context may determine the presence or absence of negative affect toward Black Americans, or in other contexts, how that affect is made meaningful. For example, background information, such as a church, may act as a contextual cue that makes an emotion concept of sympathy salient for perceivers and in turn, responses consistent with sympathy on an implicit measure. Thus, integrating models of priming and models of emotion within a psychological framework may explain how context can influence the responses of perceivers make on implicit measures. In particular, context may determine the conceptualization a perceiver generates from core affect and concept knowledge.

Practical Implications

Our findings suggest a novel way to reduce interpersonal bias. Much research has attempted to modify automatic associations in an effort to reduce discrimination that may follow from implicit biases. Although there have been some limited successes (Devine, Forscher, Austin, & Cox, 2012; Hu, Rosenfeld, & Bodenhausen, 2012; Kawakami, Dovidio, Moll, Hermsen, & Russin, 2000; Plant et al., 2009), other studies have found that attempts to modify implicit biases may not have long-lasting impact (Gawronski & Strack, 2004; Gregg, Seibt, & Banaji, 2006; Lai et al., 2016; for a discussion see Gawronski & Bodenhausen, 2006). Thus, approaches that successfully reduce implicit bias by modifying implicit associations tend to require extensive time and effort training participants to form new affective associations (e.g., Devine et al., 2012; Kawakami et al., 2000) that generalize across contexts.

However, our findings suggest that mitigating the harm caused by negative affect does not require changing well-learned, automatically activated affective associations. Instead, interventions that change the way people conceptualize their affective reactions as discrete emotions may disrupt the harmful effects of affective reactions toward outgroups without the need for long-lasting changes in affective associations. Encouraging individuals to conceptualize their negative affect toward outgroups as a prosocial discrete negative emotion may be one way to reduce discriminatory behavior toward outgroups. Because this approach does not require changing affective associations per se, it offers a tractable approach for reducing harm caused by negative affect toward outgroups in day-to-day life. In the long term, reducing discriminatory behavior may effectively reduce affective associations by changing participants' self-perceptions or perceptions of societal norms.

Limitations and Future Directions

The results of the present studies should be interpreted within the context of their limitations. First, though our studies found that conceptualizing negative affect as a prosocial emotion reduced explicit prejudice, we did not find consistent evidence that our manipulation increased prosocial emotions toward Black Americans. As discussed earlier, one possibility is that although sympathy may be experienced as negative, it is perceived to be a positive emotion. Thus, simply activating participants' emotion concept of sympathy may not be enough for them to construct the experience of sympathy from negative affect.

Further, practically speaking, discouraging discriminatory behaviors is not the same as encouraging prosocial behaviors. For example, simply because a person does not experience antisocial negative emotions toward an outgroup member, does not mean that he or she will step in to offer aid to an outgroup member facing discrimination. Thus, future studies should further investigate ways concept knowledge and situated conceptualization shape prosocial or antisocial behaviors toward outgroup members. For instance, future research might investigate whether causing participants to conceptualize negative affect as guilt can increase participants' tendency engage in prosocial behaviors toward outgroup members. However, feelings of self-conscious negative emotions such as guilt and shame can also prompt behavioral inhibition (Amodio, Devine, & Harmon-Jones, 2007; Beer, Heerey, Keltner, Scabini, & Knight, 2003; Fourie, Kilchenmann, Malcolm-Smith, & Thomas, 2012; Muris, Meesters, Bouwman, & Notermans, 2015; Tangney, Miller, Flicker, & Barlow, 1996). Behavioral inhibition may be one reason that inducing self-conscious negative emotions can make individuals appear colder in interactions with outgroups (Vorauer & Turpie, 2004). Thus, though it is possible to mitigate the link between negative affect and antisocial behavior, inducing prosocial behavior toward outgroups members may be easier said than done.

Additionally, the findings provide evidence that emotions toward outgroup members may be constructed from more basic processes (core affect and concept knowledge). However, an important question that remains is whether this generalizes beyond the construction of emotion. Although the findings of Study 3 and studies of visual processing suggest that concept knowledge and situated conceptualization play an important role in many basic

psychological processes (i.e., visual perception), future studies should investigate how concept knowledge may shape implicit evaluations and implicit biases across multiple domains. It is hypothesized that affect forms a "common currency" (Barrett & Bliss-Moreau, 2009; Cabanac, 2002; Craig, 2009) for myriad evaluative states including emotions as well as attitudes, valuation, and decision-making. Thus, it stands to reason that conceptualization of core affect can impact multiple social and affective phenomena. Such studies will be important in elucidating the mechanism by which automatic evaluations can shape behavior across multiple domains.

Conclusion

Encounters with outgroup members sometimes provoke negative affective reactions. Our research suggests that this negative affect need not result in emotions associated with antisocial behaviors. Rather, the concept knowledge that perceivers apply in making meaning of their affective state may be an important determinant of responses toward outgroup members. Understanding implicit bias through a psychological constructionist approach can elucidate when negative affect results in discriminatory behaviors toward outgroups.

References

Aminoff, E., Schacter, D. L., & Bar, M. (2008). The cortical underpinnings of context-based memory distortion. *Journal of Cognitive Neuroscience, 20,* 2226–2237. http://dx.doi.org/10.1162/jocn.2008.20156

Amodio, D. M., Devine, P. G., & Harmon-Jones, E. (2007). A dynamic model of guilt: Implications for motivation and self-regulation in the context of prejudice. *Psychological Science, 18,* 524–530. http://dx.doi.org/10.1111/j.1467-9280.2007.01933.x

Bar, M., Aminoff, E., & Schacter, D. L. (2008). Scenes unseen: The parahippocampal cortex intrinsically subserves contextual associations, not scenes or places per se. *The Journal of Neuroscience, 28,* 8539–8544. http://dx.doi.org/10.1523/JNEUROSCI.0987-08.2008

Barden, J., Maddux, W. W., Petty, R. E., & Brewer, M. B. (2004). Contextual moderation of racial bias: The impact of social roles on controlled and automatically activated attitudes. *Journal of Personality and Social Psychology, 87,* 5–22. http://dx.doi.org/10.1037/0022-3514.87.1.5

Barrett, L. F. (2006). Solving the emotion paradox: Categorization and the experience of emotion. *Personality and Social Psychology Review, 10,* 20–46. http://dx.doi.org/10.1207/s15327957pspr1001_2

Barrett, L. F. (2009). Variety is the spice of life: A psychological construction approach to understanding variability in emotion. *Cognition and Emotion, 23,* 1284–1306. http://dx.doi.org/10.1080/02699930902985894

Barrett, L. F. (2013). Psychological construction: The darwinian approach to the science of emotion. *Emotion Review, 5,* 379–389. http://dx.doi.org/10.1177/1754073913489753

Barrett, L. F., & Bliss-Moreau, E. (2009). Affect as a psychological primitive. *Advances in Experimental Social Psychology, 41,* 167–218. http://dx.doi.org/10.1016/S0065-2601(08)00404-8

Barrett, L. F., & Russell, J. A. (1998). Independence and bipolarity in the structure of current affect. *Journal of Personality and Social Psychology, 74,* 967–984. http://dx.doi.org/10.1037/0022-3514.74.4.967

Barrett, L. F., Wilson-Mendenhall, C. D., & Barsalou, L. W. (2015). The conceptual act theory: A road map. In L. F. Barrett & J. A. Russell (Eds.), *The psychological construction of emotion* (pp. 83–110). New York, NY: Guilford Press.

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science, 275,* 1293–1295. http://dx.doi.org/10.1126/science.275.5304.1293

Beer, J. S., Heerey, E. A., Keltner, D., Scabini, D., & Knight, R. T. (2003). The regulatory function of self-conscious emotion: Insights from patients with orbitofrontal damage. *Journal of Personality and Social Psychology, 85,* 594–604. http://dx.doi.org/10.1037/0022-3514.85.4.594

Blaison, C., Imhoff, R., Hühnel, I., Hess, U., & Banse, R. (2012). The affect misattribution procedure: Hot or not? *Emotion, 12,* 403–412. http://dx.doi.org/10.1037/a0026907

Cabanac, M. (2002). What is emotion? *Behavioural Processes, 60,* 69–83. http://dx.doi.org/10.1016/S0376-6357(02)00078-5

Clore, G. L., & Ortony, A. (2013). Psychological construction in the OCC model of emotion. *Emotion Review, 5,* 335–343. http://dx.doi.org/10.1177/1754073913489751

Condon, P., & Feldman Barrett, L. (2013). Conceptualizing and experiencing compassion. *Emotion, 13,* 817–821. http://dx.doi.org/10.1037/a0033747

Cooley, E., Payne, B. K., Loersch, C., & Lei, R. (2015). Who owns implicit attitudes? Testing a metacognitive perspective. *Personality and Social Psychology Bulletin, 41,* 103–115. http://dx.doi.org/10.1177/0146167214559712

Cooley, E., Payne, B. K., & Phillips, K. J. (2014). Implicit bias and the illusion of conscious ill will. *Social Psychological & Personality Science, 5,* 500–507. http://dx.doi.org/10.1177/1948550613506123

Cottrell, C. A., & Neuberg, S. L. (2005). Different emotional reactions to different groups: A sociofunctional threat-based approach to "prejudice". *Journal of Personality and Social Psychology, 88,* 770–789. http://dx.doi.org/10.1037/0022-3514.88.5.770

Craig, A. D. (2009). How do you feel—Now? The anterior insula and human awareness. *Nature Reviews Neuroscience, 10,* 59–70. http://dx.doi.org/10.1038/nrn2555

Cuddy, A. J., Fiske, S. T., & Glick, P. (2007). The BIAS map: Behaviors from intergroup affect and stereotypes. *Journal of Personality and Social Psychology, 92,* 631–648. http://dx.doi.org/10.1037/0022-3514.92.4.631

Cunningham, W. A., Dunfield, K. A., & Stillman, P. E. (2013). Emotional states from affective dynamics. *Emotion Review, 5,* 344–355. http://dx.doi.org/10.1177/1754073913489749

Cunningham, W. A., & Zelazo, P. D. (2007). Attitudes and evaluations: A social cognitive neuroscience perspective. *Trends in Cognitive Sciences, 11,* 97–104. http://dx.doi.org/10.1016/j.tics.2006.12.005

Damasio, A. R. (1994). *Descartes' error: Emotion, reason, and the human brain.* New York, NY: Quill.

Dasgupta, N., Desteno, D., Williams, L. A., & Hunsinger, M. (2009). Fanning the flames of prejudice: The influence of specific incidental emotions on implicit prejudice. *Emotion, 9,* 585–591. http://dx.doi.org/10.1037/a0015961

Davenport, J. L., & Potter, M. C. (2004). Scene consistency in object and background perception. *Psychological Science, 15,* 559–564. http://dx.doi.org/10.1111/j.0956-7976.2004.00719.x

Dawson, M. E., Schell, A. M., & Filion, D. L. (2007). The electrodermal system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (3rd ed., pp. 159–181). New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511546396.007

DeSteno, D., Dasgupta, N., Bartlett, M. Y., & Cajdric, A. (2004). Prejudice from thin air: The effect of emotion on automatic intergroup attitudes. *Psychological Science, 15,* 319–324. http://dx.doi.org/10.1111/j.0956-7976.2004.00676.x

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56,* 5–18. http://dx.doi.org/10.1037/0022-3514.56.1.5

Devine, P. G., Forscher, P. S., Austin, A. J., & Cox, W. T. (2012). Long-term reduction in implicit race bias: A prejudice habit-breaking intervention. *Journal of Experimental Social Psychology, 48,* 1267–1278. http://dx.doi.org/10.1016/j.jesp.2012.06.003

Devine, P. G., Plant, E. A., Amodio, D. M., Harmon-Jones, E., & Vance, S. L. (2002). The regulation of explicit and implicit race bias: The role of motivations to respond without prejudice. *Journal of Personality and Social Psychology, 82,* 835–848. http://dx.doi.org/10.1037/0022-3514.82.5.835

Dovidio, J. F., Kawakami, K., & Gaertner, S. L. (2002). Implicit and explicit prejudice and interracial interaction. *Journal of Personality and Social Psychology, 82,* 62–68. http://dx.doi.org/10.1037/0022-3514.82.1.62

Dumont, M., Yzerbyt, V., Wigboldus, D., & Gordijn, E. H. (2003). Social categorization and fear reactions to the September 11th terrorist attacks. *Personality and Social Psychology Bulletin, 29,* 1509–1520. http://dx.doi.org/10.1177/0146167203256923

Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion Review, 3,* 364–370. http://dx.doi.org/10.1177/1754073911410740

Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. *Advances in Experimental Social Psychology, 23,* 75–109. http://dx.doi.org/10.1016/S0065-2601(08)60318-4

Fazio, R. H. (2007). Attitudes as object-evaluation associations of varying strength. *Social Cognition, 25,* 603–637. http://dx.doi.org/10.1521/soco.2007.25.5.603

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

Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition. research: Their meaning and use. *Annual Review of Psychology, 54,* 297–327. http://dx.doi.org/10.1146/annurev.psych.54.101601.145225

Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology, 82,* 878–902. http://dx.doi.org/10.1037/0022-3514.82.6.878

Fourie, M. M., Kilchenmann, N., Malcolm-Smith, S., & Thomas, K. G. (2012). Real-time elicitation of moral emotions using a prejudice paradigm. *Frontiers in Psychology, 3,* 275–288. http://dx.doi.org/10.3389/fpsyg.2012.00275

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin, 132,* 692–731. http://dx.doi.org/10.1037/0033-2909.132.5.692

Gawronski, B., & Bodenhausen, G. V. (2011). The associative-propositional evaluation model: Theory, evidence, and open questions. *Advances in Experimental Social Psychology, 44,* 59–127. http://dx.doi.org/10.1016/B978-0-12-385522-0.00002-0

Gawronski, B., & Strack, F. (2004). On the propositional nature of cognitive consistency: Dissonance changes explicit, but not implicit attitudes. *Journal of Experimental Social Psychology, 40,* 535–542. http://dx.doi.org/10.1016/j.jesp.2003.10.005

Gawronski, B., & Ye, Y. (2014). What drives priming effects in the affect misattribution procedure? *Personality and Social Psychology Bulletin, 40,* 3–15. http://dx.doi.org/10.1177/0146167213502548

Goldstone, R. L. (1995). Effects of categorization on color perception. *Psychological Science, 6,* 298–304. http://dx.doi.org/10.1111/j.1467-9280.1995.tb00514.x

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102,* 4–27. http://dx.doi.org/10.1037/0033-295X.102.1.4

Gregg, A. P., Seibt, B., & Banaji, M. R. (2006). Easier done than undone: Asymmetry in the malleability of implicit preferences. *Journal of Personality and Social Psychology, 90,* 1–20. http://dx.doi.org/10.1037/0022-3514.90.1.1

Hansen, T., Olkkonen, M., Walter, S., & Gegenfurtner, K. R. (2006). Memory modulates color appearance. *Nature Neuroscience, 9,* 1367–1368. http://dx.doi.org/10.1038/nn1794

Hu, X., Rosenfeld, J. P., & Bodenhausen, G. V. (2012). Combating automatic autobiographical associations: The effect of instruction and training in strategically concealing information in the autobiographical implicit association test. *Psychological Science, 23,* 1079–1085. http://dx.doi.org/10.1177/0956797612443834

Hugenberg, K., & Bodenhausen, G. V. (2003). Facing prejudice: Implicit prejudice and the perception of facial threat. *Psychological Science, 14,* 640–643. http://dx.doi.org/10.1046/j.0956-7976.2003.psci_1478.x

Hutchings, P. B., & Haddock, G. (2008). Look black in anger: The role of implicit prejudice in the categorization and perceived emotional intensity of racially ambiguous faces. *Journal of Experimental Social Psychology, 44,* 1418–1420. http://dx.doi.org/10.1016/j.jesp.2008.05.002

Hutchings, V. L. (2009). Change or more of the same? Evaluating racial attitudes in the Obama era. *Public Opinion Quarterly, 73,* 917–942. http://dx.doi.org/10.1093/poq/nfp080

Inzlicht, M., Kaiser, C. R., & Major, B. (2008). The face of chauvinism: How prejudice expectations shape perceptions of facial affect. *Journal of Experimental Social Psychology, 44,* 758–766. http://dx.doi.org/10.1016/j.jesp.2007.06.004

Izard, C. E. (2007). Basic emotions, natural kinds, emotion schemas, and a new paradigm. *Perspectives on Psychological Science, 2,* 260–280. http://dx.doi.org/10.1111/j.1745-6916.2007.00044.x

Kang, S. K., & Inzlicht, M. (2012). Stigma building blocks: How instruction and experience teach children about rejection by outgroups. *Personality and Social Psychology Bulletin, 38,* 357–369. http://dx.doi.org/10.1177/0146167211426729

Kawakami, K., Dovidio, J. F., Moll, J., Hermsen, S., & Russin, A. (2000). Just say no (to stereotyping): Effects of training in the negation of stereotypic associations on stereotype activation. *Journal of Personality and Social Psychology, 78,* 871–888. http://dx.doi.org/10.1037/0022-3514.78.5.871

Kinder, D. R., & Sears, D. O. (1981). Prejudice and politics: Symbolic racism versus racial threats to the good life. *Journal of Personality and Social Psychology, 40,* 414–431. http://dx.doi.org/10.1037/0022-3514.40.3.414

Kuppens, P., Tuerlinckx, F., Russell, J. A., & Barrett, L. F. (2013). The relation between valence and arousal in subjective experience. *Psychological Bulletin, 139,* 917–940. http://dx.doi.org/10.1037/a0030811

Lai, C. K., Skinner, A. L., Cooley, E., Murrar, S., Brauer, M., Devos, T., . . . Nosek, B. A. (2016). Reducing implicit racial preferences: II. Intervention effectiveness across time. *Journal of Experimental Psychology: General, 145,* 1001–1016. http://dx.doi.org/10.1037/xge0000179

Lang, P. J., Bradley, M. M., Fitzsimmons, J. R., Cuthbert, B. N., Scott, J. D., Moulder, B., & Nangia, V. (1998). Emotional arousal and activation of the visual cortex: An fMRI analysis. *Psychophysiology, 35,* 199–210. http://dx.doi.org/10.1111/1469-8986.3520199

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology, 30,* 261–273. http://dx.doi.org/10.1111/j.1469-8986.1993.tb03352.x

Larsen, R. J., & Diener, E. (1992). Promises and problems with the circumplex model of emotion. In M. S. Clark & M. S. Clark (Eds.), *Emotion* (pp. 25–59). Thousand Oaks, CA: SAGE.

Leach, C. W., Iyer, A., & Pedersen, A. (2006). Anger and guilt about ingroup advantage explain the willingness for political action. *Personality and Social Psychology Bulletin, 32,* 1232–1245. http://dx.doi.org/10.1177/0146167206289729

Leland, E. (2015, August 1). The tragic path from a 911 call to a fatal confrontation. *The Charlotte Observer*. Retrieved from http://www.charlotteobserver.com/news/local/crime/article29700511.html

Levenson, R. W. (2011). Basic emotion questions. *Emotion Review, 3,* 379–386. http://dx.doi.org/10.1177/1754073911410743

Lindquist, K. A. (2013). Emotions emerge from more basic psychological ingredients: A modern psychological constructionist model. *Emotion Review, 5,* 356–368. http://dx.doi.org/10.1177/1754073913489750

Lindquist, K. A., & Barrett, L. F. (2008). Constructing emotion: The experience of fear as a conceptual act. *Psychological Science, 19,* 898–903. http://dx.doi.org/10.1111/j.1467-9280.2008.02174.x

Lindquist, K. A., Gendron, M., Oosterwijk, S., & Barrett, L. F. (2013). Do people essentialize emotions? Individual differences in emotion essentialism and emotional experience. *Emotion, 13,* 629–644. http://dx.doi.org/10.1037/a0032283

Loersch, C., & Payne, B. K. (2011). The situated inference model: An integrative account of the effects of primes on perception, behavior, and motivation. *Perspectives on Psychological Science, 6,* 234–252. http://dx.doi.org/10.1177/1745691611406921

MacCormack, J. K., & Lindquist, K. A. (2017). Bodily contributions to emotion: Schachter's legacy for a psychological constructionist approach. *Emotion Review, 9,* 36–45. http://dx.doi.org/10.1177/1754073916639664

Mackie, D. M., Devos, T., & Smith, E. R. (2000). Intergroup emotions: Explaining offensive action tendencies in an intergroup context. *Journal of Personality and Social Psychology, 79,* 602–616. http://dx.doi.org/10.1037/0022-3514.79.4.602

Mackie, D. M., Maimer, A. T., & Smith, E. R. (2009). Intergroup emotions theory. In T. D. Nelson (Ed.), *Handbook of prejudice, stereotyping, and discrimination* (pp. 285–307). New York, NY: Psychology Press.

Mackie, D. M., Smith, E. R., & Ray, D. G. (2008). Intergroup emotions and intergroup relations. *Social and Personality Psychology Compass, 2,* 1866–1880. http://dx.doi.org/10.1111/j.1751-9004.2008.00130.x

Maitner, A. T., Mackie, D. M., & Smith, E. R. (2007). Antecedents and consequences of satisfaction and guilt following ingroup aggression. *Group Processes & Intergroup Relations, 10,* 223–237. http://dx.doi.org/10.1177/1368430207075154

Maner, J. K., Kenrick, D. T., Becker, D. V., Robertson, T. E., Hofer, B., Neuberg, S. L., . . . Schaller, M. (2005). Functional projection: How fundamental social motives can bias interpersonal perception. *Journal of Personality and Social Psychology, 88,* 63–78. http://dx.doi.org/10.1037/0022-3514.88.1.63

McConnell, A. R., & Leibold, J. M. (2001). Relations among the Implicit Association Test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology, 37,* 435–442. http://dx.doi.org/10.1006/jesp.2000.1470

McGarty, C., Pedersen, A., Leach, C. W., Mansell, T., Waller, J., & Bliuc, A. M. (2005). Group-based guilt as a predictor of commitment to apology. *British Journal of Social Psychology, 44*(Part 4), 659–680. http://dx.doi.org/10.1348/014466604X18974

Muris, P., Meesters, C., Bouwman, L., & Notermans, S. (2015). Relations among behavioral inhibition, shame- and guilt-proneness, and anxiety disorders symptoms in non-clinical children. *Child Psychiatry and Human Development, 46,* 209–216. http://dx.doi.org/10.1007/s10578-014-0457-3

Navarrete, C. D., Olsson, A., Ho, A. K., Mendes, W. B., Thomsen, L., & Sidanius, J. (2009). Fear extinction to an out-group face: The role of target gender. *Psychological Science, 20,* 155–158. http://dx.doi.org/10.1111/j.1467-9280.2009.02273.x

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2,* 175–220. http://dx.doi.org/10.1037/1089-2680.2.2.175

Olsson, A., Ebert, J. P., Banaji, M. R., & Phelps, E. A. (2005). The role of social groups in the persistence of learned fear. *Science, 309,* 785–787. http://dx.doi.org/10.1126/science.1113551

Oosterwijk, S., Topper, M., Rotteveel, M., & Fischer, A. H. (2010). When the mind forms fear: Embodied fear knowledge potentiates bodily reactions to fearful stimuli. *Social Psychological & Personality Science, 1,* 65–72. http://dx.doi.org/10.1177/1948550609355328

Panksepp, J., & Watt, D. (2011). What is basic about basic emotions? Lasting lessons from affective neuroscience. *Emotion Review, 3,* 387–396. http://dx.doi.org/10.1177/1754073911410741

Payne, B. K., Brown-Iannuzzi, J. L., & Panter, A. T. (2010). *AMP race database*. Unpublished raw data, Department of Psychology, University of North Carolina, Chapel Hill, NC.

Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology, 89,* 277–293. http://dx.doi.org/10.1037/0022-3514.89.3.277

Payne, K., & Lundberg, K. (2014). The affect misattribution procedure: Ten years of evidence on reliability, validity, and mechanisms. *Social and Personality Psychology Compass, 8,* 672–686. http://dx.doi.org/10.1111/spc3.12148

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology, 75,* 811–832. http://dx.doi.org/10.1037/0022-3514.75.3.811

Plant, E. A., Devine, P. G., Cox, W. L., Columb, C., Miller, S. L., Goplen, J., & Peruche, B. M. (2009). The Obama effect: Decreasing implicit prejudice and stereotyping. *Journal of Experimental Social Psychology, 45,* 961–964. http://dx.doi.org/10.1016/j.jesp.2009.04.018

Prokasy, W. F., & Kumpfer, K. L. (1973). Classical conditioning. In W. F. Prokasy & D. C. Raskin (Eds.), *Electrodermal activity in psychological research* (pp. 157–202). New York, NY: Academic Press. http://dx.doi.org/10.1016/B978-0-12-565950-5.50008-9

Rankin, R. E., & Campbell, D. T. (1955). Galvanic skin response to Negro and White experimenters. *Journal of Abnormal and Social Psychology, 51,* 30–33. http://dx.doi.org/10.1037/h0041539

Ray, D. G., Mackie, D. M., Smith, E. R., & Terman, A. W. (2012). Discrete emotions elucidate the effects of crossed-categorization on prejudice. *Journal of Experimental Social Psychology, 48,* 55–69. http://dx.doi.org/10.1016/j.jesp.2011.07.011

Roseman, I. J. (2011). Emotional behaviors, emotivational goals, emotion strategies: Multiple levels of organization integrate variable and consistent responses. *Emotion Review, 3,* 434–443. http://dx.doi.org/10.1177/1754073911410744

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39,* 1161–1178. http://dx.doi.org/10.1037/h0077714

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110,* 145–172. http://dx.doi.org/10.1037/0033-295X.110.1.145

Russell, J. A. (2005). Emotion in human consciousness is built on core affect. *Journal of Consciousness Studies, 12*(8–10), 26–42.

Russell, J. A. (2009). Emotion, core affect, and psychological construction. *Cognition and Emotion, 23,* 1259–1283. http://dx.doi.org/10.1080/02699930902809375

Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology, 76,* 805–819. http://dx.doi.org/10.1037/0022-3514.76.5.805

Schmitt, M., Behner, R., Montada, L., Muller, L., & Muller-Fohrbrodt, G. (2000). Gender, ethnicity, and education as privileges: Exploring the generalizability of the existential guilt reaction. *Social Justice Research, 13,* 313–337. http://dx.doi.org/10.1023/A:1007640925819

Seger, C. R., Banerji, I., Park, S. H., Smith, E. R., & Mackie, D. M. (2017). Specific emotions as mediators of the effect of intergroup contact on

prejudice: Findings across multiple participant and target groups. *Cognition and Emotion, 31,* 923–936. http://dx.doi.org/10.1080/02699931.2016.1182893

Shapiro, J. R., Ackerman, J. M., Neuberg, S. L., Maner, J. K., Vaughn Becker, D., & Kenrick, D. T. (2009). Following in the wake of anger: When not discriminating is discriminating. *Personality and Social Psychology Bulletin, 35,* 1356–1367. http://dx.doi.org/10.1177/0146167209339627

Smith, E. R. (1993). Social identity and social emotions: Toward new conceptualizations of prejudice. In D. M. Mackie & D. L. Hamilton (Eds.), *Affect, cognition, and stereotyping: Interactive processes in group perception* (pp. 297–315). San Diego, CA: Academic Press. http://dx.doi.org/10.1016/B978-0-08-088579-7.50017-X

Smith, E. R. (1999). Affective and cognitive implications of a group becoming part of the self: New models of prejudice and of the self-concept. In D. Abrams & M. A. Hogg (Eds.), *Social identity and social cognition* (pp. 183–196). Oxford, England: Basil Blackwell.

Smith, E. R., & Mackie, D. M. (2015). Dynamics of group-based emotions: Insights from intergroup emotions theory. *Emotion Review, 7,* 349–354. http://dx.doi.org/10.1177/1754073915590614

Stern, R. M., Ray, W. J., & Quigley, K. S. (2000). *Psychophysiological recording* (2nd ed.). New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780195113594.001.0001

Swim, J. K., & Miller, D. L. (1999). White guilt: Its antecedents and consequences for attitudes toward affirmative action. *Personality and Social Psychology Bulletin, 25,* 500–514. http://dx.doi.org/10.1177/0146167299025004008

Tangney, J. P., Miller, R. S., Flicker, L., & Barlow, D. H. (1996). Are shame, guilt, and embarrassment distinct emotions? *Journal of Personality and Social Psychology, 70,* 1256–1269. http://dx.doi.org/10.1037/0022-3514.70.6.1256

Terbeck, S., Kahane, G., McTavish, S., Savulescu, J., Cowen, P. J., & Hewstone, M. (2012). Propranolol reduces implicit negative racial bias. *Psychopharmacology, 222,* 419–424. http://dx.doi.org/10.1007/s00213-012-2657-5

Tuch, S. A., & Hughes, M. (2011). Whites' racial policy attitudes in the twenty-first century: The continuing significance of racial resentment. *The Annals of the American Academy of Political and Social Science, 634,* 134–152. http://dx.doi.org/10.1177/0002716210390288

Uhlmann, E. L., Brescoll, V. L., & Paluck, E. L. (2006). Are members of low status groups perceived as bad, or badly off? Egalitarian negative associations and automatic prejudice. *Journal of Experimental Social Psychology, 42,* 491–499. http://dx.doi.org/10.1016/j.jesp.2004.10.003

Vidulich, R. N., & Krevanick, F. W. (1966). Racial attitudes and emotional response to visual representation of the Negro. *The Journal of Social Psychology, 68,* 85–93. http://dx.doi.org/10.1080/00224545.1966.9919669

Vorauer, J. D., & Turpie, C. A. (2004). Disruptive effects of vigilance on dominant group members' treatment of outgroup members: Choking versus shining under pressure. *Journal of Personality and Social Psychology, 87,* 384–399. http://dx.doi.org/10.1037/0022-3514.87.3.384

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin, 98,* 219–235. http://dx.doi.org/10.1037/0033-2909.98.2.219

Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review, 107,* 101–126. http://dx.doi.org/10.1037/0033-295X.107.1.101

Wilson-Mendenhall, C. D., Barrett, L. F., Simmons, W. K., & Barsalou, L. W. (2011). Grounding emotion in situated conceptualization. *Neuropsychologia, 49,* 1105–1127. http://dx.doi.org/10.1016/j.neuropsychologia.2010.12.032

Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology, 72,* 262–274. http://dx.doi.org/10.1037/0022-3514.72.2.262

Wittenbrink, B., Judd, C. M., & Park, B. (2001). Spontaneous prejudice in context: Variability in automatically activated attitudes. *Journal of Personality and Social Psychology, 81,* 815–827. http://dx.doi.org/10.1037/0022-3514.81.5.815

Wundt, W. (1998). *Outlines of psychology.* Bristol, UK: Thoemmes Press. (Original work published 1897)

Ziegert, J. C., & Hanges, P. J. (2005). Employment discrimination: The role of implicit attitudes, motivation, and a climate for racial bias. *Journal of Applied Psychology, 90,* 553–562. http://dx.doi.org/10.1037/0021-9010.90.3.553