Modern prejudice: Subtle, but unconscious? The role of Bias Awareness in Whites' perceptions of personal and others' biases

Sylvia P. Perry [a,*], Mary C. Murphy [b], John F. Dovidio [c]

[a] Department of Psychological Science, University of Vermont, 2 Colchester Ave., Burlington, VT 05401, USA
[b] Department of Psychological and Brain Sciences, Indiana University, 1101 E. 10th St., Bloomington, IN 47405, USA
[c] Department of Psychology, Yale University, 2 Hillhouse Ave., New Haven, CT 06520, USA

HIGHLIGHTS

- We introduce a new individual difference measure of racial bias awareness.
- Whites endorse items indicating they are bias-aware.
- Bias awareness predicts Whites' responses to bias feedback.
- Bias-aware individuals are more attuned to others' biases.
- These findings have important implications for anti-bias training.

ARTICLE INFO

ABSTRACT

Three studies introduced the construct of bias awareness and examined its effect on Whites' responses to evidence of personal and others' racial biases. Contemporary theories of prejudice suggest that awareness of personal bias is a critical step in reducing one's prejudice and discrimination. When bias is a cloaked in a way that people do not recognize, they are likely to continue to perpetuate their biased behaviors and unlikely to reduce their negative attitudes. However, when people become aware of their biases, they often adjust their attitudes and behavior to be more egalitarian. The present research investigated (a) individual differences in Whites' awareness of their propensity to express subtly biased behavior against Blacks in interracial contexts (Study 1), (b) the convergent and discriminant validity of a new individual difference measure of bias awareness (Studies 1, 2, and 3), (c) whether this measure uniquely predicts Whites' responses to a difficult race-related context—receiving feedback that they are high in implicit bias from an Implicit Association Test (R-IAT; Study 2), and (d) whether this measure uniquely predicts Whites' perceptions of others' racial bias, particularly subtle expressions (Study 3). Results revealed that the Bias Awareness Scale measures a distinct construct that uniquely predicts Whites' emotional and behavioral responses to information about their own bias, and their ability to detect bias in others.

© 2015 Elsevier Inc. All rights reserved.

Awareness of personally-held biases are widely considered a critical step in reducing an individual's prejudice and discrimination. Rokeach and Cochkane (1972), in classic research using their value confrontation procedure, demonstrated that making people aware of inconsistencies between their core value of equality and racially biased attitudes produced less prejudice and lower levels of discrimination immediately and for several weeks after the intervention (Grube, Mayton, & Ball-Rokeach, 1994; Penner, 1971). More recent work has revealed that when Whites become aware of their biases (Monteith, 1993; Monteith, Arthur, & Flynn, 2010a; Monteith, Mark, & Ashburn-Nardo, 2010b) or when their goals to be non-prejudiced are made salient

(Moskowitz & Li, 2011), they often adjust their attitudes and behavior to be more egalitarian, and they work harder to compensate for prejudiced behavior. Alternatively, when bias is cloaked in ways that people do not recognize, they are likely to continue to perpetuate their biased behaviors and unlikely to change their negative attitudes. Thus, understanding whether there are individual differences in people's sensitivity to and concerns about their expressions of subtle bias (i.e., their bias awareness) is timely and important.

The extant literature has provided a plethora of evidence that there are individual differences in Whites' (a) motivations to be non-biased (Devine, Plant, Amodio, Harmon-Jones, & Vance, 2002; Plant & Devine, 1998) and egalitarian (Moskowitz, Gollwitzer, Wasel, & Schaal, 1999), (b) acknowledgement of discrepancies between how they should versus would act in interracial contexts (Monteith & Voils, 1998), (c) defensiveness when confronted with evidence of their bias, and

* Corresponding author.
E-mail addresses: sylvia.perry@uvm.edu (S.P. Perry), mcmpsych@indiana.edu (M.C. Murphy), john.dovidio@yale.edu (J.F. Dovidio).

(d) ability to compensate when they recognize their bias (Moskowitz, Li, Ignarri, & Stone, 2011). Although these lines of work all implicate the fundamental importance of being aware of and concerned about bias, no previously published work directly illuminates individual differences in the extent to which people are aware of and concerned about their own an others' subtle bias. We term this phenomenon *bias awareness*.

Models of controlling prejudice and attaining egalitarian goals argue that a concerned awareness of bias is a critical factor for prejudice reduction (Grube et al., 1994; Monteith, 1993; Monteith et al., 2010a, 2010b; Moskowitz et al., 2011; Penner, 1971; Rokeach & Cochkane, 1972)—this makes one's ability to become aware of bias, especially subtle bias, extremely important. For example, Monteith and colleague's self-regulation of prejudice model posits that when a low- prejudiced person has a negative implicit evaluation of an outgroup member (of which he or she may or may not be aware), this evaluation leads to the recognition of a discrepancy between his or her egalitarian goals and his or her negative behavior toward the outgroup (Monteith, 1993; Monteith, Ashburn-Nardo, Voils, & Czopp, 2002; and Monteith & Mark, 2009). As a result of this discrepancy, low-prejudiced people will (a) respond by interrupting their prejudiced behavior (and/or they will display a compensatory response), (b) experience feelings of guilt, and (c) reflect on their prejudiced behavior. By contrast, high-prejudiced individuals will experience more other-directed negative affect (e.g., anger) rather than self-directed negative emotion (e.g., guilt) when their biased behavior is made salient, and consequently will not exhibit compensatory responses.

This line of research on the responses people make when bias-related discrepancies are salient underscores the critical role of a concerned awareness of bias – that is, not simply being aware but also being concerned about such bias is central – in self-regulation processes to reduce personal bias. While previous work has theorized that bias awareness is powerful because it allows one to access processes of control—discrepancy, motivation, compensation, satiation—that work has experimentally induced, but not measured concerned awareness as an individual difference variable. Therefore, we believe a measure that assesses individual differences in concerned awareness of bias provides a novel and significant contribution.

The current work extends and complements Monteith and colleague's self-regulation of prejudice model (Monteith, 1993; Monteith et al., 2002; Monteith & Mark, 2009) by illuminating whether there are systematic individual differences in concerned awareness of bias in the self and in others, and whether this awareness will affect the previously theorized self-regulatory processes associated with Whites' egalitarian goals. In the present research we develop and test a measure of the concerned awareness, the Bias Awareness Scale, and investigate how it predicts Whites' responses in race-relevant contexts in ways that go beyond existing measures. We hypothesize that the initiation of the process that triggers the previously described motives and compensatory responses will systematically differ as a function of individual differences in Whites' Bias Awareness. We propose that, because people who score relatively high on the Bias Awareness Scale are attuned to the possibility that they may exhibit subtle biases, individual differences in Bias Awareness will relate to one's ability to detect the discrepancy between one's egalitarian beliefs and prejudiced behavior. This process will subsequently influence one's ability to correct for his or her behavior.

As suggested by the literature cited above, an early recognition of one's biases may also be affected by other individual-level factors and differences, such as people's motivations to respond without prejudice or their actual levels of (implicit or explicit) prejudice (Perry, Dovidio, Murphy, & van Ryn, 2015). We propose, however, that individual differences as assessed by the Bias Awareness Scale can influence race-relevant behaviors independent of people's levels of prejudice or motivation to be non-biased. Although personal levels of prejudice (Monteith, 1993) and individual differences in internal or external

motivations to be nonprejudiced (Plant & Devine, 1998) also systematically relate to the extent and ways people regulate race-related behavior, because it is hypothesized to represent a discrete element in the self-regulation process, individual differences in Bias Awareness should affect responses in unique ways that go beyond the effects of existing scales. For instance, individuals who score relatively high on the Bias Awareness Scale may be more open to accepting and responding constructively to evidence suggesting that they have demonstrated subtle bias. This is because in general, people find feedback that is consistent (versus inconsistent) with their self-concept to be more credible (Swann, Chang-Schneider, & McClarty, 2007). Moreover, when the feedback is disconfirming of, and threatening to, their self-concepts, people often respond defensively (Ditto & Boardman, 1995; Ditto, Scepansky, Munro, Apanovitch, & Lockhart, 1998; Sherman & Cohen, 2002). Greater personal awareness of bias may also have effects beyond self-regulatory behavior. Because people are better able to process information about others when it is consistent with how they process information about themselves (Gillihan & Farah, 2005; Rogers, Kuiper, & Kirker, 1977), people who score higher on the Bias Awareness Scale may be more attuned to the existence of cues or information signaling subtle bias in their environment and thus more likely to see actions by others subtly biased.

The present research develops and tests the Bias Awareness Scale and investigates the systematic effects of bias awareness of Whites in the US on a range of race-related behaviors and perceptions, independent of level of prejudice, motivation to be nonprejudiced (Devine et al., 2002; Plant & Devine, 1998), strength of egalitarian goals (Moskowitz et al., 1999), and recognition of discrepancies between what one would do and should do in interracial contexts (Monteith et al., 2010a, 2010b; Monteith & Mark, 2009). Individual differences based on the Bias Awareness Scale are hypothesized to operate as an antecedent to intergroup responses and behavior. A concerned awareness of bias, represented by scores on the Bias Awareness Scale, for example, may critically relate to how Whites respond to potential evidence of racial bias in themselves and others. Specifically, we investigated whether Whites show individual differences on the Bias Awareness Scale, and whether bias awareness has implications for how Whites view their own and others' racial bias against Blacks. Understanding whether there are individual differences on the Bias Awareness Scale can illuminate how and why interventions may have different effects across individuals (drawing on a person × situation perspective; Snyder & Ickes, 1985) and offer insights into cross-situational patterns of responses across individuals.

In the present work, we propose and examine the hypothesis that people who score higher on the Bias Awareness Scale will be more likely to recognize and accept information that their behavior in a given race-relevant context reflects bias. In particular people scoring high on the Bias Awareness Scale are hypothesized to be more likely to acknowledge that an action is racially biased, and consequently experience more negative emotion and/or a goal to be egalitarian (Moskowitz et al., 2011). The goal, once triggered, leads to compensatory responses aimed at reducing the tension and addressing the discrepancy.

Three studies examine how individual differences on the Bias Awareness Scale relate to Whites' reactions to evidence of personal bias and their perceptions of others' bias. Specifically, we investigate the extent to which bias awareness predicts (a) Whites' acceptance of information suggesting that they may hold negative implicit attitudes against Blacks, (b) whether people are motivated (or not) to respond in non-prejudiced and/or compensatory ways, and, (c) people's detection of others' subtle racial bias.

1. The present research

The first goal of the present research was to create and validate a measure of concerned awareness of subtle bias. We developed the scale in the context of Whites' orientations toward Blacks. We focused

on Whites' potential awareness of subtle bias against Blacks because of the historical and contemporary importance of White racism in the US, the amount of research evidence in psychology currently devoted to this form of prejudice, and the fact that other relevant measures of intergroup orientations and motivations largely (albeit not exclusively) concern Whites' responses to Blacks. The Bias Awareness Scale differs from other explicit measures of contemporary prejudice in that it (a) assesses individual differences in the extent to which people have a concerned awareness about the expression of their subtle bias, and (b) assesses individual differences in how attuned people are to the expression of subtle bias in others. Study 1 assessed the extent to which the Bias Awareness Scale converges with and diverges from currently established measures of explicit bias, racial and non-racial measures related to people's concerns about socially appropriate behavior, and need for approval. Study 2 directly tested whether individual differences on the Bias Awareness Scale moderated people's responses to and willingness to accept feedback that they were implicitly racially biased. Finally, Study 3 investigated whether the Bias Awareness Scale is distinct from another, potentially related, measure that taps awareness of discrepancies from personal standards: the Should–Would Discrepancies measure (Devine, Monteith, Zuwerink, & Elliot, 1991; Monteith & Voils, 1998), and whether individual differences on the Bias Awareness Scale have consequences that extend beyond the intra-psychic domain (i.e., self-regulation and personal responses to discrepancies), to predict Whites' sensitivity to and awareness of *other* people's subtle bias as well.

1.1. Study 1

Study 1 used four different samples to create and validate the Bias Awareness Scale, which assesses the degree to which Whites have awareness, suspicion, and concerns that they may harbor subtle biases against Blacks. The Bias Awareness Scale is hypothesized to be related to but distinct from a range of other concepts and measures. Because individuals high on the Bias Awareness Scale are theoretically attuned to and aware of their own (and others') behavior in intergroup contexts, we would expect that Whites' scores on the Bias Awareness Scale would be positively related to, but distinct from (a) racial measures related to people's concerns about the expression of racially inappropriate behavior (i.e., the External Motivation to Respond Without Prejudice Scale [EMS], Should–Would Discrepancies, and the Motivation to Control Prejudiced Reactions Scale [MCPR]; Dunton & Fazio, 1997; Monteith & Voils, 1998; Plant & Devine, 1998), and (b) a general measure of sensitivity to and concerns about socially appropriate behavior (i.e., self-monitoring; Snyder, 1974). We also examined the relationship between the Bias Awareness Scale and people's need for approval (i.e., Social Desirability Scale; Crowne & Marlowe, 1960).

Given that the Bias Awareness Scale is a measure of Whites' concerned awareness about their bias against Blacks, it is possible that the measure is positively related to people's Internal Motivation to Respond Without Prejudice (IMS; Plant & Devine, 1998), because high Internal Motivation to Respond Without Prejudice individuals are particularly concerned, for personal reasons, about expressing prejudice. However, we believed that bias-aware Whites' concerns center on their awareness of a propensity display subtle bias, and thus we anticipated that the Bias Awareness Scale would be distinct from, and thus only weakly correlated with, the Internal Motivation to Respond Without Prejudice Scale.

Additionally, we posited that the Bias Awareness Scale is a measure of Whites' concerned *awareness* of their propensity to express bias against Blacks, and not a measure of the amount of bias they hold. That is, the Bias Awareness Scale is *not* hypothesized to be a measure of Whites' levels of racial prejudice. As such, we anticipated that the measure would be distinct from (i.e., weakly correlated with) explicit measures of prejudice (e.g., Modern Racism; McConahay, 1986).

Finally, we expected that if the items on the Bias Awareness Scale were submitted to a principal components analysis (PCA) along with the other measures of racial attitudes (i.e., External Motivation to

Respond Without Prejudice, Internal Motivation to Respond Without Prejudice, Modern Racism, Motivation to Control Prejudice, and Should–Would Discrepancies), they would consistently load as a unique factor with low cross-loadings, providing additional evidence for discriminant validity.

2. Method

2.1. Participants

A total of 819 White participants from four different samples participated in this study for partial fulfillment of course requirement, $5, or entry into a lottery to win a $25 Amazon.com gift card.

2.2. Samples

The four different samples allowed for an examination of the psychometric properties of the Bias Awareness Scale and examined how the scale related to other relevant measures. The data were collected across four different samples to reduce participant fatigue in any one group. Measure order and item order were randomized in each sample. Sample 1 consisted of 469 White undergraduates (63% women) at a large Midwestern university. The participants ranged in age from 17–63 with a mean age of 22.82 ($SD = 5.61$). Samples 2, 3, and 4 consisted of 130, 71, and 149 White participants from Yale School of Management's online national participant pool (Elab) (68%, 65%, and 67% of women, respectively; and ages ranged from 18–65, $M = 38$, $SD = 12.49$; and 18–73, $M = 39$, $SD = 13.68$; to 18–75, $Mode = 30$, respectively), in which community members can participate in online research for entry into a lottery for an Amazon.com gift card.

Sample 1 included the Bias Awareness Scale ($\alpha = .78$; mean inter-item $r = .45$), the Modern Racism Scale ($\alpha = .80$), and Motivation to Respond Without Prejudice (External and Internal; $\alpha s = .90$, and .80, respectively) Scales. Sample 2 included the Bias Awareness Scale ($\alpha = .82$; mean inter-item $r = .56$), Social Desirability Scale ($\alpha = .83$), Modern Racism Scale ($\alpha = .92$), and the Motivation to Respond Without Prejudice Scales (External and Internal; $\alpha s = .82$ and .87, respectively), and Sample 3 included the Bias Awareness Scale ($\alpha = .87$; mean inter-item $r = .60$), Self-Monitoring Scale (Snyder, 1974; $\alpha = .77$), and Motivation to Control Prejudice Scale ($\alpha = 81$). Sample 4 included the Bias Awareness Scale ($\alpha = .84$), Modern Racism Scale (McConahay, 1986; $\alpha = .90$), Motivation to Respond Without Prejudice Scales (External and Internal subscales; Plant & Devine, 1998; $\alpha = .83$ and .88, respectively), Social Desirability Scale (Crowne & Marlowe, 1960; $\alpha = .82$), and the Should–Would Discrepancies Scale (Monteith & Voils, 1998; should subscale $\alpha = .93$; would subscale $\alpha = .91$). These measures took approximately 15 min for participants to complete.

2.3. Materials and procedure

Each participant sample completed either an online or paper-and-pencil questionnaire consisting of between three and four measures, including the Modern Racism Scale (McConahay, 1986), the Motivation to Respond Without Prejudice Scale (Plant & Devine, 1998), Bias Awareness Scale, the Motivation to Control Prejudiced Reactions Scale (Dunton & Fazio, 1997), Self-Monitoring Scale (Snyder, 1974), the Social Desirability Scale (Crowne & Marlowe, 1960), Should–Would Discrepancies Scale (Monteith & Voils, 1998), and basic demographic questions (e.g., race, gender, age).

2.3.1. Bias Awareness

The Bias Awareness Scale was designed to assess Whites' concerned awareness about their subtle biases against Blacks. The scale consists of four items: "Even though I know it's not appropriate, I sometimes feel that I hold unconscious negative attitudes toward Blacks"; "When talking to Black people, I sometimes worry that I am unintentionally

acting in a prejudiced way"; "Even though I like Black people, I still worry that I have unconscious biases toward Blacks"; and "I never worry that I may be acting in a subtly prejudiced way toward Blacks" (reverse scored). Participants responded on 7-point scales anchored with *strongly agree* and *strongly disagree*, in which higher scores indicate higher levels of bias awareness.

2.3.2. Motivation to Respond Without Prejudice

The Motivation to Respond Without Prejudice Scale assesses the reasons or motivations people have for behaving in non-prejudiced ways toward Blacks (Plant & Devine, 1998). There are two motivation subscales. People responding to either subscale endorse the idea that they aim to be non-prejudiced toward Blacks, however the underlying motivations for doing so differ. The external-social motivations subscale consists of five items such as "Because of today's PC (politically correct) standards I try to appear non-prejudiced toward Black people," whereas the internal-personal motivations subscale consists of five items such as "I am personally motivated by my beliefs to be non-prejudiced toward Black people." For both subscales, participants responded on 7-point scales anchored with the labels *strongly agree* and *strongly disagree*, with higher scores indicating greater motivation to respond without prejudice.

2.3.3. Modern Racism

The Modern Racism Scale is an explicit measure of prejudice used to assess Whites' levels of contemporary or subtle racism as opposed to old-fashioned or more overt racism (McConahay, 1986). The scale consists of seven items such as "Discrimination against Blacks is no longer a problem in the United States." Participants responded on 7-point scales anchored with the labels *strongly agree* and *strongly disagree*, in which higher scores indicate higher levels of modern racism.

2.3.4. Self-Monitoring

The Self-Monitoring Scale is designed to assess individual differences in concern for social appropriateness, and the extent to which people use social situations as cues for monitoring and managing how they present themselves (Snyder, 1974). The scale consists of 25 items such as "When I am uncertain how to act in a social situation, I look to the behavior of others for cues." Participants responded *True* or *False*, with true answers coded as 1, and false answers coded as 0. Participants' answers were summed to form a Self-Monitoring score, and higher scores indicate higher levels of self-monitoring.

2.3.5. Social Desirability

The Social Desirability Scale is a measure of need for approval in social situations (Crowne & Marlowe, 1960). The scale consists of 33 items such as "I never hesitate to go out of my way to help someone in trouble." Participants responded *True* or *False*, with true answers coded as 1, and false answers coded as 0. Participants' answers were summed to form a Social Desirability score, and higher scores indicate higher levels of social desirability.

2.3.6. Motivation to Control Prejudiced Reactions Scale

The Motivation to Control Prejudiced Reactions Scale is an individual difference measure of the extent to which people seek to inhibit their prejudiced responses (Dunton & Fazio, 1997). The scale consists of 17 items such as, "In today's society, it is important that one not be perceived as prejudiced in any manner." Participants responded on 7-point scales anchored with the labels *strongly agree* and *strongly disagree*, with higher scores indicating greater motivation to control prejudice.

2.3.7. Should–Would Discrepancies Scale

The Should–Would Discrepancies Scale is meant to assess individual differences in the extent to which people will report a discrepancy between their non-prejudiced attitudes and prejudiced behaviors

(Monteith & Voils, 1998). The scale consists of 16 *should* items, such as "I believe that I should not think of Blacks in stereotypical ways" and 16 corresponding *would* items, such as "I sometimes have stereotypical racial thoughts." Participants responded on 7-point scales anchored with the labels *strongly agree* and *strongly disagree* for each of the respective subscales, then their ratings on the should items were subtracted from their corresponding would items to form a discrepancy proneness score. Higher scores indicate a greater discrepancy between what people say they should-versus-would do in each of the respective situations.

3. Results

3.1. Correlational analyses

The primary goal of Study 1 was to establish the discriminant and convergent validity of the new Bias Awareness measure. To this end, correlational analyses were conducted with four samples (See Tables 1 and 2 for scale correlations and *p*-values).

As expected, and consistent with our hypothesis that the Bias Awareness Scale is a measure of Whites' concerns about the expression of and consequences of their subtle biases against Blacks, we found a moderate and positive relation between the Bias Awareness Scale with the External Motivation to Respond Without Prejudice ($rs = .36$ to .40), Motivation to Control Prejudice ($r = .37$), and Self-Monitoring ($r = .41$) Scales.

Suggesting that people who are high on the Bias Awareness Scale are less concerned with a need for approval, there was a moderate and negative relation between the Bias Awareness and Social Desirability ($r = -.30$ to $-.38$) Scales. Moreover, consistent with our belief that the Bias Awareness Scale distinguishes between people who are aware of and concerned about their propensity to express cues of bias from those who are not, but does not measure people's personal motivations to be non-prejudiced, we found no significant relationship between the Bias Awareness and Internal Motivation to Respond Without Prejudice ($rs = -.09$ to $-.12$) Scales. Consistent with our hypothesis that the Bias Awareness Scale is not a measure of racial prejudice, we found no significant relation between the Bias Awareness and Modern Racism ($rs = -.03$ to .07) Scales. Finally, as expected, and consistent with our hypothesis that Bias Awareness is related to, but distinct from the Should–Would Discrepancies measure, we found that the Bias Awareness Scale was moderately and positively correlated with Should–Would ($r = .39$).

Table 1
Intercorrelations of Measures of Racial Attitudes and Bias Awareness (Study 1, Samples 1 and 2).

Measure	r(p)					
	M	SD	1	2	3	4
Sample 1 N = 469						
1. Bias Awareness	4.11	1.29	–	−.09 (.045)	.39 (<.001)	−.03 (.564)
2. Internal Motivation to Respond Without Prejudice	5.39	1.14		–	−.05 (.310)	−.55 (<.001)
3. External Motivation to Respond Without Prejudice	3.80	1.21			–	.14 (.003)
4. Modern Racism	2.68	1.02				–
Sample 2 N = 130						
1. Bias Awareness	3.65	1.39	–	−.12 (.480)	.36 (<.001)	.07 (.271)
2. Internal Motivation to Respond Without Prejudice	5.35	1.23		–	−.17 (.035)	−.66 (<.001)
3. External Motivation to Respond Without Prejudice	3.96	1.37			–	.29 (<.001)
4. Modern Racism	3.03	1.38				–

Table 2
Intercorrelations of Measures of Racial Attitudes and Bias Awareness (Study 1, Samples 3 and 4).

Sample 3 N = 71				
			r(p)	
Measure	M	SD	1	2
1. Bias Awareness	3.61	1.52	–	−.37 (.002)
2. Motivation to Control Prejudice	4.19	.76		–

Sample 4 N = 149							
			r(p)				
Measure	M	SD	1	2	3	4	5
1. Bias Awareness	3.68	1.36	–	−.03 (.937)	.40 (<.001)	.03 (.511)	.39 (<.001)
2. Internal Motivation to Respond Without Prejudice	5.30	1.22		–	−.14 (.159)	−.62 (<.001)	.17 (.045)
3. External Motivation to Respond Without Prejudice	3.94	1.34			–	.24 (.003)	.29 (<.001)
4. Modern Racism	3.09	1.24				–	.04 (.618)
5. Discrepancy Score	10.68	10.85					–

3.2. Principal components analysis

Given the moderate to small correlations between Bias Awareness and other measures of Whites' racial attitudes and orientations, we further examined the discriminant validity of the Bias Awareness measure by submitting the items from the Bias Awareness Scale and other measures of racial attitudes to principal components analyses (PCA). Racial attitudes items from each sample were submitted to a PCA with a direct oblimin rotation. An oblique rotation was employed to provide a more conservative test of whether Bias Awareness is a distinct construct (by not forcing the factors to be orthogonal; Tabachnick & Fidell, 2000). For Sample 1, all items from the Modern Racism Scale, Internal and External Motivations to Respond Without Prejudice Scales, and Bias Awareness Scale were submitted. The factor analysis revealed four dimensions with eigenvalues greater than one. The first factor included the Modern Racism items and accounted for 26.47% of the variance (eigenvalue = 5.56; with factor loadings ranging from .85 to .39); the second factor included the External Motivation to Respond Without Prejudice items and accounted for 17.75% of the variance (eigenvalue = 3.73; with factor loadings ranging from .86 to .67); the third factor included the Bias Awareness items and accounted for 8.88% of the variance (eigenvalue = 1.86; loadings ranged from .82 to .63); and the fourth factor included the Internal Motivation to Respond Without Prejudice items and accounted for 6.69% of the variance (eigenvalue = 1.40; with factor loadings ranging from .88 to .71). All items consistently loaded with the correct scale. Cross-loadings between Bias Awareness and other scale items were between .02 and .37 in magnitude.

Using Sample 2, we again submitted all items from the Modern Racism Scale, Internal and External Motivations to Respond Without Prejudice Scales, and Bias Awareness Scale to a PCA with an oblique rotation. Consistent with Sample 1, results revealed four factors with eigenvalues greater than one. The first factor included the Modern Racism items and accounted for 37.58% of the variance (eigenvalue = 7.51; with factor loadings ranging from .85 to .76); the second factor included the Bias Awareness items and accounted for 17.35% of the variance (eigenvalue = 3.47; with factor loadings ranging from .87 to .61); the third factor included the Internal Motivation to Respond Without Prejudice items and accounted for 9.47% of the variance (eigenvalue = 1.89; loadings ranged from .87 to .72); and the fourth factor included the External Motivation to Respond Without Prejudice items and accounted for 6.61% of the

variance (eigenvalue = 1.32; with factor loadings ranging from .86 to .71). All items consistently loaded with the correct scale. Cross-loadings between Bias Awareness and other scale items were between .06 and 32 in magnitude.

Using Sample 3, we submitted all items from the Motivation to Control Prejudice Scale and the Bias Awareness Scale to establish that, although Bias Awareness is related to MCP, it is a separate construct. Results revealed six factors with eigenvalues greater than one. The first factor included the Bias Awareness items and accounted for 34.68% of the variance (eigenvalue = 6.94; with factor loadings ranging from .90 to .66). However, the additional five factors included the MCP items accounting for a total of 48.22% of the variance (eigenvalues ranging from 1.01 to 3.65; with factor loadings ranging from .88 to .52; .77 to .59, .73 to .71; .57 to .48, and .78 to .36, for each respective factor). All items consistently loaded with the correct scale. Cross-loadings between Bias Awareness and other scale items were .001 between and .42 in magnitude.

Using Sample 4, we focused on the Bias Awareness and Should–Would Discrepancy items, only, submitting participants' scores to a principal components analysis with a direct oblimin rotation. Results revealed six factors with eigenvalues greater than one, with the Bias Awareness Scale loading onto one factor, and the Should–Would Discrepancy scores loading onto five separate factors. The discrepancy score items accounted for 50.10% of the variance (with eigenvalues ranging from = 1.01 to 4.91; and factor loadings ranging from .81 to .76 across the respective factors). The Bias Awareness factor accounted for 13.30% of the variance (eigenvalue = 2.66; with factor loadings ranging from .90 to .77) and cross-loadings were between .02 and .39. We also investigated whether the separate Should–Would item scores would load separately from the Bias Awareness items. Thus, we submitted the Bias Awareness items and participants' Should–Would Scale items to a principal components analyses with a direct oblimin rotation. Again, results revealed six factors with eigenvalues greater than one. The Bias Awareness Scale loaded onto one factor, and the Should–Would items loaded onto five separate factors. The Should–Would Scale items accounted for 58.12% of the variance (with eigenvalues ranging from 1.10 to 14.64) and had factor loadings ranging from .89 to .56. The Bias Awareness factor accounted for 10.83% of the variance (eigenvalue = 3.90; with factor loadings ranging from .88 to .75). Again, we found that the Bias Awareness items loaded onto one factor. Cross-loadings between Bias Awareness and other scale items were between .02 and .19 in magnitude.

4. Discussion

Study 1 provided initial evidence that the Bias Awareness Scale measures a distinctive construct. As hypothesized, the Bias Awareness items consistently loaded on a separate factor than did items from other relevant scales, demonstrated high reliability, and moderate inter-item correlations (Clark & Watson, 1995). Moreover, the scale was (a) only moderately positively related to measures that assess people's concerns about social appropriateness (EMS, social desirability, MCP, and Should–Would Discrepancies), (b) moderately negatively related to a measure of need for approval (social desirability), and (c) unrelated to measures of people's personal motivations to respond without prejudice (IMS) and self-reported racial bias (Modern Racism).

Although scores on the Bias Awareness Scale are positively associated with Whites' external motivations to be non-prejudiced, we hypothesized that in contexts in which Whites' propensity to be subtly biased is made salient, the Bias Awareness Scale should predict outcomes above and beyond the External Motivation to Respond Without Prejudice Scale (EMS). Thus, in Study 2, we investigated the unique predictive power of Bias Awareness in a context in which Whites' potential to be subtly biased was made salient—after they received bias feedback from the race-IAT (R-IAT). We used feedback on the R-IAT because we

hypothesized that scores on the Bias Awareness Scale are associated with a concern about expressing prejudice in uncontrolled, unintentional ways. Additionally, we examined the unique predictive validity of the Bias Awareness Scale above and beyond current explicit and *implicit* measures of bias. For reasons outlined below, we expected that, above and beyond established measures of racial bias and Whites' motivations to respond without prejudice, differences in Bias Awareness would moderate White's acceptance of, and emotional and behavioral responses to bias feedback and would affect their subsequent attitudes and behavior toward racial minorities. Additionally, we expected that low Bias-Aware individuals would be particularly dismissive of the high compared to low bias feedback. Finally, a measure of participants' public and private self-consciousness was included in the online questionnaire to examine the extent to which the constructs related to Bias Awareness.

4.1. Study 2

With Study 2, we sought to further validate the Bias Awareness Scale by investigating whether scores on the Bias Awareness Scale would independently predict Whites' responses to a situation in which the existence of their subtle bias against Blacks was made highly salient—after receiving feedback that they exhibited racial bias on the R-IAT. Previous theory suggests that there are often individual differences in people's responses to bias feedback (Monteith, Voils, & Ashburn-Nardo, 2001). When confronted with the possibility that they might be biased, Whites work hard to avoid violating norms, and attempt to bring their behavior more in line with egalitarian norms (Son Hing, Li, & Zanna, 2002). Given that individuals scoring higher on the Bias Awareness Scale are theorized to be particularly sensitive to and attuned to their propensity to violate egalitarian norms, we expected that, unlike their lower awareness counterparts, Whites would be more accepting of the validity of feedback suggesting they are biased against Blacks and, consequently, would be more motivated to bring their behavior in line with their egalitarian norms after receiving this feedback. Thus, the goals of Study 2 were to (a) examine the extent to which Whites' scores on the Bias Awareness Scale predicts their responses to bias feedback from the R-IAT, and (b) investigate whether Bias Awareness predicts these responses above and beyond current measures of Whites' explicit racial bias, implicit racial bias, and motivations to respond without prejudice. In this study, participants were randomly assigned to receive high or low bias feedback from the R-IAT. Afterward, we assessed the extent to which the bias feedback interacted with Whites' Bias Awareness to predict their emotional responses, perceived validity of the feedback, and desire for intergroup contact.

5. Hypotheses

People generally find feedback that is more consistent (versus inconsistent) with their self-concept to be more credible (Swann et al., 2007). In contrast, people tend to respond defensively to information inconsistent with their self-image (Ditto & Boardman, 1995; Ditto et al., 1998; Sherman & Cohen, 2002). Thus, we hypothesized that the extent to which White participants accepted the feedback indicating bias against Blacks, and their emotional and behavioral responses to this feedback would vary as a function of their individual differences in Bias Awareness. We propose that, because they are more attuned to the existence of their subtle bias, individuals who score higher on the Bias Awareness Scale will be more accepting of the bias feedback. This feedback acceptance will highlight the discrepancy between their implicit negative attitudes and their egalitarian goals, and this, in turn, will trigger a motivation to reduce their bias through compensatory behavior. In contrast, Whites who score lower on the Bias Awareness Scale will not experience the feedback as valid, and thus a perceived discrepancy (between their attitudes and their goals) will not be triggered. As a result, these individuals will not be motivated to compensate for

anything. We did not expect that the different responses for high compared to low Bias Aware individuals will happen as a function of individual differences in their personal motivations to be non-biased, but that instead these differences will occur as a function of their acceptance of the feedback as valid, or not. Our specific hypotheses are the following:

Hypotheses 1. We expected that, because of their heightened sensitivity to bias, individuals who scored higher on the Bias Awareness Scale would be more willing to accept feedback indicating they were high in implicit bias than those who were lower in Bias Awareness.

Hypothesis 2. Given that Whites who score lower on the Bias Awareness Scale are expected to respond defensively to feedback suggesting they are high in bias against Blacks, we anticipated that participants who scored higher, compared to lower, on the Bias Awareness scale would be more willing to admit and self-report negative affect in response to the bias feedback.

Hypothesis 3. Because Whites who score higher (compared to lower) on the Bias Awareness Scale are more attuned to their subtly biased tendencies against Blacks, we expected that high bias feedback would trigger a compensatory response, and as a result they would be more willing to seek contact with racial minorities.

Hypothesis 4. We expect that high Bias-Aware individuals' compensatory responses will be driven by their willingness to accept feedback indicating they are implicitly biased. That is, because Bias-Aware individuals are better able to detect bias in the self, they will be more willing to accept feedback indicating they are subtly biased, and this acceptance will in turn trigger compensatory behavior in the form of seeking out interracial contact.

Hypothesis 5. Consistent with research on egalitarian goal seeking, we do not expect the low bias feedback to trigger a compensatory response in high Bias-Aware individuals. Moreover, because low bias feedback is consistent with their self-perceptions, we do not expect that low Bias-Aware individuals will feel threatened by this feedback (or respond defensively). As such, we hypothesized that, when given low bias feedback, high and low Bias-Aware individuals would be equally willing to accept the feedback, self-report low levels of negative affect, and dedicate hours of their time toward compensatory behaviors.

6. Method

6.1. Participants

Eighty-three White undergraduates (53% women) enrolled in an introductory psychology course at a large Midwestern university participated in this study for partial fulfillment of course requirements or for $5. The participants ranged in age from 17 to 44 with a mean age of 19.41 ($SD = 2.47$).

6.2. Procedure

At least one week before participating in the experimental phase of the study, participants completed an online questionnaire. The questionnaire included the Modern Racism scale (Cronbach's $\alpha = .80$; McConahay, 1986), the Internal and External Motivation to Respond Without Prejudice Scales (Cronbach's $\alpha = .90$ and .80, respectively; Plant & Devine, 1998), the Bias Awareness Scale (Cronbach's $\alpha = .81$), the Public and Private-Self Consciousness Scales (Cronbach's $\alpha = .79$ and .54, respectively; Fenigstein, Scheier, & Buss, 1975), and demographic questions (e.g., race, gender, age). These measures took approximately 15 min to complete. The purpose of the pre-experimental

assessment was to measure participants' Bias Awareness and their baseline attitudes toward Blacks.

At the laboratory session, participants were greeted by a Black experimenter who seated them in front of a computer where they completed several tasks and questionnaires presented with Medialab© software. Even though the experimenter could not see their bias feedback, a Black experimenter was used to heighten participants' psychological experience of receiving racial bias feedback during the study. Participants first completed a "categorization task" (the R-IAT) that would measure their "implicit associations" and provide them with a "categorization score." The R-IAT consists of pleasant or unpleasant words and Black or White faces, shown one at a time in the middle of the computer screen while the category labels (e.g., "White" or "good") appear in the upper left or right corner of the screen. After participants completed two practice trials, the first block of trials required participants to categorize targets as "*White or good*" or "*Black or bad.*" In the second block, the categories were switched (*Black or good/White or bad*). Trials within these blocks were displayed in random order. When the stimulus appeared, participants responded by pressing the specified key to categorize the stimulus. Participants' reaction times to the trials were recorded so that their actual R-IAT scores could be calculated.

Following the R-IAT, participants viewed a screen where they were asked to wait for a few seconds while the computer "calculated their score." After 10 s, a new screen appeared, and participants were presented with their randomly assigned feedback—ostensibly their score on the R-IAT. Participants were randomly provided with one of two types of bias feedback that are identical to the actual feedback that people receive after completing the online R-IAT (Greenwald, Nosek, & Banaji, 2003). Participants in the "low bias feedback" condition read that "the data show[ed] that [they had] little to no automatic preference between Black and White people," whereas participants in the "high bias feedback" condition read that "the data show[ed] that [they had] a strong automatic preference for White people compared to Black people." After receiving feedback, participants' emotional responses to the feedback were assessed by the PANAS (Watson, Clark, & Tellegen, 1988). Next, participants completed a measure that assessed the extent to which they desired to participate in an activity in which they would help and interact with racial minorities. This measure asked participants to volunteer hours personally assisting with fictitious campus diversity week activities. Afterward, participants answered a question that assessed their acceptance or rejection of the feedback. The laboratory session lasted approximately 45 min.

6.3. Measures

The primary dependent variables were participants' their willingness to accept the feedback and their emotional and behavioral responses to the bias feedback (detailed below). However, participants' implicit racial attitudes (R-IAT) and explicit racial attitudes (including the Bias Awareness scale) were assessed in the experimental session and in the online pre-measure session, respectively, so that we could compare the predictive power of the Bias Awareness scale with the other measures of prejudice. Finally, a measure of participants' public and private self-consciousness was included in the online questionnaire to examine the extent to which the constructs related to Bias Awareness.

6.3.1. The Public/Private Self-Consciousness Scale

The public self-consciousness scale (Fenigstein et al., 1975) consists of seven items such as, "I'm concerned about my way of doing things," whereas the private self-consciousness subscale consists of ten items such as, "I reflect a lot about myself." Participants responded on 7-point scales anchored with the labels *strongly agree* and *strongly disagree* in which higher scores indicated higher public or private self-consciousness (Cronbach's αs = .79 and .54, respectively).

6.3.2. Race Implicit Association Test

Participants' R-IAT scores were calculated using the improved scoring algorithm (Greenwald et al., 2003), which yields an effect size estimate (the IAT *D* score, a variant of Cohen's *d*) for each participant, with positive *D* scores indicating a relative preference for Whites over Blacks and negative *D* scores indicating a relative preference for Blacks over Whites. The *D* scores represented the difference in mean response latency between the incongruent (*Black or good*; *White or bad*) and congruent (*White or good*; *Black or bad*) blocks, divided by the pooled standard deviation.

6.3.3. Personal acceptance of feedback

The question "*How likely is it that your score reflects your true levels of implicit prejudice?*" assessed participants' personal acceptance of the bias feedback. Participants responded to this item on a 7-point scale anchored with the labels *not at all likely* and *very likely*. This question allowed us to examine how people across feedback conditions and Bias Awareness levels varied in their acceptance of the feedback.

6.3.4. Emotional responses

Although we were primarily interested in participants' negative emotional responses to the bias feedback, to avoid suspicion, both positive (empowered, hopeful, proud, optimistic, enthusiastic, satisfied, content, and pleased) and negative (angry, anxious, ashamed, disgusted, embarrassed, fearful, furious, guilty, irritated, regretful, sad, sorry, and worried) emotions were assessed with items such as, "To what extent did the feedback you received make you feel [angry, happy, etc.]?"(PANAS; Watson et al., 1988). Participants responded on 4-point scales anchored by the labels *strongly agree* and *strongly disagree*, in which higher scores indicate greater negative or positive affect. Factor analysis revealed that the negative emotions loaded onto the first factor and explained 31.6% of the variance. Thus, responses to these items were averaged to create a *negative affect* score (Cronbach's α = .87).[1]

6.3.5. Willingness to seek contact with and help out racial minorities

Participants read about a fictitious campus diversity week initiative that allowed them to volunteer their time in the coming month to activities supportive of racial minorities. They learned that their university was trying to enhance diversity on campus and was relying on students to help prepare for and assist with the diversity week events. Participants responded to questions about whether they would be willing to participate in eight supportive activities, and if so, the number of hours they would be willing to dedicate. For example, participants indicated whether they were willing to volunteer to become a one-on-one peer contact that would answer questions about the university for ethnically diverse students or whether they would give minority prospective students a tour of the university. Participants could volunteer zero to twenty hours of their time toward these activities. Desire for future interracial contact was measured by the number of hours participants were willing to volunteer toward these interpersonal activities (see Fig. 1 for the questionnaire).

---

[1] Participants' negative affect results did not differ by specific emotion. We presume that this was the case because there was very little variability between the individual emotions. While we were primarily interested in the relation between the bias feedback condition and self-reported negative affect, we also analyzed participants' positive affective responses to the feedback. Our results revealed no significant differences in positive affect as a function of feedback condition by Bias Awareness measure. Specifically, results revealed that while there was a main effect of condition, such that participants who received no bias feedback compared to bias feedback reported more positive affect, $b = -0.90$ ($SE_b = .29$), $t(75) = -3.03$, $p < .001$, individual differences in Bias Awareness did not affect positive affect, $b = 0.15$ ($SE_b = .16$), $t(75) = 0.91$, $p = .37$, and the Bias Awareness by Condition interaction was nonsignificant, $b = -0.20$ ($SE_b = .20$), $t(75) = -0.95$, $p = .34$.

PIN_____

The university administration is implementing a number of different programs to help increase and maintain ethnic diversity at UIC. They are relying on student involvement to make these diversity efforts a success. Participation in these activities is completely voluntary.

Please indicate whether you would be willing to participate in UIC's diversity efforts next school year. If you indicate "Yes", please tell us the number of hours you would be willing to participate over the course of a semester (with a **maximum** of **20** hours total).

Please fill in the circle for "Yes" or "No". Would you be willing to:

		Yes	No
1.	Attend social events sponsored by the diversity month committee.	○	○
	If yes, how many hours would you be willing to commit? _____		
2.	Make phone calls to alumni to help fund a diversity scholarship?	○	○
	If yes, how many hours would you be willing to commit? _____		
3.	Sign-up to give ethnic minority prospective students tours during UIC's diversity month?	○	○
	If yes, how many hours would you be willing to commit? _____		
4.	Put fliers up around campus to advertise for a diversity event?	○	○
	If yes, how many hours would you be willing to commit? _____		
5.	Volunteer as a peer contact for ethnic minority students to answer questions about UIC?	○	○
	If yes, how many hours would you be willing to commit? _____		

		Yes	No
6.	Volunteer to sit on a student panel to discuss diversity at UIC and your personal experiences of diversity with ethnic minority prospective students?	○	○
	If yes, how many hours would you be willing to commit? _____		
7.	Write a Congress person for more state funding for diversity efforts at UIC?	○	○
	If yes, how many hours would you be willing to commit? _____		
8.	Conduct online and library research on programs that UIC might implement in its diversity efforts?	○	○
	If yes, how many hours would you be willing to commit? _____		

Please provide your email address so that we may contact you about these activities in the future: _____

Fig. 1. Willingness to interact with and help racial minorities measure (Study 2).

7. Results

7.1. Discriminant validity of the Bias Awareness Scale

Consistent with Study 1, the Bias Awareness measure had strong discriminant validity. A correlational analysis including the Bias Awareness Scale, the Modern Racism Scale, the Motivation to Respond Without Prejudice (Internal and External) Scales, and R-IAT scores replicated the finding that Bias Awareness is not strongly correlated with conventional measures of implicit or explicit prejudice, or motivations to respond without prejudice (all r's < .31; see Table 3). Moreover, Bias Awareness was not significantly related to public or private self-consciousness ($r = .16$, and $r = .03$, respectively).

We again submitted all items from the Modern Racism Scale, Internal and External Motivations to Respond Without Prejudice Scales, and Bias Awareness Scale to a principal components analysis with direct oblimin rotation. This time we requested a four-factor solution based on the observed factor structures in Study 1 (although the Bias Awareness items load onto one factor in an exploratory factor analysis). The first factor included the Internal Motivation to Respond Without Prejudice items and accounted for 24.96% of the variance (eigenvalue = 5.24; with factor loadings ranging from .87 to .77); the second factor included the External Motivation to Respond Without Prejudice items and accounted for 18.42% of the variance (eigenvalue = 3.88; with factor loadings ranging from .89 to .46); the third factor included the Bias Awareness items and accounted for 10.92% of the variance (eigenvalue = 2.29; with factor loadings ranging from .83 to .75); and the fourth factor included the Modern Racism items and accounted for 6.93% of the variance (eigenvalue = 1.46; with factor loadings ranging from .73 to .54). These results provide further evidence that Bias Awareness is an independent construct.

7.2. Preliminary analyses

Because research has suggested that people can "feel" how they are performing on the IAT (Monteith et al., 2001), participants' responses to the pre-determined bias feedback would differ as a function of their actual IAT scores. Thus, we conducted preliminary analysis to determine whether this was indeed the case. The IAT × Condition interaction revealed that there were no significant effects on participants' believability in the feedback, negative affect, or behavioral responses as a function of this interaction, $b = -0.15$ ($SE_b = 1.13$), $t(79) = -0.13$, $p = .90$; $b = -0.12$ ($SE_b = 0.31$), $t(79) = -0.38$, $p = .70$; $b = -0.54$ ($SE_b = 3.82$), $t(79) = -0.14$, $p = .88$, respectively.

Finally, we initially analyzed all higher order interactions, but they were not significant. Thus, we present the trimmed model below.

7.3. Does Bias Awareness moderate Whites' acceptance of bias feedback?

Recall that an interaction with between Feedback Condition and Bias Awareness was expected, such that, because they are more attuned to the existence of their subtle bias, participants who scored higher (compared to lower) on the Bias Awareness Scale would show more acceptance of the high bias feedback. In contrast, we expected that both high and low bias aware individuals might believe feedback signaling they are lower in implicit bias. A moderated regression analysis (following the procedures recommended by Aiken & West, 1991; Dawson & Richter, 2006) with the dummy-coded bias feedback condition (high bias = 1, low bias = 0; the moderator variable of interest) and the standardized Bias Awareness items in the first block and the Bias Awareness by condition interaction term in the second block revealed a significant main effect for bias feedback condition, such that participants were less accepting of their feedback in the high, compared to low, bias feedback condition, $b = -2.18$ ($SE_b = .47$), $t(77) = -4.64$, $p < .001$. However, this main effect was qualified by the predicted interaction, $b = 1.23$ ($SE_b = .49$), $t(77) = 2.54$, $p = .01$ (see Fig. 2). Simple slopes analyses at

Table 3
Intercorrelations of Measures of Racial Attitudes and Bias Awareness (Study 2).

Study 2 N = 83							
Measure	r(p)						
	M	SD	1	2	3	4	5
1. Bias Awareness	3.99	1.32	–	−.08 (.971)	.02 (.821)	.31 (<.003)	.08 (.873)
2. R-IAT Score	0.51	0.45		–	−.13 (.482)	−.12 (.757)	.16 (.298)
3. Internal Motivation to Respond Without Prejudice	5.33	1.24			–	−.26 (.013)	−.45 (<.001)
4. External Motivation to Respond Without Prejudice	3.88	1.22				–	.20 (.267)
5. Modern Racism	3.00	1.06					–
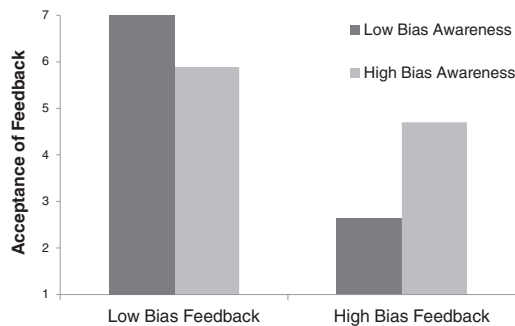
Fig. 2. Acceptance of bias feedback as a function of bias awareness and bias feedback condition, where high bias feedback is one, and low bias feedback is zero (Study 2).

high and low (−1 *SD*) levels of bias feedback revealed that Bias Awareness did not affect acceptance when participants were in the low bias feedback condition, $b = -0.20$ ($SE_b = 0.37$), $t(77) = -0.55$, $p = .56$. These participants were equally likely to accept their bias feedback, regardless of their level of Bias Awareness. Participants in the high bias feedback condition however, were more accepting of the feedback when they were higher in Bias Awareness, $b = -1.05$ ($SE_b = 0.33$), $t(77) = 3.32$, $p = .002$. That is, participants lower in Bias Awareness were more likely to discount and deny the high bias feedback than those who were higher in Bias Awareness[2,3].

7.4. Does Bias Awareness moderate Whites' emotional and behavioral responses to bias feedback?

We tested our hypotheses regarding participants' affective and behavioral responses to bias feedback above and beyond existing

measures using hierarchical moderated regression. Following the procedures recommended by Aiken and West (1991) and Dawson and Richter (2006), the first regression block included (a) the dummy-coded bias feedback condition (high bias = 1, low bias = 0; the independent variable of interest), (b) the standardized Modern Racism, R-IAT, and Internal/External Motivation to Respond Without Prejudice scores, and (c) the standardized Bias Awareness scores (the moderator of interest). The second block included the interaction terms of the Modern Racism, R-IAT, Internal/External Motivation to Respond Without Prejudice, and Bias Awareness scores with the bias feedback condition. If Bias Awareness predicted the outcomes above and beyond the other measures of explicit and implicit bias, as we expected it to, the Bias Awareness by Feedback Condition interaction term should remain significant after controlling for all of the other variables.

Table 4
Unstandardized Regression Weights for the Effects of Bias Awareness on Contact Behavior and Self-Reported Negative Affect (Controlling for External/Internal Motivation to Respond Without Prejudice, Modern Racism, R-IAT Scores, and Each Measure by Condition; Study 2).

Block		Diversity activities	Self-Reported Negative Affect
1	Bias Feedback Condition	1.14	0.50**
	External Motivation to Respond Without Prejudice (EMS)	−0.04	−0.49
	Internal Motivation to Respond Without Prejudice (IMS)	1.87	−0.04
	Modern Racism	0.49	0.12
	R-IAT	−2.91	0.12
	Bias Awareness	0.64	−0.01
2	R^2	.32	0.50**
	Bias Feedback Condition	0.59	0.60**
	External Motivation to Respond Without Prejudice (EMS)	0.82	0.20
	Internal Motivation to Respond Without Prejudice (IMS)	−0.02	−0.02
	Modern Racism	0.98	0.02
	R-IAT	−3.31	0.16
	Bias Awareness	−2.24	−0.04
	Bias Feedback Condition × R-IAT	0.51	−0.20
	Bias Feedback Condition × IMS	2.53	0.30
	Bias Feedback Condition × EMS	−1.11	−0.44*
	Bias Feedback Condition × Modern Racism	−0.93	0.19
	Bias Feedback Condition × Bias Awareness	4.27*	0.32*
	R^2	.45*	.59*

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

---

[2] Even though Study 1 shows that, across multiple samples, the Bias Awareness items load onto one factor, we acknowledge that two of the items ("Even though I know it's not appropriate, I sometimes feel that I hold unconscious negative attitudes toward Blacks" and "Even though I like Black people, I still worry that I have unconscious biases toward Blacks") are more complexly worded than the other two items ("When talking to Black people, I sometimes worry that I am unintentionally acting in a prejudiced way" and "I never worry that I may be acting in a subtly prejudiced way toward Blacks" [reverse scored]). Supplementary analysis revealed that each of the 2-item sets produced the same patterns of results, and consistent with the results for the 4-item scale. A similar analysis for Study 3 also showed the same pattern of findings. Additional details are available from the first author.

[3] Previous research has suggested that IMS and EMS may produce multiplicative effects in addition to additive effects. However, in the present study, when the IMS × EMS × Condition 3-way and the IMS × EMS 2-way interactions were entered into the previously presented model the results were nonsignificant (*p*s > .77) while the Bias Awareness × Condition interactions remained significant [$b = 0.31$, ($SE_b = 0.14$), $t(71) = 2.19$, $p = .03$; behavior, $b = 4.49$, ($SE_b = 1.94$), $t(71) = 2.32$, $p = .042$].

7.4.1. Negative affect

A significant main effect of bias feedback condition emerged, such that participants reported more negative affect when they received high compared to low bias feedback (see Table 4). More importantly, and consistent with the hypothesis, Bias Awareness significantly qualified the relationship between bias feedback condition and participants' negative affect (see Table 4). To examine this interaction, we tested the simple slopes of Bias Awareness in the high (condition = 1) and low (condition = 0) bias feedback conditions (Aiken & West, 1991; Holmbeck, 2002), adjusting for participants' Modern Racism, Motivation to Respond Without Prejudice (Internal and External), and R-IAT scores, as well as the four interaction terms of these measures with the bias feedback condition. In support of our hypothesis, and consistent with the literature on defensive responding (Ditto & Boardman, 1995; Ditto et al., 1998; Sherman & Cohen, 2002), even when adjusting for other measures of prejudice, and importantly people's External Motivations to Respond Without Prejudice, individuals high compared to low in Bias Awareness reported significantly more negative affect when they were told they were high in implicit racial bias, $b = 0.27$ ($SE_b = .09$), $t(71) = 3.14$, $p = .002$. In contrast, when participants were told they were low in implicit bias, there was no difference as a result of their levels of Bias Awareness, $b = -0.04$ ($SE_b = .11$), $t(71) = -.37$, $p = .71$ (see Fig. 3).
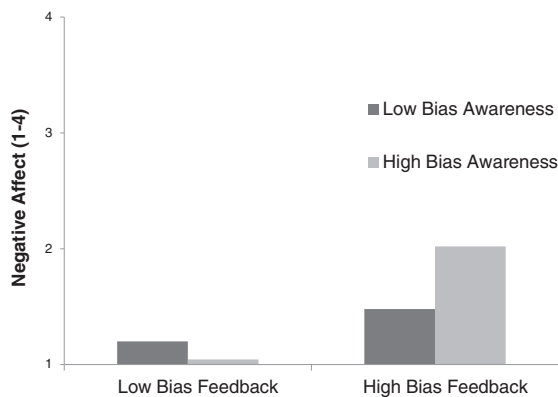

Fig. 3. Self-Reported Negative Affect as a function of bias awareness and bias feedback condition, where high bias feedback is one, and low bias feedback is zero (Study 2).

7.4.2. Willingness to seek contact with and help out racial minorities

The hierarchical moderated regression revealed only the predicted interaction between feedback condition and Bias Awareness on participants' willingness to interact with and help racial minorities (see
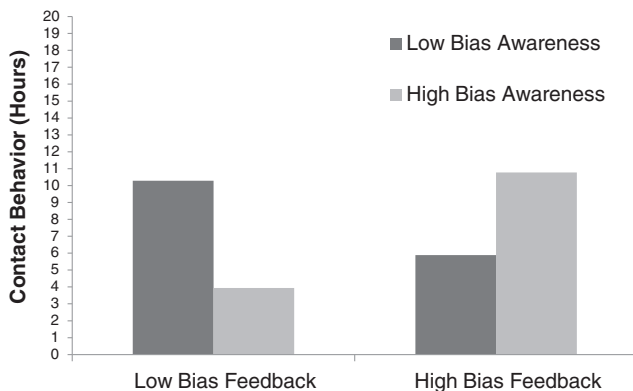

Fig. 4. Hours dedicated to diversity activities as a function of bias awareness and bias feedback condition, where high bias feedback is one, and low bias feedback is zero (Study 2).

Table 4). To examine this interaction, we tested the simple slopes of Bias Awareness in the high (condition = 1) and low (condition = 0) bias feedback conditions, adjusting for participants' Modern Racism, Motivation to Respond Without Prejudice (Internal and External), and R-IAT scores, as well as the four interaction terms of these measures with the bias feedback condition. Consistent with our hypothesis, individuals high compared to low in Bias Awareness were more likely to seek contact after receiving feedback that they were high, compared to low, in implicit bias, $b = 2.10$ ($SE_b = 1.19$), $t(71) = 1.76$, $p = .08$. In contrast, when participants were told they were low in implicit bias, there was no difference in intention to seek contact as a result of their levels of Bias Awareness, $b = -2.23$ ($SE_b = 1.54$), $t(71) = -1.45$, $p = .15$ (see Fig. 4).

7.5. Is the effect of the Bias Awareness by feedback condition interaction on behavior mediated through acceptance of feedback?

Next, we investigated our primary hypothesis regarding whether the effect of the Feedback Condition by Bias Awareness interaction on participants' willingness to dedicate hours toward increasing and maintaining diversity was explained through their acceptance of the bias feedback. Specifically, we expected that the interactional effect would be explained through feedback acceptance, but only in the condition in which participated received the high bias feedback. To examine this question, we used Model 8 of the Hayes (2013) PROCESS method. This model allowed us to test the conditional indirect effect of bias awareness (in the high and low bias feedback conditions) on participants' compensatory behavior, adjusting for participants' Modern Racism, Motivation to Respond Without Prejudice (Internal and External), and R-IAT scores, as well as the four interaction terms of these measures with the bias feedback condition. Analyses based on 5000 bootstrap samples revealed an indirect effect of Bias Awareness on hours dedicated through feedback acceptance in the high, $b = 1.03$, 95% CI (0.108, 2.82), but not low, $b = 0.26$, 95% CI ($-.25$, 1.42) feedback condition. Consistent with this result, the Index of Moderated Mediation did not include zero, confirming a significant moderated mediation, $b = 0.77$ 95% CI (0.0060, 2.78). Thus, consistent with our hypothesis, a moderated mediation analysis revealed that participants' willingness to accept feedback explained the relation between the Bias Awareness by Feedback Condition interaction on participants' willingness to dedicate hours of their time toward increasing and maintaining diversity on their university campus.

8. Discussion

Study 2 examined the emotional and behavioral responses of Whites when they received bias feedback from the R-IAT. Bias Awareness uniquely moderated Whites' emotional responses and their willingness to seek contact with racial minorities in the face of bias feedback above and beyond previously established measures of implicit and explicit prejudice, and Whites' motivations to respond without racial bias. Consistent with our hypothesis, when told they were high in racial bias, White participants high, compared to low, in Bias Awareness showed acceptance of the feedback. Moreover, these individuals, relative to those lower in awareness, reported higher levels of negative affect and were more likely to seek out interracial contact that would help increase and maintain ethnic diversity at their university after receiving high bias feedback. Importantly however, it was the personal acceptance of bias feedback that drove the positive relation between participants' Bias Awareness and their desire for intergroup contact in the high bias feedback condition. When given feedback that they were biased, high, compared to low, Bias-Aware individuals were more likely to accept this feedback, and as a result, were more willing to seek out intergroup contact as a compensatory response. Overall, consistent with Study 1, Study 2 further validates the Bias Awareness Scale in the context of White–Black race relations, providing evidence that Bias Awareness is

an important and useful predictor of Whites' emotional and behavioral responses in difficult race-related contexts above and beyond current measures of explicit and implicit prejudice and, importantly, Whites' External Motivations to Respond Without Prejudice.

Although the literature on anti-bias interventions suggests that confronting Whites about their prejudice may be an important prejudice-reduction method (Ashburn-Nardo, Morris, & Goodwin, 2008), to date only a limited number of studies have investigated how Whites respond to feedback that they are racially biased in an implicit way (including Amodio, Devine, & Harmon-Jones, 2008; Czopp & Monteith, 2003; Vorauer, 2012). The present findings contribute to this growing literature by suggesting that bias feedback may be useful for some individuals, but ineffective for others, and that this difference depends on people's levels of Bias Awareness. When White individuals have awareness of their subtle biases, they are more likely to accept feedback that they are biased, as well as increase their willingness to seek interracial contact, presumably to demonstrate their egalitarian values or gain practice in reducing their biases. However, when Whites lack personal awareness of their subtly biased attitudes, they reject feedback that that they hold implicit racial bias, and their behaviors and emotions remain unaffected.

Although it has been said that intervention methods to reduce old-fashioned, overt forms of prejudice should differ from methods used to reduce more contemporary and subtle forms of prejudice (Dovidio & Gaertner, 1999), the present findings suggest that intervention methods designed to target even contemporary forms of prejudice may need to take a more tailored approach. That is, unless people have awareness of their subtle biases, bias feedback may lead to unintended negative consequences for future interracial interactions as a result of people rejecting the feedback.

Although we established in Study 1 that Bias Awareness is moderately related to but distinct from a measure of discrepancies (Should–Would), we are unable to rule out the possibility that discrepancies alone explain Study 2's results. It could be argued that Bias Awareness is simply a measure of Whites' knowledge of the discrepancies between their egalitarian attitudes and biased behaviors. Thus, with Study 3 we investigated whether Bias Awareness is distinct from a previously established measure of discrepancies (the Should–Would Discrepancies measure; Monteith & Voils, 1998), and whether, as we previously theorized, Bias Aware individuals are more attuned to examples of bias (and subtle biases, in particular) in others.

8.1. Study 3

The goal of Study 3 was to further validate the Bias Awareness measure in the context of Whites' orientations toward Blacks by evaluating whether Bias Awareness is a unique construct relative to a theoretically similar one that also measures discrepancies—the Should–Would Discrepancies measure (Monteith & Voils, 1998). Factor analyses in Study 1 provided initial evidence that Bias Awareness and Should–Would are distinct constructs. This study, however, examined the measures' comparative predictive validity.

More specifically, we investigated whether Bias Awareness predicted Whites' general sensitivity to the expression of subtle bias against Blacks by others in Sample 4 from Study 1. Consistent with the self-reference literature, which suggests that people are better able to process information about others when it is consistent with how they process information about themselves (Gillihan & Farah, 2005; Rogers et al., 1977), we anticipated that because relatively high Bias-Aware individuals are particularly self-aware of their propensity to be subtly biased, they would be better able to perceive others' subtly biased behaviors as racist as than those low in Bias Awareness. Moreover, to the extent that Bias Awareness is distinct from Should–Would discrepancies, the predicted effect of Bias Awareness on perceptions of others' bias should emerge even when controlling for the Should–Would Scale.

The Should–Would measure was created to assess whether people have the self-awareness necessary to self-report their prejudiced tendencies. More specifically, following theory from Devine (1989), Monteith and Voils (1998), and Monteith et al. (2001), the Should–Would measure is meant to assess whether people will report a discrepancy between their nonprejudiced attitudes and prejudiced behaviors, and whether these discrepancies are predictive of their behaviors. People who have a high discrepancy score, on average, show a large gap between their *actual* behaviors and their *intended* behaviors, and this difference ostensibly suggests that these individuals believe they *should not* be biased, despite the reality that they *are* biased.

Consistent with our hypothesis that Bias Awareness is not primarily about people's concerns about the outward expression of their prejudice, but instead assesses people's *awareness* of and concerns about their expression of their *subtle* bias, we found in Study 2 that Bias Awareness continued to predict Whites' responses to feedback that they were implicitly biased even when controlling for Whites' Internal and External Motivations to Respond Without Prejudice. However, it could be argued that, similar to the Should–Would measure, Bias Awareness is also a measure of discrepancies between people's actual and intended behaviors. That is, people may choose to endorse Bias Awareness questions that suggest that even though they wish they did not behave in a subtly biased way (what one *should* do), they nonetheless do behave in a subtly biased way (what one *would* do). If Bias Awareness mainly assesses perceived discrepancies between Whites' intended and actual behavior toward Blacks, Bias Awareness and the Should–Would scale would show similar correlates and effects. Thus, the goals of Study 3 were to (a) further validate the Bias Aware measure by testing the hypothesis that Bias-Aware individuals' awareness extends beyond the self and into the "other" domain, and (b) assess the unique effects of Bias Awareness controlling for Should–Would Discrepancies.

Previous work by Sommers and Norton (2006) has shown that Whites, compared to Blacks, are poor at recognizing the existence of subtle bias in others. However, because high Bias-Aware individuals are theorized to be sensitive to and attuned to subtle forms of bias and aware of the expression of their own biases, we expected that this awareness might extend to their sensitivity and awareness of others' biases as well. That is, we expect that Bias-Aware individuals will be better at detecting bias not only in themselves (as indicated by their endorsement of the Bias Awareness items), but also in other people's behavior. To investigate this notion, we asked people to evaluate the extent to which various items previously found to be exemplars of blatant or subtle bias (Sommers & Norton, 2006), were in fact examples of bias.

We hypothesized that when people were asked to evaluate whether these behaviors were racist, Bias Awareness would predict unique variance in people's identification of the subtly biased behaviors as racist above and beyond other measures of racism and the Should–Would measure. Moreover, we hypothesized that this predictive effect of Bias Awareness would be especially evident in situationally "weak" contexts where bias is subtle than in situationally "strong" contexts where bias is blatant (Snyder & Ickes, 1985; Wang, Stroebe, & Dovidio, 2012). For example, Wang et al. (2012) found that individual differences in stigma consciousness were more influential in an ambiguous prejudice condition than in a blatant prejudice condition, with the latter representing a "strong" situation in which prejudice would be clearly evident to all targets (regardless of their stigma consciousness levels). Thus, consistent with Wang et al. (2012) (as well as Mischel, 1973, 2004; Snyder & Ickes, 1985), we expected that the situationally strong context of blatant behaviors would mitigate any individual difference effects for Bias Awareness. Strong situations composed of blatantly biased behaviors should be perceived similarly for those both high and low in Bias Awareness. However, we expected that individual differences in Bias Awareness would predict people's perceptions when the context was more ambiguous or subtle. Therefore, Bias-Aware

individuals should be more attuned to these subtle behaviors and label them as prejudiced more so than those who are relatively unaware.

9. Method

9.1. Participants and procedure

After completing the pre-measures previously outlined in Study 1 (including the Should–Would Discrepancies Scale, Internal Motivation to Respond Without Prejudice, External Motivation to Respond Without Prejudice, Modern Racism, Bias Awareness, and the Social Desirability scale), participants from Sample 4 (n = 149) were asked to read 32 items and indicate on a scale from 1 to 7 the extent to which they thought the behaviors were examples of racism (Sommers & Norton, 2006). Specifically, participants were told "We [were] interested in [their] perceptions of racism…[and that they should]…use the scale below to indicate the extent to which [they] agree[d] or disagree[d] that the person described [was] racist". In previous research, the items loaded onto three different factors, which Sommers and Norton described as examples of *blatant racism*, *subtle racism*, and *denial of racism*, respectively. However, for the purpose of this study, participants answered all 32 items, and the blatant and subtle examples were retained for analyses. To reduce demand characteristics, participants answered the questions in random order, and they were unaware of the extent to which the items had previously been established as examples of "subtle" compared to "overt" bias.

9.2. Measures

We measured perceptions of blatant and subtle racial bias.

9.2.1. Blatant bias
The blatant bias subscale consists of 9 items such as, "A White person denies group membership to Blacks on account of race." Participants responded on 7-point scales anchored with the labels *strongly agree* and *strongly disagree*, in which higher scores indicate greater perception of the behavior as racist ($\alpha = .95$).

9.2.2. Subtle bias
The subtle bias subscale consists of 10 items such as, "A White person has trouble distinguishing Black people from one another." Participants responded on 7-point scales anchored with the labels *strongly agree* and *strongly disagree*, in which higher scores indicate greater perception of the behavior as racist ($\alpha = .90$).

Before investigating our primary hypothesis, we first investigated whether the blatant and subtle bias items indeed loaded onto two separate factors using a PCA with an oblique rotation (the same rotation employed by Sommers and Norton). With the exception of one item, "A White person prefers to not be around Black people," which, inconsistent with Sommers and Norton (2006), loaded more strongly on the blatant bias factor (.69) than the subtle bias factor (.30), the factor loadings were consistent with the loadings reported in Sommers and Norton (2006). Factor loadings for the subtle and blatant bias items ranged from .82 to .57, and .95 to .56, respectively, with low cross-loadings. To maintain consistency with Sommers and Norton (2006), the subtle and overt scales were kept the same and the respective 9 and 10 items were averaged together to form the subtle and blatant bias subscales ($Ms = 4.04$ and 5.07 and $SDs = 1.14$ and 1.45, respectively).

10. Results

We tested our hypothesis regarding participants' perceptions of blatant and subtle racism using regression analyses. Because we were interested in whether Bias Awareness could uniquely predict Whites'

perceptions of subtly racist behaviors above and beyond other measures of racial attitudes, including Should–Would, the regression model included participants' Modern Racism, Internal/External Motivation to Respond Without Prejudice, and their Should–Would Discrepancy scores. Additionally, when blatant racism was the outcome, the predictors included participants' assessment of the subtle racism items; and vice-versa, when subtle racism was the outcome, participants' assessments of the blatant racism items were included in the predictors. Finally participants' Bias Awareness scores were included as a predictor. If Bias Awareness predicted the outcomes above and beyond the other measures of bias, as we expected it to, the effect would remain significant even when controlling for all of the other variables.

10.1. Blatant racism

We expected that in the strong situational context of blatant behaviors, Bias Awareness would not predict participants' perceptions of blatant racism. Indeed, only modern racism predicted unique variance in the extent to which participants perceived overtly racist behaviors as racist, $b = -0.30$, $(SE_b = 0.09)$, $t(141) = -3.20$, $p = .002$. Bias Awareness had no significant effect, $b = 0.001$, $(SE_b = 0.08)$, $t(141) = 0.008$, $p = .993$. Specifically, the higher people were in modern racism, the less likely they were to perceive overtly racist behaviors as racist.

10.2. Subtle racism

Conversely, we anticipated that in weaker situational contexts comprised of subtle behaviors, Bias Awareness would uniquely predict participants' perceptions of subtle racism. Participants' Internal Motivation to Respond Without Prejudice and Modern Racism scores predicted unique variance in the extent to which people perceived subtly racist behaviors as racist, $b = 0.15$, $(SE_b = 0.07)$, $t(141) = 2.09$, $p = .04$; $b = 0.23$ $(SE_b = 0.08)$, $t(141) = 3.01$, $p = .003$. However, and consistent with our hypothesis, the results also revealed that, after controlling for other measures of racial attitudes, including Should–Would Discrepancies, Bias Awareness predicted significant and unique variance in participants' perceptions of subtly biased behaviors as racist, $b = 0.13$, $(SE_b = 0.06)$, $t(141) = 2.21$, $p = .029$, whereas the Should–Would measure did not, $b = -0.005$, $(SE_b = 0.007)$, $t(141) = -0.70$, $p = .486$[1,4].

11. Discussion

Consistent with our hypothesis, Bias Awareness explained unique variance in Whites' judgments of subtle bias toward Blacks above and beyond Should–Would Discrepancies. Additionally, although Internal Motivation to Respond Without Prejudice also explained unique variance in people's perceptions, it did not overshadow that explained by Bias Awareness. This is not surprising, given that Internal Motivation to Respond Without Prejudice and Bias Awareness are consistently non-correlated, and because Internal Motivation to Respond Without Prejudice did not predict unique variance in Whites' behavioral responses to bias feedback from the R-IAT (Study 2). These results further confirm our hypothesis that Bias Aware individuals are more generally attuned to the existence of bias whether it manifests in themselves or in others' behaviors.

It will be important for future research to investigate the extent to which high Bias-Aware individuals are more *accurate* than others in identifying subtle bias by examining outcomes such as the correspondence between participants' perceptions and targets' reported attitudes

---

[4] Previous research on Should–Would Discrepancies suggests that, if participants' *Should* scores are entered separately from their *Would* scores, their *Should* scores may significantly predict participants' identification of subtle racial bias as racism. Additional analyses revealed that, when *Should* scores were entered separately from *Would* scores in the model, they did not predict significant variance in participants' perceptions of subtle racism as racist ($p = .585$) and the effect of Bias Awareness remained significant on perceptions of subtle bias, $b = 0.13$, $(SE_b = 0.06)$, $t(140) = 2.06$ $p = .042$.

or intentions. An alternative is that high Bias-Aware individuals might be prone to projecting their own attitudes about bias onto others (West & Kenny, 2011). The projection literature suggests that people may sometimes rely *too* heavily on their own behaviors and attitudes to predict others' attitudes and behaviors (Hoch, 1987; Krueger, 2000). Consistent with the *Truth and Bias Model*, that suggests that accuracy and projection are independent components (West & Kenny, 2011), Bias Awareness might include some level of projection, but high Bias-Aware individuals may also be particularly accurate at identifying subtle bias in others because—as Study 3 demonstrates—they acknowledge the association between subtle behavioral cues and bias (e.g., distancing, fidgeting; Richeson & Shelton, 2003). If Whites who are high in Bias Awareness are particularly sensitive to cues of bias against Blacks, we would expect that, irrespective of their levels of projection, they would show greater accuracy relative to those low in Bias Awareness.

12. General discussion

Across three studies and four samples, the present research introduced a new measure of Bias Awareness reflecting individual differences in the extent to which people are aware of and concerned about the expression and consequences of their bias. In the present research, we focused on a particular instantiation of bias awareness, Whites' concerned awareness about being biased toward Blacks. Together, these studies demonstrated the internal consistency of the Bias Awareness measure (Studies 1–3), its convergent and discriminant validity (Studies 1–3), and its ability to uniquely explain how different people respond to feedback of personal bias (Study 2) as well as how they perceive bias in others (Study 3). As predicted, Bias Awareness was only modestly related to External Motivation to Respond Without Prejudice (Plant & Devine, 1998), Should–Would Discrepancies (Monteith & Voils, 1998), and Self-Monitoring (Snyder, 1974). It weakly related to Public and Private-Self Consciousness (Fenigstein et al., 1975), Modern Racism (McConahay, 1986), and Internal Motivation to Respond Without Prejudice (Plant & Devine, 1998); and it was negatively related to Social Desirability (Crowne & Marlowe, 1960). People higher in Bias Awareness were more likely to accept feedback of personal bias as credible and take action to reduce their bias. They were also more likely to perceive acts of subtle bias as racial discrimination.

In addition to the particular effects reported here, understanding and assessing differences in Bias Awareness have both conceptual and practical implications. Conceptually, considerable research has revealed that while blatant expressions of prejudice have declined over time, implicit and subtle forms of bias persist (Dovidio & Gaertner, 2004; Nosek, Banaji, & Greenwald, 2002). Moreover, these subtler forms of prejudice may be automatically activated, often with limited intentionality or control. However, people may "correct" their thoughts and actions to become more egalitarian when they become aware of their potential bias (Dovidio, Gaertner, & Kawakami, 2003). Repeatedly engaging in self-regulatory actions can help people develop automatic inhibitory mechanisms over the long term (Monteith et al., 2010a, 2010b; Monteith & Mark, 2009; Moskowitz & Ignarri, 2009). Whereas previous work has emphasized contextual features that sensitize people to their bias (Dovidio & Gaertner, 2004) or particular interventions that enhance people's awareness of what they should and would do in intergroup contexts (Monteith & Voils, 1998), Bias Awareness assesses chronic sensitivity to bias that differs systematically across individuals. The present work reveals that Bias Awareness as an individual difference can critically moderate how people respond in a range of intergroup situations, increasing explanatory power beyond conventional measures of prejudice alone.

For example, the idea of bias awareness has been implicated as an important moderator in the "shooter" task in which Whites are asked to shoot a Black or White target when they are holding a weapon, but not when they are holding an everyday object. Regardless of whether the target is holding a weapon or everyday object, Whites are more likely to shoot the Black target than the White target (Correll, Park, Judd, & Wittenbrink, 2007). However, police officers are less likely to show this shooter bias compared to lay people (Correll et al., 2007). In explaining this compelling finding, Correll and colleagues suggested that lay people and police officers exhibit different behavioral responses because police officers have developed an "automatic control of bias" through experience and self-knowledge. Specifically, Correll et al. (2007) intuited that police officers were more likely to be aware of the fact that they may associate Blacks with danger through repeated exposure to Blacks in criminal situations. Thus, officers have learned to suppress their stereotypical thoughts and behaviors, making them "experts" at correcting their bias and engaging in controlled processes compared to community members. Although the moderating role of individual differences on the Bias Awareness Scale on people's bias to shoot was not directly tested, Correll et al.'s (2007) theoretical insight suggests that this individual difference should play an important role in people's behavior when their propensity to be subtly biased is heightened—such as when they receive bias feedback (e.g., Study 2).

Building on this work, future research could directly examine how Bias Awareness shapes the dynamics of interracial interactions between Whites and Blacks and whether Bias Awareness has positive or negative implications for intergroup contact. Our results revealed that people who were higher in Bias Awareness were more likely to seek interracial contact after high bias feedback (Study 2); however, when told they were *not* racially biased, Bias Aware individuals reported *reduced* desire for contact—as if preferring to "play it safe" by avoiding future contact. Aversive racism theory has argued that aversive racists may be motivated to seek contact only when they feel the need to prove that they are not racially biased (Gaertner & Dovidio, 1986). Moreover, work by Moskowitz et al. (2011) on the dynamics of egalitarian goals has shown that people with egalitarian goals (such as Whites high in Bias Awareness) are strongly motivated to pursue their egalitarian goal through a compensatory act when their egalitarian goal is unfulfilled (e.g., challenged by high bias feedback; Moskowitz, 2002) but relax their motivation when their egalitarian goals is fulfilled (e.g., when feedback confirms they are low in prejudice; Moskowitz et al., 2011).

In interracial interactions, then, Bias Awareness may systematically affect the goals Whites have for these encounters (e.g., to reaffirm a nonprejudiced image) and consequently how Whites engage Blacks in their exchanges. Whites' efforts to appear nonprejudiced are often interpreted by Blacks as indications of racial bias (Shelton, Richeson, Salvatore, & Trawalter, 2005), and Blacks who anticipate greater rejection (e.g., those higher in racial Rejection Sensitivity; Mendoza-Denton, Downey, Purdie, Davis, & Pietrzak, 2002) may be particularly sensitive to such cues. Thus, additional research might productively consider the independent and joint effects of individual differences in Whites' Bias Awareness and in Blacks' Rejection Sensitivity on the processes and outcomes of their interracial interactions.

Future research should investigate the extent to which individuals who are willing to endorse items indicating they are aware of and concerned about their subtle bias (high Bias-Aware Whites) have more solidified egalitarian goals than those who do not endorse these items (low Bias-Aware Whites). However, these kinds of egalitarian expectations are difficult to maintain during ongoing interracial interaction and often disrupt Whites' motivations and resources, as well as the behavioral dynamics of the interaction (Murphy, Richeson, & Molden, 2011; Richeson & Shelton, 2003; Richeson & Trawalter, 2005). Thus, Bias Awareness may moderate the extent to which Whites approach interracial interactions with a promotion focus, in which they adopt a "learning" orientation versus a prevention focus, in which they respond in a guarded fashion and experience intrusive, biased thoughts that interfere with the effectiveness of the interaction and have negative consequences for the experiences of both participants (Mallett & Wilson, 2010; Migacheva & Tropp, 2012; Trawalter & Richeson, 2006; Vorauer, Gagnon, & Sasaki, 2009).

As noted earlier, we studied bias awareness in the present work solely in the context of Whites' orientations toward Blacks. We acknowledge, however, that the dynamics of Whites' biases toward Blacks has several unique qualities. For example, expressing prejudice toward Blacks is generally perceived to be much less socially acceptable than exhibiting bias against other groups, such as overweight people (Crandall, Eshleman, & O'Brien, 2002). Moreover, racial discrimination is prohibited by federal law; weight discrimination, for example, is not. Also, Whites' bias against Blacks has a particular history (e.g., related to slavery), progression of legal issues (e.g., related to "separate but equal" forms of segregation), and centrality in the social, political, and economic systems in the US that makes it different in many ways from other forms of racial and ethnic bias (Jones, Dovidio, & Vietze, 2014). Thus, future work might consider the ways in which the dynamics and outcomes of individual differences in Bias Awareness, which we studied in the context of Whites' orientations toward Blacks, are similar or different than Bias Awareness applied to other intergroup relations.

Practically, understanding the role of individual differences in Bias Awareness, alone and/or in combination with people's level of prejudice and motivations to reduce prejudice, can help create more effective and targeted anti-bias interventions and programs. Many anti-bias education programs have the goal of increasing people's awareness of their personal bias as a key component of their approach (Stephan & Stephan, 2001). Whether people are already receptive to such information (i.e., are high in Bias Awareness) may influence whether they will be receptive and responsive to the training or instead, demonstrate backlash (Legault, Gutsell, & Inzlicht, 2011; Plant & Devine, 2001). Indeed, anti-bias education programs have had mixed success in sustaining longer-term effects (Kalev, Dobbin, & Kelly, 2006). However, understanding how individual differences in Bias Awareness might critically moderate responses to anti-bias training may help guide the development of more effective interventions tailored to the orientations and needs of particular audiences.

Acknowledgments

References

Aiken, L.S., & West, Stephen G. (1991). *Multiple regression: Testing and interpreting interactions.* London: Sage Publications.

Amodio, D.M., Devine, P.G., & Harmon-Jones, E. (2008). Individual differences in the regulation of intergroup bias: The role of conflict monitoring and neural signals for control. *Journal of Personality and Social Psychology, 94*, 60–74.

Ashburn-Nardo, L., Morris, K.A., & Goodwin, S.A. (2008). The confronting prejudiced responses (CPR) model: Applying CPR in organizations. *Academy of Management: Learning and Education, 7*, 332–342.

Clark, L., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7*, 309–319.

Correll, J., Park, B., Judd, C.M., & Wittenbrink, B. (2007). The influence of stereotypes on decisions to shoot. *European Journal of Social Psychology, 37*, 1102–1117.

Crandall, C.S., Eshleman, A., & O'Brien, L. (2002). Social norms and the expression and suppression of prejudice: The struggle for internalization. *Journal of Personality and Social Psychology, 82*, 359–378.

Crowne, D.P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*, 349.

Czopp, A.M., & Monteith, M.J. (2003). Confronting prejudice (literally): Reactions to confrontations of racial and gender bias. *Personality and Social Psychology Bulletin, 29*, 532–544.

Dawson, J.F., & Richter, A.W. (2006). Probing three-way interactions in moderated multiple regression: Development and application of a slope difference test. *Journal of Applied Psychology, 91*, 917–926.

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

Devine, P.G., Monteith, M.J., Zuwerink, J.R., & Elliot, A.J. (1991). Prejudice with and without compunction. *Journal of Personality and Social Psychology, 60*, 817–830.

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.

Ditto, P.H., & Boardman, A.F. (1995). Perceived accuracy of favorable and unfavorable psychological feedback. *Basic and Applied Social Psychology, 16*, 137–157.

Ditto, P.H., Scepansky, J.A., Munro, G.D., Apanovitch, A.M., & Lockhart, L.K. (1998). Motivated sensitivity to preference-inconsistent information. *Journal of Personality and Social Psychology, 75*, 53–69.

Dovidio, J.F., & Gaertner, S.L. (1999). Reducing prejudice: Combating intergroup biases. *Current Directions in Psychological Science, 8*, 101–105.

Dovidio, J.F., & Gaertner, S.L. (2004). Aversive racism. *Advances in Experimental Social Psychology, 36*, 1–52.

Dovidio, J.F., Gaertner, S.L., & Kawakami, K. (2003). Intergroup contact: The past, present, and the future. *Group Processes & Intergroup Relations, 6*, 5–21.

Dunton, B.C., & Fazio, R.H. (1997). An individual difference measure of motivation to control prejudiced reactions. *Personality and Social Psychology Bulletin, 23*, 316–326.

Fenigstein, A., Scheier, M.F., & Buss, A.H. (1975). Public and private self-consciousness: Assessment and theory. *Journal of Consulting and Clinical Psychology, 43*, 522–527.

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

Gillihan, S.J., & Farah, M.J. (2005). Is self special? A critical review of evidence from experimental psychology and cognitive neuroscience. *Psychological Bulletin, 131*, 76–97.

Greenwald, A.G., Nosek, B.A., & Banaji, M.R. (2003). Understanding and using the implicit association test: I. an improved scoring algorithm. *Journal of Personality and Social Psychology, 85*, 197–216.

Grube, J.W., Mayton, D.M., & Ball-Rokeach, S.J. (1994). Inducing change in values, attitudes, and behaviors: Belief system theory and the method of value self-confrontation. *Journal of Social Issues, 50*, 153–173.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach.* Guilford Press.

Hoch, S.J. (1987). Perceived consensus and predictive accuracy: The pros and cons of projection. *Journal of Personality and Social Psychology, 53*, 221–234.

Holmbeck, G.N. (2002). Post-hoc probing of significant moderational and mediational effects in studies of pediatric populations. *Journal of Pediatric Psychology, 27*, 87–96.

Jones, J.M., Dovidio, J.F., & Vietze, D.L. (2014). *The psychology of diversity: Beyond prejudice and racism.* Malden, MA: Wiley Blackwell.

Kalev, A., Dobbin, F., & Kelly, E. (2006). Best practices or best guesses? Assessing the efficacy of corporate affirmative action and diversity policies. *American Sociological Review, 71*, 589–617.

Krueger, R.F. (2000). Phenotypic, genetic, and nonshared environmental parallels in the structure of personality: A view from the multidimensional personality questionnaire. *Journal of Personality and Social Psychology, 79*, 1057–1067.

Legault, L., Gutsell, J.N., & Inzlicht, M. (2011). Ironic effects of antiprejudice messages how motivational interventions can reduce (but also increase) prejudice. *Psychological Science, 22*, 1472–1477.

Mallett, R.K., & Wilson, T.D. (2010). Increasing positive intergroup contact. *Journal of Experimental Social Psychology, 46*, 382–387.

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

Mendoza-Denton, R., Downey, G., Purdie, V.J., Davis, A., & Pietrzak, J. (2002). Sensitivity to status-based rejection: Implications for African Americans' college experience. *Journal of Personality and Social Psychology, 83*, 896–918.

Migacheva, K., & Tropp, L.R. (2012). Learning orientation as a predictor of positive intergroup contact. *Group Processes & Intergroup Relations*, 1–19.

Mischel, W. (1973). *Toward a cognitive social learning reconceptualization of personality.*

Mischel, W. (2004). Toward an integrative science of the person. *Annual Review of Psychology, 55*, 1–22.

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

Monteith, M. J., & Mark, A. (2009). The self-regulation of prejudice. In T. D. Nelson (Ed.), *Handbook of prejudice, stereotyping, and discrimination* (pp. 507–523). New York: Psychology Press.

Monteith, M., Arthur, S., & Flynn, S. (2010a). Self-regulation and bias. In J.F. Dovidio, M. Hewstone, P. Glick, & V.M. Esses (Eds.), *SAGE handbook of prejudice, stereotyping, and discrimination* (pp. 493–507). London: Sage.

Monteith, M.J., Ashburn-Nardo, L., Voils, C.I., & Czopp, A.M. (2002). Putting the brakes on prejudice: On the development and operation of cues for control. *Journal of Personality and Social Psychology, 83*, 1029–1050.

Monteith, M.J., Mark, A.Y., & Ashburn-Nardo, L. (2010b). The self-regulation of prejudice: Toward understanding its lived character. *Group Processes & Intergroup Relations, 13*, 183–200.

Monteith, M.J., & Voils, C.I. (1998). Proneness to prejudiced responses: Toward understanding the authenticity of self-reported discrepancies. *Journal of Personality and Social Psychology, 75*, 901–916.

Monteith, M.J., Voils, C.I., & Ashburn-Nardo, L. (2001). Taking a look underground: Detecting, interpreting, and reacting to implicit racial biases. *Social Cognition*, *19*, 395–417.

Moskowitz, G.B. (2002). Preconscious effects of temporary goals on attention. *Journal of Experimental Social Psychology*, *38*, 397–404.

Moskowitz, G.B., Gollwitzer, P.M., Wasel, W., & Schaal, B. (1999). Preconscious control of stereotype activation through chronic egalitarian goals. *Journal of Personality and Social Psychology*, *77*, 167.

Moskowitz, G.B., & Ignarri, C. (2009). Implicit volition and stereotype control. *European Review of Social Psychology*, *20*, 97–145.

Moskowitz, G.B., & Li, P. (2011). Egalitarian goals trigger stereotype inhibition: A proactive form of stereotype control. *Journal of Experimental Social Psychology*, *47*, 103–116.

Moskowitz, G.B., Li, P., Ignarri, C., & Stone, J. (2011). Compensatory cognition associated with egalitarian goals. *Journal of Experimental Social Psychology*, *47*, 365–370.

Murphy, M.C., Richeson, J.A., & Molden, D.C. (2011). Leveraging motivational mindsets to foster positive interracial interactions. *Social and Personality Psychology Compass*, *5*, 118–131.

Nosek, B.A., Banaji, M.R., & Greenwald, A.G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dynamics*, *6*, 101–115.

Penner, L.A. (1971). Interpersonal attraction toward a black person as a function of value importance. *Personality: An International Journal*, *2*, 175–187.

Perry, S.P., Dovidio, J.F., Murphy, M.C., & van Ryn, M. (2015). The joint effect of bias awareness and self-reported prejudice on intergroup anxiety and intentions for intergroup contact. *Cultural Diversity & Ethnic Minority Psychology*, *21*, 89–96.

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

Plant, E.A., & Devine, P.G. (2001). Responses to other-imposed pro-black pressure: Acceptance or backlash? *Journal of Experimental Social Psychology*, *37*, 486–501.

Richeson, J.A., & Shelton, J.N. (2003). When prejudice does not pay: Effects of interracial contact on executive function. *Psychological Science*, *14*, 287–290.

Richeson, J.A., & Trawalter, S. (2005). Why do interracial interactions impair executive function? A resource depletion account. *Journal of Personality and Social Psychology*, *88*, 934–947.

Rogers, T.B., Kuiper, N.A., & Kirker, W.S. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology*, *35*, 677–688.

Rokeach, M., & Cochrane, R. (1972). Self confrontation and confrontation with another as determinants of long term value change. *Journal of Applied Social Psychology*, *2*, 283–292.

Shelton, J.N., Richeson, J.A., Salvatore, J., & Trawalter, S. (2005). Ironic effects of racial bias during interracial interactions. *Psychological Science*, *16*, 397–402.

Sherman, D.K., & Cohen, G.L. (2002). Accepting threatening information: Self-affirmation and the reduction of defensive biases. *Current Directions in Psychological Science*, *11*, 119–123.

Snyder, M. (1974). Self-monitoring of expressive behavior. *Journal of Personality and Social Psychology*, *30*, 526–537.

Snyder, M., & Ickes, W. (1985). Personality and social behavior. *Handbook of social psychology*, *2*, 883–947.

Sommers, S.R., & Norton, M.I. (2006). Lay theories about White racists: What constitutes racism (and what doesn't). *Group Processes & Intergroup Relations*, *9*, 117–138.

Son Hing, L.S., Li, W., & Zanna, M.P. (2002). Inducing hypocrisy to reduce prejudicial responses among aversive racists. *Journal of Experimental Social Psychology*, *38*, 71–78.

Stephan, W.G., & Stephan, C.W. (2001). *Improving intergroup relations*. London: Sage Publications, Inc.

Swann, W.B., Jr., Chang-Schneider, C., & McClarty, K.L. (2007). Do people's self-views matter. *American Psychologist*, *62*, 84–94.

Tabachnick, B.G., & Fidell, L.S. (2000). *Using multivariate statistics* (4th ed.). Boston: Allyn & Bacon.

Trawalter, S., & Richeson, J.A. (2006). Regulatory focus and executive function after interracial interactions. *Journal of Experimental Social Psychology*, *42*, 406–412.

Vorauer, J.D. (2012). Completing the implicit association test reduces positive intergroup interaction behavior. *Psychological Science*, *23*, 1168–1175.

Vorauer, J.D., Gagnon, A., & Sasaki, S.J. (2009). Salient intergroup ideology and intergroup interaction. *Psychological Science*, *20*, 838–845.

Wang, K., Stroebe, K., & Dovidio, J.F. (2012). Stigma consciousness and prejudice ambiguity: Can it be adaptive to perceive the world as biased? *Personality and Individual Differences*, *53*, 241–245.

Watson, D., Clark, L.A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect scale: The PANAS scale. *Journal of Personality and Social Psychology*, *6*, 1063–1070.

West, T.V., & Kenny, D.A. (2011). The truth and bias model of judgment. *Psychological Review*, *118*, 357–378.