The Influence of Facial Feedback on Race Bias

Tiffany A. Ito,1 Krystal W. Chiao,2 Patricia G. Devine,3 Tyler S. Lorig,4 and John T. Cacioppo2

1University of Colorado, 2University of Chicago, 3University of Wisconsin-Madison, and 4Washington and Lee University

ABSTRACT—Two studies were conducted to examine whether facial feedback can modulate implicit racial bias as assessed by the Implicit Association Test (IAT). Participants were surreptitiously induced to smile through holding a pencil in their mouth while viewing photographs of unfamiliar Black or White males or performed no somatic configuration while viewing the photographs (Study 1 only). All participants then completed the IAT with no facial manipulation. Results revealed a spreading attitude effect, with significantly less racial bias against Blacks among participants surreptitiously induced to smile during prior viewing of Black faces than among participants surreptitiously induced to smile during prior viewing of White faces.

Procedures that change the cognitive associations activated by a group exemplar or label demonstrate that even easily activated implicit associations can be modified (e.g., Blair & Banaji, 1996; Kawakami, Moll, Hermsen, Dovidio, & Russin, 2000; Macrae, Bodenhausen, & Milne, 1995; Wittenbrink, Judd, & Park, 2001). This finding has been valuable for understanding implicit associations, but because evaluative responses also affect group perceptions, it is also important to examine how evaluations can be moderated. Recent research on conditioning and what has been called the spreading attitude effect suggests a potentially powerful method for changing the evaluations associated with a group by subtly changing the evaluative associations of individual exemplars. The spreading attitude effect refers to an indirect conditioning process. Walther (2002) demonstrated it by first associating two evaluatively neutral, unfamiliar faces through repeated, paired presentation. In the subsequent evaluative conditioning phase, one of the faces from the initial preassociation phase became a conditioned stimulus and was repeatedly paired with a face of someone who was either liked or disliked (the unconditioned stimulus). A conditioning effect was obtained, with the previously neutral face acquiring the evaluative quality of the unconditioned stimulus with which it was paired.

Of greater interest are evaluations of the preassociate, that is, the neutral face that was previously associated with the face that went on to become the conditioned stimulus. Note that the preassociate was itself never directly paired with the unconditioned stimulus, but it nevertheless acquired the same evaluative tone. For instance, liking increased for a neutral face that was paired with a face subsequently paired with a positive unconditioned stimulus. In this way, the evaluative conditioning effect spread to a stimulus never directly associated with the unconditioned stimulus (see also Hammerl & Grabitz, 1996).

Although Walther (2002) demonstrated the spreading attitude effect for evaluations of individuals, this mechanism may provide a means for changing evaluations of an entire social group. A particularly attractive feature of the effect is that the type of conditioning it involves—evaluative conditioning—occurs without conscious awareness of the contingency between the conditioned and unconditioned stimuli, holding promise that relatively subtle manipulations of evaluation could be effective in changing evaluations of a group. This may be important in avoiding the increases in bias that are typically associated with more powerful manipulations of positive affect. That is, when manipulations of positive affect succeed in changing subjective affective states, increased reliance on heuristics, including stereotypes, is typically seen (Bodenhausen, Kramer, & Susser, 1994; Stroessner & Mackie, 1992). This is thought to occur because positive affective states signal satisfaction with and safety in the environment, while also creating a desire to maintain the positive state, all of which should decrease careful or effortful information processing (Bodenhausen et al., 1994; Lambert, Khan, Lickel, & Fricke, 1997). These past studies suggest that if evaluative associations with group members are to be improved, this may need to be accomplished in ways that do not elevate subjective affective state.
Drawing on research demonstrating the spreading attitude effect, the two studies reported here examined the effects of indirect evaluative conditioning on reactions to social groups. In addition to extending Walther’s (2002) findings from evaluations of individuals to evaluations of a group, we were interested in assessing whether subtle changes in bodily configuration would be sufficient to produce the evaluative changes. This interest derived from research showing that motor processes related to affect and emotion can moderate evaluative reactions. Cacioppo, Priester, and Berntson (1993) demonstrated that arm flexion, associated with approach–positive motivational states, and arm extension, associated with withdrawal–negative motivational states, can surreptitiously activate different implicit evaluative processes. Participants who were exposed to one group of neutral attitude objects (Chinese ideographs) during arm flexion and another group of ideographs during arm extension produced a subsequent preference for the ideographs associated with arm flexion. A similar outcome has been obtained with changes in facial configuration. Participants instructed to hold a pen between their teeth in a manner that contracted the zygomaticus major muscles typically involved in smiling rated cartoons as more humorous than participants who relaxed these muscles (Strack, Martin, & Stepper, 1988). These changes are thought to occur in the absence of changes in subjective affective state, raising the possibility that such manipulations could be used to produce more positive evaluative reactions toward members of social groups.

To increase positive evaluative associations with individual members of a social group, we surreptitiously induced participants to smile (following the method of Strack et al., 1988) while they viewed either primarily Black or primarily White faces. Other participants viewed the faces without any special somatic manipulation (Study 1 only). In a separate phase, we measured implicit racial bias with the Implicit Association Test (IAT) and explicit bias with the Attitudes Toward Blacks Scale (ATB; Brigham, 1993). Both implicit and explicit measures were included so we could examine whether effects of the subtle evaluative manipulation differed as a function of level of measurement. Studies finding a deleterious effect of positive affective states on bias have tended to measure relatively deliberative responses (e.g., judgments of guilt or similarity among group members; Bodenhausen et al., 1994; Stroessner & Mackie, 1992). By contrast, the types of judgments on which somatic influences have been obtained (e.g., ratings of funniness) may be considered less deliberative (Strack et al., 1983). This pattern of findings is consistent with models of attitude change specifying that implicit evaluative processes have a bigger impact under conditions of low rather than high elaboration and reflection (Petty & Cacioppo, 1981, 1986).

The two studies we conducted used similar procedures and obtained similar outcomes, so they are described together. The main difference between the studies was whether a vigilance task was implemented during the somatic manipulation to explicitly direct attention toward the faces of Blacks and Whites that were being viewed. This task was omitted in Study 2 so we could examine whether explicit attention to the stimuli paired with the positive evaluative state was needed to improve reactions toward Blacks.

**METHOD**

**Participants and Overview**

Thirty-three (6 Asian American, 2 Hispanic American, 22 Caucasian American, 3 other) and 40 (3 Asian American, 4 Hispanic American, 29 Caucasian American, 4 other) undergraduates from the University of Chicago participated in Studies 1 and 2, respectively. Participants who identified themselves as African American (3 in Study 1 and 2 in Study 2) took part in the study but were not included in the data analyses.1 All participants were tested individually and completed three main tasks: (a) viewing of matrices of Black and White faces with or without the facial feedback manipulation, (b) the IAT, and (c) the ATB.

**Materials**

Forty-four monochromatic pictures of Black and White males were cropped to show just the eyes and nose. Thirty-two of each of these sets were used in the face-matrix task. Each individual face was sized to $341 \times 256$ pixels and placed in a matrix of faces. Participants indicated the location of a detected dot by pressing one of nine keys on the numeric keypad (e.g., pressing the “1” for a dot appearing on the lower left face). Dots in Study 1, participants were told they would complete a vigilance task by searching for a red dot presented among a 3 × 3 matrix of faces. Participants indicated the location of a detected dot by pressing one of nine keys on the numeric keypad (e.g., pressing the “1” for a dot appearing on the lower left face). Dots

---

1 In addition, data from 5 participants were excluded from Study 2 because they made errors on more than 22% of all IAT trials ($n = 3$) or did not change the position of the pencil in their mouths between blocks of trials as instructed ($n = 2$).
appeared on half of the matrices, near the center of one of the faces. The position of the dot was randomly determined for each trial with the constraint that it never appeared on the center face. Each matrix was presented for 2,500 ms, and successive presentations were separated by a 350-ms intertrial interval. To ensure that any moderating effects of the somatic manipulation did not depend on carefully attending to the racial stimuli, we designed Study 2 as a conceptual replication that omitted the instructions for the vigilance task. Consequently, the dots were omitted from the face matrices, and participants were simply asked to look at the faces. Participants in both studies completed two blocks of face-matrix trials consisting of 24 matrices each. The two blocks differed in whether the majority of the faces in the matrices were Black or White, with block order counterbalanced across participants.

The somatic manipulation occurred during the viewing of the face matrices. In Study 1, 10 participants were induced to contract the muscles involved with smiling (zygomaticus major) during the Black-majority block and relax those muscles during the White-majority block. We refer to this condition as smiling-Black. Twelve other participants contracted the muscles involved in smiling during the White-majority block and relaxed those muscles during the Black-majority block. We refer to this condition as smiling-White. The remaining 11 participants, assigned to the control condition, performed the vigilance task without performing any somatic manipulations. The control condition was omitted in Study 2, with all participants induced to perform the smiling manipulation during viewing of either the Black-majority (n = 20) or the White-majority (n = 20) block.

Following Strack et al. (1988), the somatic configuration associated with smiling was achieved by asking participants to hold a pencil in their teeth such that the point was directed away from them. They were told to avoid touching their lips to the pencil, thus forming a smile. Note, however, that a smile was never mentioned, and instructions were given only in terms of the orientation of the pencil and action of the lips and teeth. The pencil was held in this manner throughout the assigned 24-matrix block. In the nonsmiling somatic configuration in Study 1, participants were instructed to hold the pencil horizontally between their teeth such that the point was directed to their right or left, in the horizontal plane, and to clench the pencil lightly with their teeth and lips. In Study 2, the nonsmiling manipulation was changed slightly to directly replicate the procedure of Strack et al., with participants holding the pencil such that the point was directed away from them. The pencil was to be held tightly with their lips, making no contact with their teeth.

The two face-matrix blocks were separated by completion of a questionnaire packet consisting of the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988) embedded in 40 filler items. Participants used a 5-point scale to indicate the extent to which they were experiencing each affective state “right now (that is, at this moment).” Participants in a somatic-configuration condition were asked to remove the pencil from their mouth and relax during completion of this survey. Participants took approximately 5 min to complete the survey, which provided a break between the activation periods of different facial muscle groups.

Following the face-matrix blocks, all participants performed a 2-min distractor task consisting of circling Es in a scrambled field of letters.

Measurement of Implicit Bias

After the distractor task, implicit racial bias was measured with the IAT, described to participants as a key-pressing task. The task consisted of classifying words as positive or negative and faces as Black or White, and involved five blocks of trials following the method of Greenwald et al. (1998): (a) face-only social categorization; (b) word-only evaluation; (c) critical trials combining the two tasks, using the same response mappings as the first two blocks; (d) practice trials with the reversed response mapping for evaluation; and (e) critical trials for the reversed combined task. In the IAT, racial bias is indicated if participants respond more quickly when the response keys used to indicate “White” and “Black” are the same as the keys used to indicate “positive” and “negative,” respectively, as compared with when “White” is paired with “negative” and “Black” is paired with “positive.” These blocks of trials are referred to as compatible and incompatible, respectively. Order for the face-only and word-only blocks and for the compatible and incompatible blocks varied across participants. Faces used in the IAT differed from those used during facial feedback manipulation.

Self-Reported Bias

Following the IAT, participants completed the ATB, then were probed for suspicion regarding the experimental hypotheses and debriefed. No participants revealed any knowledge about the actual purpose of the pencil manipulations.

RESULTS AND DISCUSSION

Data in Study 1 were analyzed with a 3 (smile configuration: smiling-Black, smiling-White, no somatic manipulation) × 2 (order of face-matrix blocks: Black-majority matrices first, White-majority matrices first) analysis of variance. In Study 2, the smile-configuration variable had only two levels (smiling-Black, smiling-White).
Self-Reported Mood
Because the PANAS was completed between the two face-matrix blocks, it provides a between-participants assessment of the initial somatic-configuration manipulation each participant experienced. As we expected, affective state was not elevated by the somatic configuration associated with smiling. There were no effects of smile configuration (or block order) on level of positive or negative affect in Study 1 (all Fs < 1.). Although positive affect was affected by the somatic manipulation in Study 2, $F(1, 36) = 6.24, \eta^2_p = .15, p < .05$, participants who performed the somatic configuration associated with smiling experienced lower positivity than those who performed the nonsmiling configuration ($M$s = 2.04 and 2.35).

Implicit Prejudice
For IAT data, we calculated D scores, according to the recommendations of Greenwald, Nosek, and Banaji (2003). This score represents the difference in mean response latencies to the incompatible and compatible blocks, divided by the pooled standard deviation across blocks. As in much IAT research, D was significantly different from 0 in both studies, $M = 0.29$, $t(32) = 5.74, d = 1.00, p < .001$, and $M = 0.33, t(39) = 4.93, d = 0.78, p < .001$. These results indicate implicit preference for Whites as compared with Blacks in both studies.

Next, we examined IAT D scores for moderation by the somatic manipulation. The only significant effects were main effects of smile configuration in both studies, $F(2, 27) = 4.56, \eta^2_p = .25$, and $F(1, 39) = 6.03, \eta^2_p = .14, ps < .05$. Mean IAT D scores for the two studies are shown in Table 1. The effect of smile configuration was consistent with our prediction that evaluations of a group can change by prior positive associations with individual exemplars: Bias was significantly reduced in the smiling-Black condition as compared with the smiling-White condition. In Study 1, which included the somatic-manipulation control condition, bias was significantly lower in the smiling-Black condition relative to both the smiling-White and the control conditions, $t(20) = 3.50, d = 1.50, p < .005$, and $t(19) = 2.15, d = 0.94, p < .05$; the latter two conditions did not differ, $t < 1$.

Explicit Prejudice
The results were consistent with our hypothesis that the somatic manipulation would have bigger effects on implicit than explicit bias, as the somatic manipulation had no effect on ATB scores in either study, $F$s < 1. This dissociation of implicit and explicit effects is consistent with prior research on somatic influences (Cacioppo et al., 1993, Strack et al., 1988).³

### Table 1

Mean D Score on the Implicit Association Test as a Function of Smile Configuration

<table>
<thead>
<tr>
<th>Study</th>
<th>Smiling-Black</th>
<th>Smiling-White</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study 1</td>
<td>0.09ᵃ</td>
<td>0.40ᵇ</td>
<td>0.037ᵇ</td>
</tr>
<tr>
<td>Study 2</td>
<td>0.17ᵃ</td>
<td>0.49ᵇ</td>
<td>—</td>
</tr>
</tbody>
</table>

Note. Within a row, means with different subscripts differ at $p < .05$.

GENERAL DISCUSSION

The present research demonstrates that repeatedly viewing Black faces while being surreptitiously induced to smile diminishes implicit racial bias. As in past research examining somatic influences on attitudes (Cacioppo et al., 1993), changes in implicit bias were obtained in the absence of changes in explicit bias, a dissociation consistent with prior conclusions about the important role of nondeclarative information in attitudes (Cacioppo et al., 1993; Greenwald & Banaji, 1995; Wilson, Lindsey, & Schoorler, 2001). When coupled with the absence of improvements in subjective mood or awareness of the relation between the facial manipulation and measurement of bias, this dissociation suggests that the subtle changes in motor processes influenced racial attitudes via links with nondeclarative information.

Because bias was measured when no somatic configuration was being performed, and using exemplars different from the ones viewed during the somatic configuration, the results go beyond merely demonstrating that affectively laden motor processes influence current evaluations. In accord with the spreading attitude effect, implicit evaluative responses were generalized beyond the exemplars seen in the face matrices to their associated social category. This effect is thought to depend on unreinforced associations among stimuli before the conditioning phase (sensory preconditioning; Hammerl & Grabitz, 1996; Walther, 2002). In the present case, the prior association of individual exemplars with a social group could provide the initial sensory preconditioning. When some of the exemplars undergo conditioning and acquire the evaluative tone of the unconditioned stimulus, that evaluation spreads to other exemplars linked through initial sensory preconditioning.

The spreading attitude effect suggests a potentially powerful way to extend past research on the moderation of implicit bias. Consider the research by Dasgupta and Greenwald (2001), which showed that initial exposure to liked Blacks (e.g., Denzel Washington) and disliked Whites (e.g., Jeffrey Dahmer) decreased subsequently measured bias against Blacks. As in the present studies, the exemplars involved in the assessment of implicit bias differed from the liked Blacks and disliked Whites to whom participants were initially exposed. In addition, exposure to liked Blacks and disliked Whites had no effect on explicit prejudice.

³Additional analyses examining whether ATB scores moderated the effects of smiling configuration on IAT scores were not significant in either study.
By using an evaluative conditioning approach, the present studies build on Dasgupta and Greenwald’s (2001) work in several ways. First, in Dasgupta and Greenwald’s study, the effectiveness of exposure to liked Blacks and disliked Whites presumably depended on participants’ prior familiarity with the individuals. Their presentation then either changed the salience of exemplars brought to mind during the subsequent prejudice measurement or changed the actual stored abstract representation of the groups. In the present studies, none of the exemplars in the face-matrix blocks were known to participants, so the beneficial effect of surreptitiously induced smiling was based on new associations to previously unknown individuals. Second, the intensity of evaluative associations to previously known positive and negative exemplars is likely stronger than the intensity of the association between surreptitiously induced smiling and the unknown exemplars in the present studies. Nevertheless, the latter was sufficient to moderate racial bias. Finally, the decrease in implicit prejudice that Dasgupta and Greenwald demonstrated occurred after exposure to both liked Blacks and disliked Whites. The present studies involved only a somatic manipulation associated with a positive affective state, and it was paired with only one racial group. The results show that increasing the positivity of associations with Blacks is sufficient to moderate implicit bias.

More generally, evaluative conditioning occurs quickly, with relatively few pairings of a conditioned and unconditioned stimulus, and does not require conscious awareness of the contingency between the two. Other procedures intended to change explicit associations with a group, such as practice denying a stereotypic association (Kawakami et al., 2000) or the internalization of low-prejudice standards of behavior (e.g., Devine, 1989; Devine & Monteith, 1999), can affect bias, but the procedures used in the studies reported here demonstrate that changes in evaluative associations achieved in relatively subtle ways can also moderate bias. Questions do remain about the use of evaluative conditioning and the spreading attitude effect as a means to change racial bias. One issue not yet addressed is the temporal stability of the change. Evaluative conditioning in general and the spreading attitude effect in particular are known to be resistant to extinction (Baeyens, Crombez, Van den Bergh, & Eelen, 1988; Walther, 2002), but this issue has not been directly addressed in the case of conditioning of responses toward social groups. Nevertheless, both the general mechanism of evaluative conditioning and the specific manipulation of surreptitious changes in somatic configuration appear to hold promise in moderating racial attitudes.

Acknowledgments—This research was supported by National Science Foundation Grant No. BCS-0086314.

REFERENCES


(Received 1/19/05; Revision accepted 3/24/05; Final materials received 4/5/05)