

Research Report

Ambiguity in Social Categorization

The Role of Prejudice and Facial Affect in Race Categorization

Kurt Hugenberg and Galen V. Bodenhausen

Northwestern University

ABSTRACT—Two studies tested the hypothesis that perceivers' prejudice and targets' facial expressions bias race categorization in stereotypic directions. Specifically, we hypothesized that racial prejudice would be more strongly associated with a tendency to categorize hostile (but not happy) racially ambiguous faces as African American. We obtained support for this hypothesis using both a speeded dichotomous categorization task (Studies 1 and 2) and a rating-scale task (Study 2). Implicit prejudice (but not explicit prejudice) was related to increased sensitivity to the targets' facial expressions, regardless of whether prejudice was measured after (Study 1) or before (Study 2) the race categorizations were made.

Beginning with Darwin's (1872/1965) seminal work, scholars have made an increasingly compelling argument that the human face has been shaped by selection pressures to serve as a medium of communication. The complex upper layer of facial musculature and its associated neural structures seem to communicate internal emotional states effortlessly, sometimes more clearly than one might wish (see Dimberg & Öhman, 1996). However, the vast and nuanced array of possible facial expressions may communicate more than just the target's emotional state. We argue that those selfsame expressions may also influence the social category to which a target is assigned.

How a person is categorized is a matter of considerable consequence. If assigned to a negatively stereotyped category, an individual may be subjected to discrimination in a variety of subtle or blatant manifestations (Fiske, 1998; Macrae & Bodenhausen, 2000). Most research examining the effects of categorization has focused on cases in which category membership is unambiguous and category assignments can be made easily, even automatically and potentially outside of awareness. However, not all cases of social categorization are quite so clear-cut. Some stigmatized social identities can be

concealed (e.g., homosexuality), and some guesswork is thus required in establishing category membership (Ambady, Hallahan, & Conner, 1999). Even visually marked categories, such as race, are not always unambiguous. Variations in skin tone, hair color, and eye color are not perfect markers of race; in such ambiguous cases, we contend that a target's facial expression may be used to disambiguate the social category in question. Specifically, we argue that facial expressions stereotypic of a particular social category will be used as cues for assigning a target to that category. For example, the stereotype of African Americans includes hostility (Devine, 1989). If our hypothesis is correct, a target with an ambiguous racial identity would be more likely to be categorized as an African American when that target displays a hostile facial expression than when it displays a happy one.

We further hypothesize that the extent to which stereotypes will affect categorization in ambiguous situations will be moderated by the perceiver's level of prejudice. As high-prejudice individuals are more likely than low-prejudice individuals to activate and apply stereotypic information (Lepore & Brown, 1997; Wittenbrink, Judd, & Park, 1997), this bias in categorization should be strongest among people who are relatively high in prejudice. Thus, high-prejudice perceivers should be most likely to categorize angry, racially ambiguous targets as African American. Previous research, however, indicates that measures of implicit prejudice, based on differences in reaction times to attitude-relevant stimuli, may be better than explicit measures in capturing the aspects of prejudiced attitudes that are most relevant in the rapid parsing of nonverbal displays (e.g., Hugenberg & Bodenhausen, 2003). Therefore, this categorization bias may be evident only when prejudice is indexed using an implicit measure.

STUDY 1

Study 1 was designed to test the hypothesis that implicit prejudice interacts with a target's facial expression in determining the race to which a target is assigned. We employed a speeded dichotomous decision task in which individuals were presented a series of racially ambiguous computer-generated faces and were asked to categorize each target as either Caucasian or African American. Each of the 15 faces was

Address correspondence to Kurt Hugenberg, Department of Psychology, 110F Benton Hall, Miami University, Oxford, OH 45056; e-mail: hugenbk@muohio.edu.

presented twice: once with a clearly happy facial expression and once with a clearly angry facial expression. Participants then completed measures of their explicit attitudes toward Caucasians and African Americans and finally completed an implicit association task (IAT; see Greenwald, McGhee, & Schwartz, 1998) designed to measure implicit racial attitudes. We predicted that as prejudice increased, so would the tendency to categorize angry faces as African American. However, prejudice was expected to be unrelated to categorizations of happy faces.

Method

Participants and Design

Twenty European American undergraduates (8 female) participated in this study. Both implicit and explicit measures of prejudice served as predictors of categorization of happy and angry targets. Expression of the target was manipulated on a within-subjects basis.

Materials and Procedure

Stimuli consisted of 15 male faces generated using the Poser 4™ three-dimensional character animation software, which permitted the construction of racially ambiguous targets by manipulation of the targets' facial structure, skin tone, and hair style and color. Each of these faces was then further manipulated to create two versions, one with a clearly happy facial expression and the other with a clearly hostile expression. The software allowed for manipulation of expression without changing the facial physiognomy of the targets. Pretesting indicated that the 15 target faces were difficult to categorize by race when emotionally neutral and that each of the happy and angry expressions was clearly identifiable as such. Each picture measured 12 × 12 cm and displayed the target's full face, from the front, as well as hair and neck regions.

After giving informed consent, participants were seated at computers in individual cubicles and were instructed that they would see a series of faces that were to be categorized along a randomly determined dimension (e.g., emotion, gender, race). The computer then ostensibly randomly determined the dimension along which the targets were to be categorized. In all cases, the computer instructed participants to categorize the subsequent targets by race, as either "Caucasian" or "African American." Participants were further instructed to make the categorizations as quickly as possible. After performing two practice trials, participants engaged in the race categorization task, which included all 30 targets, with stimulus order randomized for each participant. Following the dichotomous decision task, participants were presented with "feeling thermometers" measuring explicit attitudes about six different social groups, including Caucasians and African Americans. Participants indicated how warmly or coldly they felt about each group on a scale from 1 to 100, with higher responses indicating more warmth.

Finally, participants performed the IAT, which was described as an ostensibly unrelated word categorization task. The IAT consisted of five trial blocks. The first two blocks were practice blocks in which participants learned to map White names to one response key and Black names to another (the first block) and to map pleasant and unpleasant words to those same two response keys (the second block). Each of these practice blocks contained only names (the first block) or pleasant and unpleasant words (the second block). The selected names and words were taken from Greenwald et al. (1998). A third block

consisted of "compatible" trials, on which White names and pleasant words were mapped to the same response key and Black names and unpleasant words were mapped to another key. Thus, this third block involved both names and valenced words. A fourth block involved learning a new mapping for the pleasant and unpleasant words. The fifth and final block consisted of "incompatible" trials, on which White names and unpleasant words were mapped to the same key and Black names and pleasant words were mapped to another key. Thus, this fifth block again involved both names and valenced words. On this task, implicit prejudice is indicated by the extent to which performance on the incompatible trials (i.e., Black-good/White-bad) is impaired, relative to performance on the compatible trials (i.e., Black-bad/White-good). After completing all tasks, participants were debriefed.

Results and Discussion

The main dependent measure was the frequency with which participants categorized the targets as African American. We hypothesized that high-prejudice European Americans would categorize more of the ambiguous faces as African American than low-prejudice European Americans would, but only when the faces were angry.

To test this hypothesis, we employed a multiple regression analysis wherein centered values of implicit prejudice,¹ explicit prejudice ($M = .10$, $SD = 7.99$),² and their interaction term were used as predictors of frequency of categorization as African American for both happy and angry faces, with facial affect as a within-subjects factor (i.e., treating the difference in categorization of angry and happy faces as African American as the criterion variable). This analysis revealed the hypothesized Implicit Prejudice × Target Expression interaction, standardized $\beta = .46$, $F(1, 16) = 4.56$, $p = .049$ (see Fig. 1). No other effects approached statistical significance, all $ps > .25$. This interaction confirms that, as expected, the relationship between prejudice and categorization was stronger for angry than for happy faces. Follow-up simple slope tests revealed that implicit-prejudice scores were marginally positively related to the frequency of categorization of angry faces as African American, standardized $\beta = .35$, $F(1, 18) = 2.44$, $p = .068$ (one-tailed), but the relationship between implicit prejudice and categorization was much weaker for happy faces, standardized $\beta = .11$, $F(1, 18) = 0.21$, $p = .65$.

STUDY 2

The first study indicated that individuals high in implicit (but not explicit) prejudice are likely to categorize racially ambiguous target faces as African American when the faces are angry but not when they are happy. Thus, it seems that when facial expressions are congruent with the stereotypic content of a racial category (hostility, in the case of the African American stereotype), individuals relatively high in implicit prejudice are disproportionately likely to categorize those

¹Implicit prejudice was computed separately for each participant by subtracting the participant's mean latency in milliseconds on compatible trials ($M = 331$, $SD = 144$) from his or her mean latency on incompatible trials ($M = 1,059$, $SD = 213$), following the conventions used by Greenwald et al. (1998).

²The IAT is a relative measure of prejudice. In order to make the explicit measure of prejudice analogous, we subtracted the feeling-thermometer score for African Americans from that for Caucasians to obtain a relative measure of explicit prejudice, with higher scores indicating relative preference for Caucasians. (Analyses performed with the absolute feeling-thermometer scores for African Americans yielded results quite similar to those presented.)

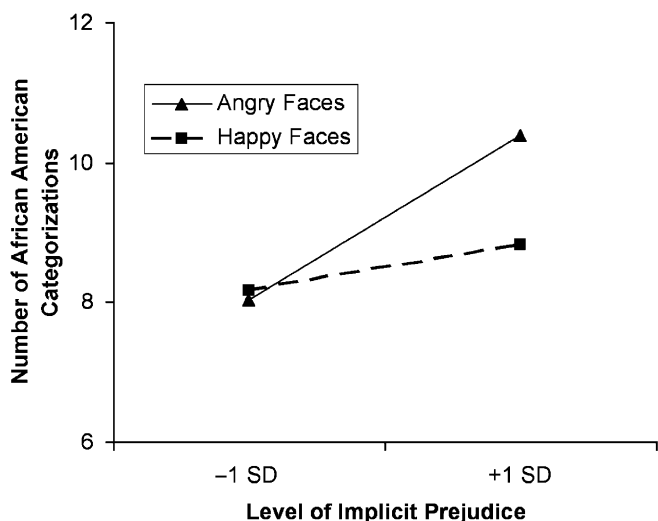


Fig. 1. Results of the regression analysis in Study 1: frequency of categorization of both happy ($SE = 0.69$) and angry ($SE = 0.76$) targets as African American as a function of level of implicit prejudice. Plotted values of implicit prejudice are at 1 *SD* above and below the mean.

faces as belonging to that category. We undertook a second study to replicate and extend these findings. In this instance, the measures of prejudice were collected in a separate experimental session, and two types of categorization judgments were obtained: a speeded dichotomous judgment (replicating Experiment 1) and a nonspeeded judgment collected on a continuous rating scale. Except for these changes, Study 2 employed a method virtually identical to that of Study 1.

Method

Participants and Design

Fifty-seven European American undergraduates (26 female) participated in the study. The design mirrored that of Study 1.

Materials and Procedure

The materials and procedure were identical to those used in Study 1 except as follows. The study included two separate experimental sessions. Feeling-thermometer and IAT measures of prejudice were assessed in Session 1, using the same measures as in Study 1. A minimum of 1 week later, participants returned for a second session that contained both a speeded dichotomous categorization task, exactly as in Study 1, and a nonspeeded categorization task, in that order. The second categorization task was quite similar to the dichotomous categorization task in that participants were presented with the same 30 faces and were asked to categorize each of the faces on a scale from 1 to 7, with 1 indicating *definitely Caucasian* and 7 indicating *definitely African American*. Unlike in the dichotomous categorization task, participants were instructed to take as much time as needed to make accurate categorizations.

Results and Discussion

Dichotomous Categorization Task

A multiple regression analysis analogous to that employed in Study 1 was used to test the hypothesis that racially ambiguous faces

displaying hostility (as compared with those displaying happiness) would be categorized more frequently as African American as implicit prejudice increased. The dependent measure of interest was again the frequency with which faces were categorized as African American versus Caucasian. This analysis yielded the predicted Implicit Prejudice \times Target Expression interaction, standardized $\beta = .38$, $F(1, 53) = 8.55$, $p = .005$ (see Fig. 2, top panel). Again, neither explicit prejudice nor its interaction with implicit prejudice was reliably associated with categorization, $ps > .3$.

As Studies 1 and 2 used exactly the same dichotomous categorization task and found the same pattern of results, we used meta-analytic procedures to combine the effects from the two studies to create a more powerful test of the relationship between prejudice and categorization of happy and angry faces. As predicted, these analyses revealed that implicit-prejudice scores were positively related to categorization of target faces as African American when the faces were angry, $r(75) = .21$, $p = .03$ (one-tailed). No relationship was found between implicit prejudice and categorization for happy faces, $r(75) = .07$, $p = .53$.

Rating-Scale Categorization Task

The dependent measure of interest for the rating-scale categorization task was the mean rating of the race of the targets, for both angry and

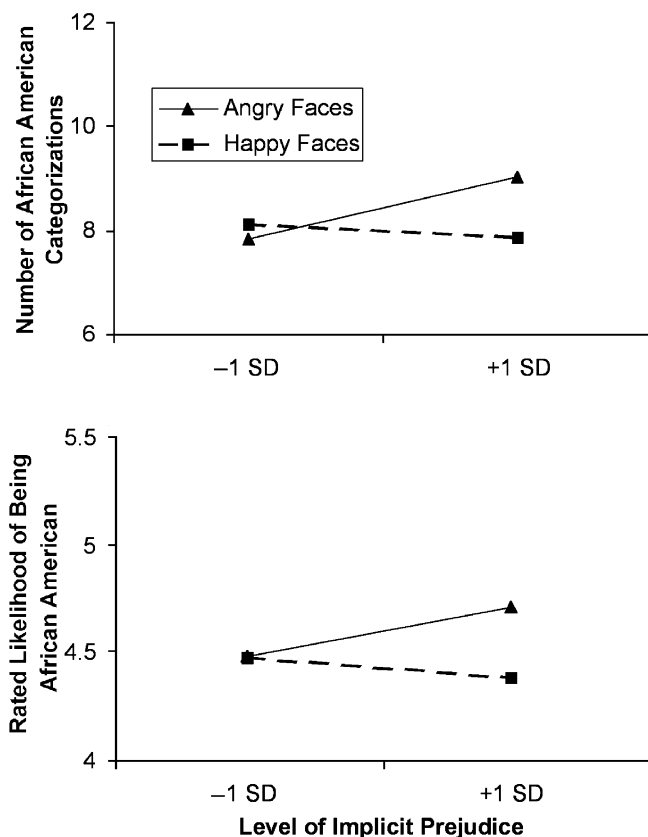


Fig. 2. Results of the regression analysis in Study 2: frequency of categorization (top panel) and mean ratings (bottom panel) of both happy ($SE_{\text{categorization}} = 0.47$, $SE_{\text{rating}} = 0.13$) and angry ($SE_{\text{categorization}} = 0.48$, $SE_{\text{rating}} = 0.14$) targets as African American as a function of level of implicit prejudice. Plotted values of implicit prejudice are at 1 *SD* above and below the mean.

happy faces. A multiple regression analysis (using the same predictor variables as for the dichotomous categorization task) confirmed the predicted Implicit Prejudice \times Target Expression interaction, standardized $\beta = .32$, $F(1, 53) = 5.65$, $p = .02$ (see Fig. 2, bottom panel). No other effects approached statistical significance, all p s $> .35$. Inspection of the bottom panel in Figure 2 shows a pattern similar to that obtained with speeded categorizations, indicating that more deliberative judgments were subject to the same relationship: As implicit prejudice increased, categorization decisions were more powerfully influenced by targets' facial affect.

GENERAL DISCUSSION

Across both studies, the relationship between implicit prejudice and categorization of racially ambiguous faces was dependent on facial expression. When faces were seen to display relatively hostile expressions (stereotypic of African Americans), individuals high in prejudice tended to categorize them as African American. This relationship between prejudice and categorization was not evident for happy faces. Thus, it seems that in cases of uncertain categorization, the stereotypic biases of individuals relatively high in prejudice take hold and drive categorization.

It is interesting to consider the present findings in light of prominent theories of face processing. One such theory (Bruce & Young, 1986; see also Haxby, Hoffman, & Gobbini, 2000) proposes that two separate neural systems are involved in parsing faces. One system is dedicated to processing cross-situationally stable aspects of faces (e.g., identity, gender, race), and one is dedicated to processing dynamic aspects (e.g., momentary emotional displays). Recent findings indicate that these two neural systems interact in a bidirectional manner. The current studies suggest that ambiguity regarding stable aspects of target identity can be resolved by reference to dynamic aspects of the face, such as emotional expression. In previous research (Hugenberg & Bodenhausen, 2003), we showed that the converse relationship holds as well: Emotional ambiguity can be resolved by reference to stable aspects of a target's identity (i.e., race). Taken together, these and other recent findings suggest that perceptions of emotional displays and social categories may be intimately linked.

The present results have particular implications for the construal of biracial individuals. The susceptibility of such individuals to societal prejudice depends on whether or not they are categorized as members of a culturally stigmatized group. Previous research has linked the well-being of biracial persons to exposure to such prejudice (e.g., Jacobs, 1992). The current findings suggest that the risk of prejudice and discrimination is heightened whenever biracial individuals' affective displays are stereotype congruent. Of course, angry displays

are likely to result in problematic social interactions in general, but biracial individuals may be saddled with an additional burden: the activation of a host of otherwise dormant negative racial stereotypes that could influence interpersonal interactions in a way that makes a positive resolution of conflict more difficult to achieve (see Bodenhausen, Macrae, & Garst, 1998). Although the current results do suggest that categorization can be driven by stereotypic biases, more research is needed to examine the downstream consequences of these affective influences on racial categorization.

REFERENCES

- Ambady, N., Hallahan, M., & Conner, B. (1999). Accuracy of judgments of sexual orientation from thin slices of behavior. *Journal of Personality and Social Psychology*, *77*, 538–547.
- Bodenhausen, G.V., Macrae, C.N., & Garst, J. (1998). Stereotypes in thought and deed: Social-cognitive origins of intergroup discrimination. In C. Sedikides, J. Schopler, & C.A. Insko (Eds.), *Intergroup cognition and intergroup behavior* (pp. 311–335). Mahwah, NJ: Erlbaum.
- Bruce, V., & Young, A.W. (1986). Understanding face recognition. *British Journal of Psychology*, *77*, 305–327.
- Darwin, C. (1965). *The expression of emotions in man and animals*. Chicago: University of Chicago Press. (Original work published 1872)
- Devine, P.G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*, 5–18.
- Dimberg, U., & Öhman, A. (1996). Behold the wrath: Psychophysiological responses to facial stimuli. *Motivation and Emotion*, *20*, 149–182.
- Fiske, S.T. (1998). Stereotyping, prejudice, and discrimination. In D.T. Gilbert & S.T. Fiske (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 357–411). New York: McGraw-Hill.
- Greenwald, A.G., McGhee, D.E., & Schwartz, J.L.K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*, 1464–1480.
- Haxby, J.V., Hoffman, E.A., & Gobbini, M.I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*, 223–233.
- Hugenberg, K., & Bodenhausen, G.V. (2003). Facing prejudice: Implicit prejudice and the perception of facial threat. *Psychological Science*, *14*, 640–643.
- Jacobs, J.H. (1992). Identity development in biracial children. In M.P.P. Root (Ed.), *Racially mixed people in America* (pp. 190–206). Newbury Park, CA: Sage.
- Lepore, L., & Brown, R. (1997). Category and stereotype activation: Is prejudice inevitable? *Journal of Personality and Social Psychology*, *72*, 275–287.
- Macrae, C.N., & Bodenhausen, G.V. (2000). Social cognition: Thinking categorically about others. *Annual Review of Psychology*, *51*, 93–120.
- Wittenbrink, B., Judd, C.M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology*, *72*, 262–274.

(RECEIVED 2/18/03; ACCEPTED 3/28/03)