What Are We Really Priming? Cue-Based Versus Category-Based Processing of Facial Stimuli

Robert W. Livingston and Marilynn B. Brewer
The Ohio State University

Results from 5 experiments provide converging evidence that automatic evaluation of faces in sequential priming paradigms reflects affective responses to phenotypic features per se rather than evaluation of the racial categories to which the faces belong. Experiment 1 demonstrates that African American facial primes with racially prototypic physical features facilitate more automatic negative evaluations than do other Black faces that are unambiguously categorizable as African American but have less prototypic features. Experiments 2, 3, and 4 further support the hypothesis that these differences reflect direct affective responses to physical features rather than differential categorization. Experiment 5 shows that automatic responses to facial primes correlate with cue-based but not category-based explicit measures of prejudice. Overall, these results suggest the existence of 2 distinct types of prejudice.

During a trip to Scandinavia, one of the authors met a Norwegian acquaintance at a rustic seaside café for lunch. In the course of an otherwise enjoyable conversation, the acquaintance began to disclose rather unsavory sentiments toward a particular category of Mediterranean men, recounting in disturbing detail her utter repugnance for what she perceived to be the group's dispositional, cultural, and behavioral tendencies. In the wake of this vehement tirade, she paused, took a sip of wine, and, in a markedly different tone of voice, expressed a certain regret for her contempt toward this group, given the undeniable physical attraction she felt toward its male constituents!

The foregoing anecdote raises intriguing questions concerning the complex nature of intergroup perception and the extent to which racial attitudes may vary depending on whether evaluative responses are triggered by concrete perceptual features (cue based) or by abstract group representations (category based). Social categories such as "Mediterranean" may elicit divergent reactions depending on whether evaluation is based on the physical features of typical exemplars of the group or the stereotypic traits associated with the group as a whole. The current article sought to explore this dissociation between cue-based and category-based evaluation of racial groups and its potential implication for automatic paradigms using facial primes.

The tacit (and sometimes explicit) assumption of researchers investigating automatic prejudice has generally been that facial primes serve as symbolic equivalents to racial category labels, rendering the two types of primes functionally interchangeable. In fact, however, faces are ontologically ambiguous; they can serve as symbolic representations of abstract social categories (concepts) or as configurations of concrete physical features per se (percepts), independent of any higher order semantic meaning. Hence, one question that arises is whether automatic responses to faces necessarily reflect category-based evaluations of the abstract social groups that the faces represent (e.g., African Americans) or whether they reflect cue-based, affective responsiveness to concrete, physical features (e.g., dark skin, full lips).

There are both theoretical and empirical bases to suggest that automatic evaluation of faces is cue based rather than category based. First, according to Brewer's (1988) dual-process model of person perception, "identification," or simple recognition of perceptual features, occurs at the automatic level, whereas category-based person perception requires more controlled levels of pro-

Robert W. Livingston and Marilynn B. Brewer, Department of Psychology, The Ohio State University.

This research was supported in part by a National Science Foundation predoctoral fellowship awarded to Robert W. Livingston and a Russell Sage Foundation grant awarded to Marilynn B. Brewer. Portions of the data in the present article were reported in a "Hot Topic" presentation given at the 12th Annual Meeting of the American Psychological Society in Miami, Florida.

We extend thanks to Stephanie Butler, Nicole Meiring, Amy Mitrey, and Scott Charles for help in conducting the experiments and to Richard J. Shakarchi for his technical support. We also give thanks to John Cacioppo, Jon Krosnick, Rich Petty, John Skowronski, Bill von Hippel, and Tim Wilson for valuable comments on drafts of this article.

Correspondence concerning this article should be addressed to Robert W. Livingston, who is now at the Department of Psychology, University of Wisconsin, 1202 West Johnson Street, Madison, Wisconsin 53706. E-mail: rwlivingston@facstaff.wicsc.edu

Journal of Personality and Social Psychology, 2002, Vol. 82, No. 1, 5–18
 0022-3514/02/$5.00 DOI: 10.1037//0022-3514.82.1.5

cessing.[1] Because categorization requires controlled processing, automatic evaluation of facial primes likely reflects cue-based rather than category-based, or semantic, processing. Specifically, Brewer and Feinstein (1999) stated that "research indicating that racial cues produce automatic evaluation suggests that this affective response occurs at the initial stage of identification, prior to selection of category-based or person-based processing mode" (p. 266). Bargh (1997) also supported this position, maintaining that automatic evaluative processing reflects "direct associative connections . . . between stimulus feature representations and the affective response mechanism, bypassing the cognitive processing mechanism entirely" (p. 16). Further, Giner-Sorolla, Garcia, and Bargh (1999) argued on the basis of empirical evidence that "automatic evaluation of pictures represents a direct response to the picture stimulus, and not a response to the word or words that the stimulus calls to mind (p. 91)." Overall, then, these positions maintain that, at the automatic level, responses to faces reflect affective responsiveness to perceptual features per se rather than more abstract evaluation of the social category that the features represent.

Because physical traits and category membership have been largely confounded in previous research, it is unclear whether negative evaluative responses to Black facial primes, for example, reflect negative evaluation of the category "African Americans" (and its associated category stereotypes), or negative affective responses to African-type phenotypic features per se.[2] One way to empirically disentangle cue-based versus category-based processing is to use facial primes of the same racial category that vary on physical cues. Within the African American population, for example, there exists considerable diversity with respect to race-relevant cues such as skin color, facial features, hair texture, and so on. Such differences in physical "prototypicality" provide a means of varying physical features while holding racial category membership constant.[3] The following experiment sought to capitalize on this within-race diversity in physical cues as a means of separately examining the effects of perceptual features and category membership on automatic evaluation of facial primes.

Experiment 1

We used a priming paradigm essentially similar to that developed by Fazio, Jackson, Dunton, and Williams (1995). The basic design involves the pairing of Black or White facial primes with positively or negatively valenced target words. The underlying assumption is that reaction time to target words is moderated by attitude congruence between the prime and target. Racial prejudice is reflected in shorter response latencies to negative target words that follow African American compared with White facial primes and longer response latencies to positive target words that follow African American compared with White facial primes.

Although facial stimuli in the original Fazio et al. (1995) experiment were not prerated for racial features or other physical attributes, African American facial primes in the current study were selected systematically to vary on prototypicality of racial features. Thus, the priming stimulus set included an equal number of low-prototypic (LP) Black and high-prototypic (HP) Black faces matched with each other and with a set of White (W) faces. Moreover, to ensure that differential responses to facial stimuli reflect affective reactions to perceptual cues rather than semantic facilitation to category-based content, we used affectively loaded nouns that were semantically unrelated to racial categories as target words rather than stereotypic adjectives.

This selection of African American faces makes it possible to assess cue-based (LP Black vs. HP Black) as well as category-based (White vs. Black [LP Black + HP Black]) differences in response times to positive and negative words. If automatic evaluation of facial stimuli primarily reflects category-based evaluative responses rather than evaluative responses to physical cues, response times would be expected to vary as a function of racial category membership rather than physical features. More specifically, the between-race contrast (when both HP and LP Blacks are combined in comparison to Whites) would be greater than the within-race contrast (LP Black vs. HP Black) if evaluation is category based. However, if evaluation is cue based, the opposite should be true. That is, the within-race contrast should be greater than the between-race contrast.

Method

Participants. Fifty-nine introductory psychology students participated in this study in partial fulfillment of requirements for course credit. Three participants were dropped because of computer malfunctions, and 2 were omitted for expressing specific suspicions regarding our hypotheses. Additionally, all 5 African American participants were excluded. Hence, our final sample consisted of 49 participants (20 men, 29 women).

Stimulus materials. Facial stimuli consisted of 48 color photographic images. All photos were head shots of males taken against a common background. Before selecting the face primes to be used in the experiment, a large pool of photographs of White and African American males were

[1] Brewer (1988) is often (mis-)cited as arguing that racial groups are automatically categorized. However, Brewer's (1988) model clearly posits that categorization is a controlled process (see Brewer's Figure 1.1, p. 5). Perhaps one source of confusion surrounds the meaning of the term "categorization." Technically speaking, "identification" (or "classification") is a type of categorization insofar as individuals are recognized or classified on the basis of certain common physical features (e.g., black skin). However, this classification is not social categorization in the sense that the physical classification has activated semantic meaning or social significance. "Identification" is essentially similar to what Bruner (1957) referred to as "primitive categorization," which entails the simple "perceptual isolation of an object or an event with certain characteristic qualities . . . the event may have no more 'meaning' than that it is an 'object' " (p. 131). "Categorization," on the other hand, refers to more elaborate semantic processing of stimuli in terms of what its features represent at a more conceptual level.

[2] At first blush, it may seem that an abstract group and the physical features associated with that group are isomorphic with one another. However, as the opening anecdote illustrates, it may be possible to have positive reactions to aesthetic perceptual features associated with a particular group such as "Mediterraneans" (e.g., curly dark hair, olive skin, hazel eyes), while holding negative evaluations of the conceptual qualities of the group (e.g., loud, aggressive, dishonest). Furthermore, in the case of African Americans, there is some degree of independence between racial group membership and possession of physical features associated with the group. In other words, there are many African Americans who do not possess very African-type physical features.

[3] Racial prototypically is defined here in terms of extremity of features rather than representativeness. Therefore, Blacks with more African-type features would be considered highly prototypic even if such individuals are not the most central (common) instances of the category African American.

pretested on a number of dimensions. First, a group of eight raters categorized all photos as being African American, Caucasian, or racially ambiguous. Any face classified as being racially ambiguous by any one of the eight raters was excluded from the pool. Thus, all African American faces used in the experiment were unanimously judged as being unambiguously African American. A group of 14 raters, composed of both men and women and both Blacks and Whites, then rated the remaining pool of White and African American faces on physical attractiveness using a 5-point rating scale. All faces were also rated on hostility of appearance using a 5-point rating scale.

In addition, prototypicality ratings were obtained for the African American faces. These ratings were made on a 5-point scale with endpoints labeled as 1 (*low prototypic*) an d5 (*high prototypic*). Raters were informed that judgments of African American prototypicality primarily entailed skin color but should also take into account facial features, eye color, hair texture, and so on, with high prototypicality indicating more "African-type" facial features and lower prototypicality indicating less "African-type" facial features.

On the basis of these ratings, eight African American photos with mean prototypicality ratings below 3.0 were matched on attractiveness and hostility of appearance with eight African American photos with mean prototypicality ratings above 3.0. Mean prototypicality ratings were 2.34 for the eight LP Blacks and 4.42 for the eight HP Blacks.[4] Each pair of African American faces was then matched with a White face of equal physical attractiveness and hostility of appearance to form eight matched trios of W, LP, and HP primes (24 faces). Attractiveness ratings for the 8 Whites ranged from 2.13 to 3.80 with a mean of 3.11; attractiveness ratings for the 8 LP Blacks ranged from 2.18 to 3.87 with a mean of 3.16; and attractiveness ratings for the 8 HP Blacks ranged from 2.13 to 4.00 with a mean of 3.14. Hostility ratings for Whites ranged from 2.00 to 3.77 with a mean of 2.97; hostility ratings for LP Blacks ranged from 1.90 to 4.00 with a mean of 3.04; and hostility ratings for HP Blacks ranged from 1.93 to 3.93 with a mean of 2.94.

By systematically selecting photos that were matched on these dimensions of attractiveness and hostility, we ruled out the likelihood that any differences in affective responses to the different photo sets would reflect differences in physical appearance other than the race-related features we were interested in. One consequence of the matching process was that all of the critical facial stimuli were rather mediocre in physical attractiveness; no highly attractive (or highly unattractive) faces were included in the critical sets.[5] The final sample of faces used in the experiment consisted of the matched sets of 8 Whites, 8 LP Blacks, and 8 HP Blacks in addition to 16 White fillers and 8 "other" fillers.

The target words used in the study were selected from a series of 50 nouns that were prerated on a 7-point Likert scale of positivity–negativity. Nouns were chosen instead of adjectives to minimize possible semantic links between primes and target words (to be sure that any match between face and word was based on affect rather than content). Therefore, any nouns considered by any one of five raters to have possible semantic associations with racial categories (e.g., crime, music, poverty) were excluded. From the remaining list of nouns, we selected eight positive (love, vacation, joy, romance, paradise, success, beauty, smile) and eight negative (garbage, vomit, poison, sewage, pest, despair, cockroach, disgust) nouns. The mean number of letters and syllables was 6.00 and 1.88, respectively, for positive words and 6.38 and 1.88, respectively, for negative words. Mean valence ratings were 6.36 for positive words and 1.93 for negative words.

Procedure. Following procedures developed by Fazio et al. (1995), each experimental session consisted of five phases, with the fourth phase serving as the crucial priming task. All experimental instructions were computerized. The first phase involved a word-meaning task whose purpose was to obtain baseline ratings for the 16 nouns used as targets in the later priming task. Students were instructed to press a key labeled *good* or a key labeled *bad* as quickly as possible to indicate their judgment of the word. Students were instructed to maximize the speed and accuracy of their responses. A row of asterisks preceded the presentation of each noun, serving as a warning signal that the target noun was about to appear. A given noun remained on the screen until the student responded or for a maximum of 1.75 s, with a 2.5-s interval separating each trial. The valence response was recorded along with the latency of response (from noun onset to response) to the nearest millisecond. After a practice block of 12 nouns intended to familiarize students with the task, students performed two blocks of trials. Each block consisted of the 16 critical nouns. The average latency of the two trials for any given noun served as the baseline latency for that word.

The next two phases were intended to prepare students for the priming task that would involve presentation of faces as primes and nouns as targets. The second phase was an ostensible face-learning task, and the third phase was a face-detection task. In the face-learning task, participants simply attended to a series of eight practice faces. All faces presented in practice trials were different from the 48 faces presented in the later crucial trials. Each face was presented twice, once in each of two blocks. In the face-detection phase, participants were shown a series of 8 faces (4 previously presented, 4 foils) and were instructed to press keys labeled *yes* if they recognized the face from the face-learning presentation or *no* if they did not recognize the face.

The fourth phase involved the actual priming task. Students were informed that the previous tasks would be combined, and that our objective was to determine the extent to which the judgment of word meaning was an automatic skill. The instructions informed participants that if word judgment was truly an automatic skill individuals should be able to perform just as well as in the first phase of the experiment even if they had to do something else at the same time. The instructions and procedures were identical to the first task, except that students were also told that it was important that they attend to the faces presented because they would be asked to recall the faces in the next task. The row of asterisks was replaced by the 48 color photographic images that served as primes. On any given trial, a prime was presented for 315 ms followed by a 135-ms interval before onset of the target noun. After an initial practice block involving different faces and nouns, four blocks of trials were presented. Each block consisted of 48 trials in which each of the 48 primes appeared once followed by 1 of the 16 nouns. Over the course of the four blocks, each prime was paired with 2 positive and 2 negative nouns. Although trials were presented in random order, each matched trio was followed by the same nouns across the four trial blocks. Trials involving the 8 matched trios (24 faces) constituted the actual trials of the experiment. All other trials involving the remaining 24 faces served as fillers and were included to reduce the overall proportion of Black faces, thereby minimizing the likelihood that students would become aware of any special interest in the race shown in the photograph.

The last phase of the experiment was the detection (recognition) task that students had been led to expect during the instructions for the fourth phase. This detection task involved the presentation of a series of eight images. Students were instructed to press the key labeled *yes* if the face was presented in the priming task or the key labeled *no* if the face had not been

[4] Because all faces were unambiguously rated as being African American, the range of prototypicality in our sample of faces does not approximate the full range of African American phenotypes. Therefore, these ratings should be interpreted in relative rather than absolute terms.

[5] Because our White faces were selected to be equivalent in rated attractiveness to African American faces, if there were an in-group bias (on the part of our White raters) in judgments of facial attractiveness, it is possible that the White photographs used in our experiments here were less attractive, on average, than the photographs used in previous experiments with Black and White faces as stimuli. Copies of sample photographs used in the current experiment are available on request from us.

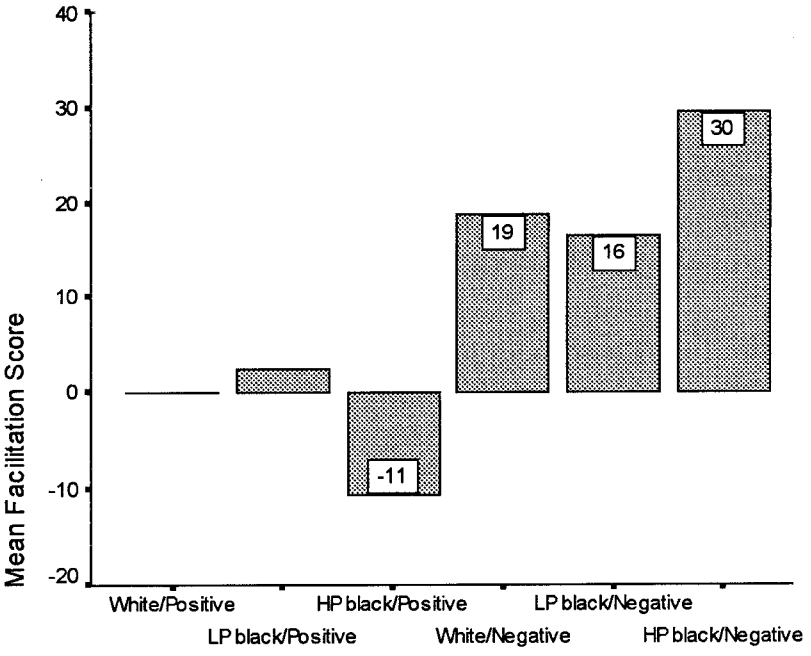

Figure 1. Experiment 1: Facilitation and inhibition as a function of word valence and facial prime. Positive numbers indicate facilitation. Negative numbers indicate inhibition. HP = high prototypic; LP = low prototypic.

presented in the priming task. After completing the face-recognition task, participants were debriefed, thanked for their participation, and dismissed.

Results

The raw (untransformed) latency for any given target noun when preceded by a given face (as measured in the fourth phase) was subtracted from the baseline latency of that noun (as measured in the first phase) to arrive at a facilitation score; positive values indicate that responses to the target noun were facilitated compared with baseline and negative scores indicate inhibition.[6] Mean facilitation scores for positive target nouns and negative target nouns were computed for each face and then averaged across the eight faces of each type. This yielded a total of six facilitation scores for each participant (i.e., W, LP, or HP primes followed by either positive or negative target nouns).

A 3 (race of prime: W, LP Black, or HP Black) × 2 (valence of target: positive vs. negative) within-subjects analysis of variance (ANOVA) conducted on these facilitation scores revealed a significant main effect of valence, $F(1, 48) = 11.59$, $p < .001$, such that participants were faster to respond to negative nouns after facial primes than to positive nouns. This valence effect may have been caused by the "neutral" facial expressions of our primes, which may have appeared slightly unpleasant compared with expressions normally displayed in posed photos. More importantly, this valence effect was qualified by a significant Race × Valence interaction, $F(2, 47) = 4.39$, $p < .02$ (Figure 1). This critical interaction was decomposed into between-race and within-race orthogonal contrasts to determine the locus of the priming effects.

The first contrast compared White faces with Black faces (LP + HP) to determine whether the significant Prime × Valence effect was due to between-category differences in racial category membership. This between-races comparison produced no significant interaction between prime and valence ($F < 1$). The second contrast compared LP Blacks with HP Blacks to assess the degree to which responses reflected differential affective responsiveness to physical cues per se, controlling for racial category membership. Comparing LP and HP primes within race, the interaction between physical prototypicality and valence was highly significant, $F(1, 48) = 8.74$, $p < .005$, such that, compared with HP primes, LP primes produced more facilitation to positive words, $t(48) = 2.33$, $p = .02$, and less facilitation to negative words, $t(48) = -1.78$, $p = .08$.

Although the Black–White comparison was not significant when responses to LP and HP Black faces were combined, a post hoc contrast between W and HP Blacks revealed a significant interaction, $F(1, 48) = 4.05$, $p = .05$. White primes produced more facilitation to positive words and less facilitation to negative words compared with HP Black facial primes. Thus, the pattern of race priming effects previously reported in the literature was replicated when HP Blacks represented the African American category. However, as seen in Figure 1, LP Blacks produced a pattern of facilitation similar to that of Whites, with even a nonsignificant pattern of greater facilitation to positive targets and less facilitation to negative targets. Treating race of the photo set as a 3-point scale (HP-LP-W), trend analyses were conducted for the Race × Valence interaction. These analyses revealed a significant linear trend for the race factor, $F(1, 48) = 4.07$, $p < .05$, qualified by a

[6] Incorrect valence responses and responses shorter than 200 ms were omitted from the data. This resulted in a total of 4.26% of the responses being excluded. Additionally, the distribution of difference scores was examined for skew but proved to be approximately normally distributed.

significant quadratic component, $F(1, 48) = 4.32$, $p < .05$, reflecting the positive elevation of outcomes for the LP primes.

Discussion

Previous research has assumed that facial stimuli serve as symbolic representations of their respective social categories, such that differential responsiveness to facial primes reflects evaluations of racial categories. The results of Study 1 indicate instead that within-group differences in affective responses to physical cues were greater than between-group differences in affective responses to category membership. When responses to HP and LP photos are combined, the Black–White racial category difference in automatic evaluations is nonsignificant. Such a finding is inconsistent with the notion that faces produce automatic categorization and category-based evaluation.

An alternative to the cue-based hypothesis is that White perceivers in Experiment 1 simply perceived the LP Blacks as being (suntanned) Whites rather than African Americans. Although all of the LP primes were unanimously categorized as being African American under conscious pretesting conditions, they may have appeared White under more brief exposure conditions. To address this possibility, Experiment 2 was designed to test whether LP Black facial primes are indeed categorizable as African American even under the brief (315-ms) exposure time conditions of Experiment 1.

Experiment 2

In this experiment, reaction time and accuracy were recorded for categorization judgments of W, LP, and HP facial primes. Because LP primes are, by definition, less clear examples of the category African American (compared with HP primes), we reasoned that perceivers might take longer to categorize LP primes compared with HP primes (Mervis & Rosch, 1981). Nonetheless, we hoped to demonstrate that LP primes could be accurately categorized as African American rather than White under these low-exposure conditions.

Method

Participants. Thirty participants enrolled in partial fulfillment of course requirements. Data for two participants were omitted as a result of computer malfunctions as were data for 1 participant who failed to complete the study. The 2 African American participants were also excluded. Thus, final analyses were conducted on 25 participants (12 men, 13 women).

Procedure. As in Experiment 1, participants were informed by computerized instructions that they would be taking part in a face-recognition task that would test how well people recognize faces under low-threshold conditions. However, in this study, they were informed that their job was to categorize each face as being White or Black because this would help them to remember the faces later on. On any given trial, each facial prime was presented for only 315 ms, after which students judged the race of the face by pressing a key labeled *Black* or a key labeled *White*.[7] Students were instructed to maximize the speed and accuracy of their responses. The accuracy (race) of the response was recorded along with the latency of response (from face onset to response) to the nearest millisecond. After an initial practice session designed to familiarize participants with the task, participants completed the critical trial block. This block consisted of 32 facial primes, which included the 8 critical W primes, 8 critical LP Black primes, and 8 critical HP Black primes used in Experiment 1 along with 8 filler W primes, which were included to equate the proportion of White and Black facial primes being presented. All trials were presented in random order. After completing this categorization task, participants were thoroughly debriefed, thanked for their participation, and dismissed.

Results

Categorization accuracy. The primary issue that Experiment 2 sought to resolve is whether LP primes are accurately categorized as African American (as opposed to White) under 315-ms exposure conditions. Table 1 reports the percentage of accurate categorization judgments for each type of facial photo. Overall accuracy was extremely high and did not significantly differ across the three groups of faces, $F(2, 23) = .40$, $p = .68$.

Response time. These were averaged for each of the eight facial primes within each of the three critical groups (W, LP, and HP primes). As shown in Table 1, participants were significantly faster to categorize HP primes as compared with LP primes, $t(24) = 4.75$, $p < .001$, or White primes, $t(24) = 3.95$, $p < .001$.

Discussion

Experiment 2 was designed to investigate the possibility that the differential evaluative responses to LP and HP faces found in Experiment 1 could have been caused by participants' perception of LP primes as White instead of Black under low-threshold conditions. The results of Experiment 2 reveal, however, that LP faces were accurately categorizable as Black even at brief exposure. Thus, in all likelihood, the results of Experiment 1 were not due to miscategorization of LP primes as White rather than Black. Nevertheless, because the absolute categorization times for LP Blacks was slightly longer than the 450-ms stimulus onset asynchrony in the Fazio task of Experiment 1, it is still possible that LP Black faces could have been uncategorized or miscategorized in that experiment. We argue that the results of Experiment 1 were due not to differences in categorization of HP and LP photos but rather to differences in evaluative responses to the phenotypic features of the faces themselves. However, we cannot completely rule out the alternative possibility that HP faces were automatically categorized as Black, whereas LP faces (even though they could be accurately identified as Black) were simply uncategorized, as were the White faces. In that case, the relative facilitation of negative evaluations after the HP primes reflects negative affect to the racial category, whereas responses after the LP and White primes reflect idiosyncratic evaluations to the faces per se (or some neutral baseline).

The results of Experiment 2 do not provide direct support for or against an automatic categorization of HP primes hypothesis because there is no a priori basis for determining what threshold speed distinguishes automatic as opposed to more effortful decision making. Hence, we have to seek more indirect evidence that differential categorization of LP versus HP faces is not playing a

[7] The "Black" key was always on the right and the "White" key was always on the left of the keyboard. Primes remained on the computer screen for only 315 ms before they disappeared. If participants did not respond after 1.75 s, responses were recorded as missing values. Missing values constituted less than 1% of all categorization responses.

Table 1
Accuracy and Reaction Time Data for Facial Primes

Ethnic group	Accuracy (%)	Reaction time (ms)
LP primes	96.00_a	498_a
HP primes	96.50_a	445_b
White primes	97.50_a	494_a

Note. Means with different subscripts differ at $p < .05$. LP = low prototypic; HP = high prototypic.

role in the automatic evaluation effects. If HP faces are automatically categorized as African American (whereas LP faces are not), then HP faces should show category-based facilitation to stereotypic target words, whereas LP primes should not. Automatic evaluation tasks (i.e., "good"–"bad" word judgments), such as that used in Experiment 1, tap evaluative or affective facilitation rather than semantic facilitation. A lexical-decision task (i.e., "word"–"not word" judgments), on the other hand, is more sensitive to stereotypic or semantic associations between prime and target (Klauer, 1998). Therefore, this task is more appropriate for assessing stereotypic or semantic facilitation. If it is the case that HP primes elicited automatic categorization, then they should show automatic facilitation to stereotypic words, because stereotyping is category based.

In short, if HP Blacks are automatically categorized as Black (whereas LP primes are not), then they should activate the African American stereotype, whereas LP Blacks should not. Thus, the current study aims to assess the effects of HP versus LP primes on relative semantic facilitation to stereotypic words, as one test of the automatic categorization hypothesis. If processing of HP primes (or any facial prime) is category based, then they should show semantic facilitation to the relevant stereotypes. However, if processing of faces is cue based rather than category based, then stereotypes should not be activated for any of the primes.

Experiment 3

Our explanation for the findings of Experiment 1 assumes that the automatic evaluation effects obtained reflect direct affective responses to the racial cues without mediation by category activation. If, however, the evaluations were mediated by categorization, responses to target words that follow HP compared with LP facial primes should be affected by the stereotypic (semantic) content as well as the valence of the words being primed. More specifically, HP primes should facilitate responses to traits that are part of the African American stereotype (particularly negative stereotypic traits), more so than LP primes or White primes. In short, if the automatic evaluation effects obtained in Experiment 1 are mediated by differences in automatic category activation, then decision time for words preceded by HP versus LP faces should vary as a function of the stereotype relevance of the target word.

Method

Participants. Fifty-five participants took part in the experiment in partial fulfillment of requirements for course credit. Five African American participants were excluded from further analyses as were 3 participants who expressed suspicion regarding our hypotheses and 1 participant who failed to follow instructions. Thus, our final sample consisted of 46 participants (20 men, 26 women).

Procedure. Experiment 3 differs from Experiment 1 in two important respects: (a) Experiment 1 used affective nouns as target words, whereas Experiment 3 uses stereotypic adjectives as targets, and (b) Experiment 3 uses a lexical-decision task, whereas Experiment 1 used an automatic evaluation task. The reason for the first change is fairly evident. In Experiment 1 we intentionally eliminated any stereotypic content from the word judgment task to obtain an unconfounded assessment of automatic evaluation per se. In Experiment 3 we are explicitly interested in the activation of stereotypes themselves. The reason for the second change is that, as previously mentioned, the lexical-decision task assesses semantic content, whereas the automatic evaluation task taps valence or affect (Klauer, 1998). In this study, we are interested in semantic links between category and target words (stereotypes), so the lexical-decision task is more appropriate.

Accordingly, judgment decisions are changed to *word* (right key) and *not word* (left key). Additionally, target words in Experiment 3 were modified to consist of eight pronounceable nonwords (e.g., packles, brabend) and eight stereotype-related adjectives, including two positive Black stereotypes (athletic, religious), two negative Black stereotypes (hostile, lazy), two positive White stereotypes (wealthy, ambitious), and two negative White stereotypes (snobbish, materialistic). Stereotypic target words were taken from prior studies on racial stereotyping (e.g., Dovidio, Evans, & Tyler, 1986; Lepore & Brown, 1997; Wittenbrink, Judd, & Park, 1997). All facial primes were identical to those used in Experiment 1.

Apart from the modification of the word-judgment task, the experimental sessions were conducted as in Experiment 1. Baseline reaction times were obtained for each target word judgment in the first phase of the experiment and were judged again after presentation of facial primes in the fourth phase.

Results

A 3 (race of prime: W, LP Black, or HP Black) × 2 (stereotypicality of target word: Black vs. White) × 2 (valence of target word: positive vs. negative) within-subjects ANOVA was performed on facilitation scores (as calculated in Study 1). As observed in Experiment 1, there was a significant main effect of valence, $F(1, 45) = 21.66$, $p < .001$, indicating that participants were quicker to respond to negatively valenced words than to positively valenced words when primed by the facial photos. There was also a significant main effect of stereotype, $F(1, 45) = 62.27$, $p < .001$, such that participants were quicker to respond to White stereotype target words than to Black stereotype words. There was also a theoretically uninteresting Stereotype × Valence interaction, $F(1, 45) = 37.61$, $p < .001$. Negative Black stereotype words took longer to respond to than negative White stereotypes, but there was no difference in response time between positive White and positive Black stereotypes.

Of primary interest were reaction times to stereotypic targets as a function of race of prime. However, the aforementioned differences in decision time for different word types were not modified by the race of prime. None of the interactions involving race

approached significance (all Fs < 1.08),[8] indicating that racial characteristics of the facial stimuli had no priming effect on judgments of stereotypic target words.

Discussion

Experiment 3 demonstrated that the same facial stimuli that produced evaluative facilitation to affective nouns in Experiment 1 did not produce semantic facilitation to stereotypic words in Experiment 3. The lexical-decision task, which taps semantic content, failed to produce any evidence of stereotypic facilitation to HP primes. In fact, there was absolutely no evidence of stereotypic facilitation for any of the faces, although there should have been if faces elicited category-based processing. Whereas Experiment 1 provides support for cue-based processing, the results of Experiment 3 fail to provide any support for the automatic categorization hypothesis.

On the surface, the results of Experiment 3 may seem somewhat inconsistent with the results of Experiment 1 or even the automatic stereotyping and prejudice literature in general. One source of the confusion may be the widespread conflation of the concepts of automatic stereotyping and automatic prejudice (affect). Experiment 1 taps automatic prejudice (affective facilitation), whereas Experiment 3 tests automatic stereotyping (semantic facilitation). Our entire hypothesis here is that faces do not, under normal default conditions, elicit automatic category-based processing. Because stereotyping is, by definition, category based, we would not expect facial primes to produce automatic stereotyping under normal conditions.

We are unaware of any studies in the literature that have found evidence of automatic, uncued stereotyping using facial primes in sequential priming paradigms. To be sure, studies such as Fazio et al. (1995) or Dovidio, Kawakami, Johnson, Johnson, and Howard (1997), for instance, are sometimes cited as evidence of "automatic stereotype activation." In actuality, however, these studies do not tap stereotyping but rather automatic affect, using facial primes in an automatic evaluation task with nonstereotypic target words. There are, of course, sequential priming studies that do measure automatic stereotyping (e.g., Kawakami, Dion, & Dovidio, 1998; Wittenbrink et al., 1997). However, these studies use lexical-decision or pronunciation tasks (as opposed to automatic evaluation tasks), with category labels as primes (as opposed to facial primes) and stereotypic adjectives as target words.[9] In short, Experiment 1 is consistent with the literature on automatic prejudice, whereas Experiment 3 it is not necessarily inconsistent with the literature on automatic stereotyping.

It is interesting however, that a 1997 study by Macrae, Bodenhausen, Milne, Thorn, and Castelli did find evidence of stereotype activation using facial primes but only in a condition in which participants were first given the specific goal of categorizing faces. However, consistent with the results of Experiment 3, null effects emerged when participants simply attended to the faces without a categorization goal. The authors hypothesized, based on assumptions of goal-dependent automaticity, that mere exposure to female faces would be insufficient for automatic gender stereotype activation. They reasoned, however, that stereotype activation might occur under conditions in which participants were given the explicit goal of categorizing the faces. Results from two studies revealed null effects, in both studies, for activation of gender stereotypes in the "exposure" condition, in which each participant's goal was merely to attend to stimuli on the computer screen, but a significant stereotyping effect, in both studies, in the "semantic-judgment" condition, in which each participant's goal was to categorize visual stimuli.[10] Consistent with our findings, then, these results demonstrate that category-based processing does not occur spontaneously at the automatic level. However, category-based processing did occur in the event of a processing goal.

Although Brewer's (1988) dual-process model posits that categorization is generally a controlled process, category-based processing may occur at the automatic level if processing goals are introduced, consistent with the tenets of "conditional automaticity" (Bargh, 1989) and the results of Macrae et al. (1997). However, in the absence of processing goals, "default" automatic responses to faces will reflect cue-based processing. Specifically, we predict that HP–LP differences in automatic evaluation will emerge when facial primes are processed as perceptual stimuli (as in Experiment 1) but will disappear if a categorization goal is introduced (because LP and HP primes belong to the same category). In other words, the categorization goal should lead to stronger between-race differences than within-race differences, whereas the no-goal condition should lead to stronger within-race differences than between-race differences. To test this idea explicitly, we conducted a fourth experiment in which, at the outset of the sequential priming task, participants were randomly assigned to either an "exposure" condition (identical to Experiment 1) or a "goal" condition, in which they were instructed to categorize each facial prime (as Black or White) as it appeared on the computer screen.

Experiment 4

Experiment 4 was designed to test whether participants' automatic evaluation of facial primes can be influenced by processing goals. Specifically, we predicted that the LP versus HP contrast would be significant only under default (exposure-only) conditions in which no categorization judgment was required. However,

[8] Orthogonal contrasts involving the within-race (HP vs. LP) comparison or the White–Black comparison were also nonsignificant ($F < 1$). The means for HP and LP facilitation scores to positive Black stereotypes were 20.26 and 21.50, respectively, and -24.07 and -20.13, respectively, for negative Black stereotypes. For the between-race (W vs. LP + HP) comparison, the means for W and Black facilitation scores to positive Black stereotypes were 36.66 and 20.88, respectively, and -15.02 and -22.10, respectively, for negative Black stereotypes.

[9] Studies by Bargh, Chen, and Burrows (1996) and Chen and Bargh (1997), which measured behavioral priming as opposed to automatic stereotyping per se, reported evidence that subliminal exposure to Black primes activated the hostility stereotype. However, it is not clear from their behavioral evidence whether participants' behavior was due to activation of the hostility stereotype per se (i.e., they were behaving in a hostile manner) or whether exposure to Black faces simply put them in a bad mood (i.e., they were irritable and annoyed). In any case, the primary focus of these studies was automatic behavioral activation or self-fulfilling prophecy rather than assessing automatic beliefs or attitudes per se.

[10] The categorization condition involved classification of objects as animate or inanimate. However, the categories "animate" and "women" were completely confounded because all animate stimuli in the study were also women.

differences in evaluative responses to LP versus HP primes should diminish when participants are given the goal of categorizing facial primes. That is, the within-race contrast should be greater than the between-race contrast in the no-goal condition, whereas the between-race contrast should be greater than the within-race contrast in the category goal condition.

Comparisons between no-categorization and categorization-processing goals should also provide another indirect test of the competing hypothesis that HP Blacks are automatically categorized. If that were the case, then the priming effect of HP photos should not be influenced by the categorization task requirements. Facilitation of negative evaluations should reflect category-based affect in either case, and only responses to White and LP photos should be differentially influenced by the categorization processing goal. On the other hand, if (as we expect) the task manipulation induces a switch from cue-based to category-based processing, then White photos should show more facilitation of positive evaluations under categorization compared with no-categorization conditions, whereas HP and LP photo effects should be equalized under categorization instructions, which could involve an increase in negative facilitation for LP Blacks or a decrease in negative facilitation for HP Blacks, or both.

Method

Participants. Ninety students participated in partial fulfillment of course requirements. Three participants were excluded because of computer malfunctions as were 2 others who failed to follow instructions. The 3 African American participants were also excluded. Thus, final analyses consisted of data from 82 participants (32 males, 50 females).

Procedure. All participants were randomly assigned to either the no-goal (exposure) or goal (categorization) condition. The exposure condition was essentially similar to the procedure described in Experiment 1. Whereas participants in the exposure condition were instructed to simply attend to and remember the faces being presented, participants in the categorization condition were given computerized instructions to mentally categorize facial primes as being either Black or White as they were attending to the faces. They were told that it was very important to categorize each face in this way to successfully complete the task that was to follow. Unlike Experiment 2, the categorization task in this study was intrapsychic and did not involve pressing keys. Apart from the categorization instructions, all other experimental aspects were identical to those in the exposure condition (as well as Experiment 1).

Results

As in Experiment 1, the raw latency for any given target noun when preceded by a given face was subtracted from the baseline latency of that noun to arrive at a facilitation score. Mean facilitation scores for positive target nouns and negative target nouns were computed for each face and then averaged across the eight faces of each type. Because of the number of conditions in this study, we subtracted facilitation scores for positive target words from facilitation scores for negative target words.[11] This yielded a single index of net negative evaluation for each of the three groups (W, LP, and HP). Higher scores indicate more automatic negativity relative to positivity.

A 3 (race: W, LP, HP) $\times$ 2 Processing Goal (categorization vs. no goal) mixed ANOVA was performed, with the last factor as a between-subjects measure and net negative evaluation facilitation as the dependent measure. Results revealed a significant two-way Race $\times$ Goal interaction, $F(2, 79) = 3.21$, $p < .05$ (Figure 2).

In the default exposure condition, the simple effect for race of photo was significant, $F(2, 38) = 7.30$, $p < .002$. Breaking this down into orthogonal contrasts, the within-race (LP vs. HP) contrast was highly significant, $F(1, 39) = 14.92$, $p < .0001$, whereas the between-race contrast was not, $F(1, 39) = .23$, $p = .63$. Replicating the findings of Experiment 1, participants in this condition showed significantly more negativity toward HP Blacks than toward LP Blacks. Further, HP Blacks elicited more automatic negativity than White faces, $t(40) = -1.73$, $p = .09$, whereas LP Blacks elicited significantly less negativity (i.e., greater positivity) than the White faces, $t(40) = 2.23$, $p < .04$. Thus, the trend that was observed in Experiment 1 in the direction of more positivity toward LP Blacks vis-à-vis Whites was replicated (and statistically significant) in Experiment 4. Although the W-HP contrast was only marginally significant in Experiment 4, the overall pattern of W-HP difference is highly reliable when collapsed across Experiments 1 and Experiment 4 (exposure condition), $F(1, 88) = 7.08$, $p < .009$.

The pattern of automatic negativity was quite different in the condition in which an explicit categorization goal was introduced. In this condition, the within-race (HP vs. LP) contrast did not approach significance, $F(1, 41) = .02$, $p = .89$. Instead, there was a marginally significant between-race contrast, $F(1, 41) = 1.72$, $p < .20$.

Comparing between task conditions, the direction of differences in negative facilitation was consistent with a switch from cue-based to category-based evaluations. The White faces produced less negative facilitation under categorization instructions than under the exposure-only task condition, $t(80) = 1.64$, $p = .10$. The effect of categorization for HP faces was also in the positive direction, $t(80) = 1.64$, $p = .10$, whereas the effect for LP faces was nonsignificantly negative, $t(80) = -.35$, *ns*, although both Black primes produced more negative facilitation than the White primes under categorization instructions.

Discussion

The important finding demonstrated in this experiment was the significant difference in the HP-LP effect under the two different processing goal conditions. Without explicit categorization instructions, automatic evaluation was highly responsive to physical features of the facial photographs, which varied in racial protoypicality. When a racial categorization task was introduced, this physical prototypicality effect was completely wiped out. Automatic evaluation of HP and LP faces was equivalent under categorization goal conditions. This means that the slightly longer time it takes to categorize LP Blacks compared with HP Blacks (Experiment 2) had no impact on the evaluative reactions once the categorization was made.

Results of Experiment 4 also indicate that categorization of Black and White faces by race is not automatic in the absence of specific instructions to make such category judgments. HP faces

[11] As in Experiment 1, there was a significant valence main effect such that facilitation to negative words was greater than facilitation to positive words regardless of racial category of the prime.

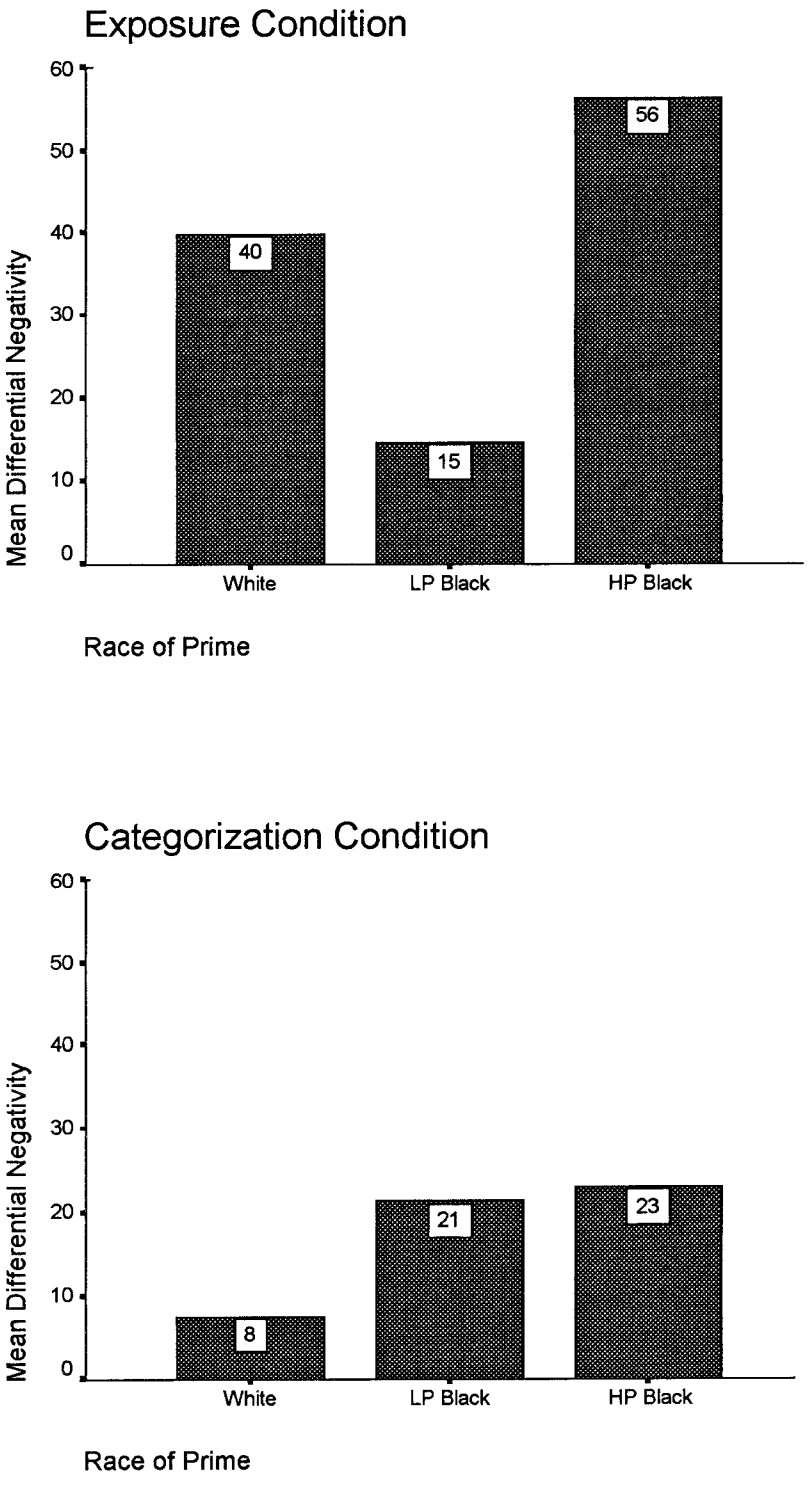

Figure 2. Experiment 4: Differential negativity toward facial primes as a function of race. Higher scores indicate more negativity. HP = high prototypic; LP = low prototypic.

were actually responded to less negatively when categorization intervened than when affective responses were made without categorization instructions. This finding is not consistent with the idea that the negativity obtained in Experiment 1 after HP primes was mediated by categorization. Rather, it suggests that the automatic evaluations were based on the value attached to the physical features of the faces themselves rather than the category memberships they imply.

The findings from this experiment also have implications for understanding why different measures of implicit prejudice may produce different results, at least when faces are used as racial stimuli. It has been reported, for instance, that individual differences in implicit racial prejudice as assessed by the sequential priming paradigm are not strongly correlated with assessments based on the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998; Cameron, Alvarez, & Bargh, 2000). Automatic evaluation assessed by the sequential priming procedures used by Fazio et al. (1995) does not require categorization of the facial stimuli used as primes. Our results suggest that responses to the faces will be cue based rather than category based in the absence of explicit demands to categorize based on race. The procedures associated with the IAT paradigm, on the other hand, do explicitly require a categorization of a face or first name.[12]

Although significant Black–White differences in implicit evaluation have been obtained with both the IAT and the Fazio paradigm, we expect that category-based racial effects would be more robust and consistent across studies using the IAT compared with those using the Fazio procedure, in which idiosyncratic variations in facial cues per se are likely to have more impact. This difference, however, may be an artifact of the different processing goals associated with the two procedures rather than any inherent difference in the reliability or validity of the methods themselves.

In short, these findings may have implications for potential hidden moderators of correlations among implicit measures or between implicit and explicit measures of prejudice. Although some previous research has proposed that implicit and explicit measures of prejudice do not correlate with one another because they represent qualitatively different cognitive systems (Greenwald & Banaji, 1995; Merikle, 1992), it has been empirically demonstrated that implicit and explicit measures of prejudice sometimes do correlate with one another (Kawakami et al., 1998; Lepore & Brown, 1997; Locke, MacLeod, & Walker, 1994; Wittenbrink, et al., 1997). The correlations obtained within and between explicit and implicit measures of prejudice may have less to do with whether they are controlled versus automatic than with the type of prejudice that they happen to be tapping.

Roediger (1990) argued that correlations between implicit and explicit measures are driven by "transfer appropriateness" (i.e., perceptual based vs. conceptual based) between measures rather than level of explicitness per se (conscious vs. automatic). In Roediger's words, "because standard explicit tests reflect meaning or conceptual elaboration, they can be referred to as conceptually-driven tests. On the other hand, many implicit tests reflect perceptual processing . . . these [transfer-appropriate processing] assumptions account well for the cross-over dissociation between explicit and implicit tests" (p. 1049). According to this perspective, correlations between different measures of prejudice are not determined by whether the measure is implicit or explicit but rather the type of evaluative processing that the measure evokes (i.e., perceptual vs. conceptual).

In short, if automatic priming paradigms using faces tap perceptual-based evaluation of physical cues, they should not necessarily correlate with more conceptual, category-based measures of prejudice such as the Modern Racism Scale (MRS; McConahay, 1986). On the other hand, if automatic priming paradigms using category labels tap evaluations of the category, as opposed to perceptual-based responsiveness of physical cues, then these measures should correlate with other category-based measures of prejudice, such as the MRS. It is interesting that the literature on automatic prejudice reveals a pattern of findings consistent with this formulation. That is, automatic paradigms using faces have generally not been correlated with the MRS (e.g., Dovidio et al., 1997; Fazio et al., 1995; but see Fazio & Dunton, 1997), whereas many automatic paradigms using category labels as primes have been correlated with the MRS (e.g., Kawakami, et al., 1998; Lepore & Brown, 1997; Locke et al., 1994; Wittenbrink et al., 1997).

In light of this suggestive evidence, Experiment 5 was designed to test whether an explicit measure of "perceptual prejudice" (Livingston, 2001) would be more related to automatic responses to facial primes than more conceptual-based evaluations of racial categories, such as the MRS.[13] If differential automatic responses to LP versus HP faces reflect affective responses to physical cues, then individuals who are more sensitive to evaluation based on physical cues should show the effect more than those who do not possess such a chronic tendency. If, on the other hand, automatic responses to HP faces reflect automatic categorization, then those who are more willing to express negativity toward Blacks as a category should show the effect more strongly than those who do not negatively evaluate Blacks as a group.

Experiment 5

Using an individual difference approach, Experiment 5 sought to provide further evidence that automatic responses to facial primes reflect cue-based as opposed to category-based evaluation. Drawing on Roediger's (1990) notion of transfer appropriateness, the logic of this experiment is that if differences in evaluation of LP and HP Blacks simply reflect greater categorization (and hence category-based evaluation) of HP primes, then this difference should be moderated by individual differences in conscious attitudes toward Blacks, as measured by the MRS. However, if evaluative differences between LP and HP primes reflect instead differential evaluation of physical cues per se, then this differential evaluation should be moderated by individual differences in perceptual-based prejudice.

Method

Procedure. The basic priming procedure used in this experiment was essentially similar to that used in Experiments 1 and 4 (exposure condition). After completing the automatic evaluation task, participants were given a computerized version of a "global opinion survey." This survey contained the seven-item MRS, which served as the measure of conceptual-based prejudice, and the nine-item Perceptual Reliance Index (PRI; Liv-

[12] It is worthwhile to mention that the IAT is not a sequential priming paradigm. Instead, it measures response competition to two evaluative categories (participants are first trained to categorize the stimulus faces as Black or White and then later to respond with a good–bad and/or black–white judgment of those faces simultaneously) rather than evaluative responses to each individual stimulus itself (whether it is a specific name or face). Thus, results obtained with faces on the IAT may not be directly comparable to results obtained with faces in the Fazio paradigm, which aggregates direct responses to individual stimuli.

[13] We have data indicating that scores on the MRS and explicit feeling thermometer ratings of the label "Blacks" are highly correlated ($r = .64$), suggesting that the scale does, in fact, tap explicit evaluations of Blacks as a category.

ingston, 2001), which taps individual differences in reliance on perceptual cues in forming social judgment. Examples of PRI items include "You can tell a lot about someone's character just by looking at them," "In general, light-skinned Blacks are more attractive than dark-skinned Blacks," and "You can usually tell if a women is a lesbian by her physical appearance."

Participants. One hundred ten introductory psychology students participated in this study in partial fulfillment of requirements for course credit. Data for 5 African American students were excluded from subsequent analyses as were data for 5 participants who expressed suspicion regarding the hypotheses and for 4 others who either did not follow instructions or experienced computer malfunctions. Therefore, the sample for final analyses consisted of 96 participants (36 men, 60 women).

Results

The measure of reliance on perceptual cues (PRI) had an internal reliability of .77 in this sample. Each participant was classified as high or low on the basis of a median split on his or her scores on the perceptual reliance index. A 2 Prototypicality of Prime (HP vs. LP) × 2 Word Valence (positive vs. negative) × 2 PRI (median split: high vs. low) mixed ANOVA was performed with PRI as the between-subjects factor. As with the prior experiments, this analysis revealed a significant main effect of valence, $F(1, 94) = 4.42$, $p < .04$, indicating that negative words were responded to more quickly than positive words. This main effect was qualified by a significant Prototypicality × Valence interaction, $F(1, 94) = 5.65$, $p < .02$. As found in Experiments 1 and 4 (exposure condition), HP primes produced significantly more facilitation to negative words and less facilitation to positive words compared with LP primes. More importantly, however, this two-way interaction was qualified by a three-way Prime × Valence × PRI interaction, $F(1, 94) = 4.32$, $p < .04$.

This three-way interaction was broken down and analyzed separately for high and low PRI participants. For high PRI participants, there was a significant two-way Prime × Valence simple interaction, $F(1, 45) = 7.15$, $p < .01$, indicating that HP primes produced significantly more facilitation to negative words than LP primes, $t(45) = -2.13$, $p < .04$, with no difference in facilitation to positive words, $t(45) = 1.15$, $p < .26$ (Figure 3). However, the Prime × Valence simple interaction for low PRI participants did not approach significance, $F(1, 49) = .07$, $p = .80$ (see Figure 3).

To test whether automatic responsiveness to facial prototypicality was moderated by the MRS as well, a 2 Prototypicality of Prime (HP vs. LP) × 2 Word Valence (positive vs. negative) × 2 MRS (median split: high vs. low) mixed ANOVA was performed. The three-way interaction with the MRS did not obtain significance, $F(1, 94) = 1.76$, $p = .19$. Further, given the significant correlation between the PRI and the MRS ($r = .36$), a 2 Prototypicality of Prime (HP vs. LP) × 2 Word Valence (positive vs. negative) × 2 PRI (median split: high vs. low) mixed analysis of covariance (ANCOVA) was performed with MRS scores entered as a covariate. The three-way Prime × Valence × PRI interaction remained marginally significant, $F(1, 93) = 3.19$, $p = .077$, when controlling for modern racism.

Discussion

Experiment 5 provides yet further evidence that automatic responses to faces reflect evaluation of physical cues per se rather than category membership. That the robust, differential overall evaluative responses to LP and HP Blacks observed in Experiments 1, 4, and 5 can be further qualified by individual differences in sensitivity to perceptual cues provides additional evidence that differential evaluative responses to these facial primes reflect cue-based rather than category-based evaluation. Further supporting this idea is the finding that explicit category-based evaluations of Blacks had no impact on participants' differential responses to LP versus HP faces. Finally, the perceptual reliance index moderated responses to facial primes even controlling for modern racism. These findings are consistent with the idea that the correlation between explicit and implicit measures depends on transfer appropriateness (Roediger, 1990). In this case, an explicit measure of sensitivity to perceptual cues did correlate with differences in implicit responses to phenotypically different faces. However, an explicit measure of conceptual prejudice did not correlate with effects on the implicit measure.

General Discussion

Using a variety of methodologies and techniques—from varying the nature of the priming paradigm to manipulating goals of the perceiver to measuring individual differences—the five experiments reported here provide consistent and convergent support for the idea that automatic evaluations of facial primes reflect affective responses to perceptual cues per se rather than conceptual evaluation of the racial categories that the cues represent. Moreover, these findings are entirely consistent with both prior theory and current research (e.g., Bargh, 1997; Brewer, 1988; Brewer & Feinstein, 1999; Giner-Sorolla et al., 1999; Smith & DeCoster, 2000).

For instance, a model proposed by Smith and DeCoster (2000) hypothesizes two distinct memory systems that account nicely for such a dissociation. According to this model, memory is characterized by two independent, qualitatively distinct modes of processing, namely the associative route, which primarily represents affective associations to certain features, and the rule-based route, which represents more cognitive processing of "culturally transmitted knowledge" (see Smith & DeCoster, 2000). More specific to the present article, Smith and DeCoster (2000) discussed evidence of "dissociations between people's positive or negative evaluative responses to *pictures* of Black stimulus persons and their favorable or unfavorable attitudes of Blacks as a *group*" (p. 117, emphasis added). They proceeded to argue that these dissociations may be due to the fact that "specific and general judgments call differentially on the two processing modes . . . It seems likely that more richly detailed, specific stimuli are better cues for responses to the associative system, whereas more general and abstract stimuli—often verbally represented—better afford processing by symbolic rules" (Smith & DeCoster, 2000, p. 117).

This distinction between evaluation associated with perceptual-based processing and that associated with conceptual-semantic processing has gained further credence with experimental evidence that category and exemplar processing are dominated by different hemispheres of the brain (Zarate, Sanders, & Garza, 2000). This research distinguished between social perception involving perceptual-episodic information and conceptual-abstracted information and found evidence that processing of the latter occurs primarily in the left hemisphere, whereas processing of the former occurs in the right hemisphere. In addition, using the mere exposure paradigm, the researchers found that "the association of positive affect to previously viewed target photos . . . occurs only in the right hemisphere" (p. 223). In short, their experiments provide neurological evidence that (a) perceptual and conceptual

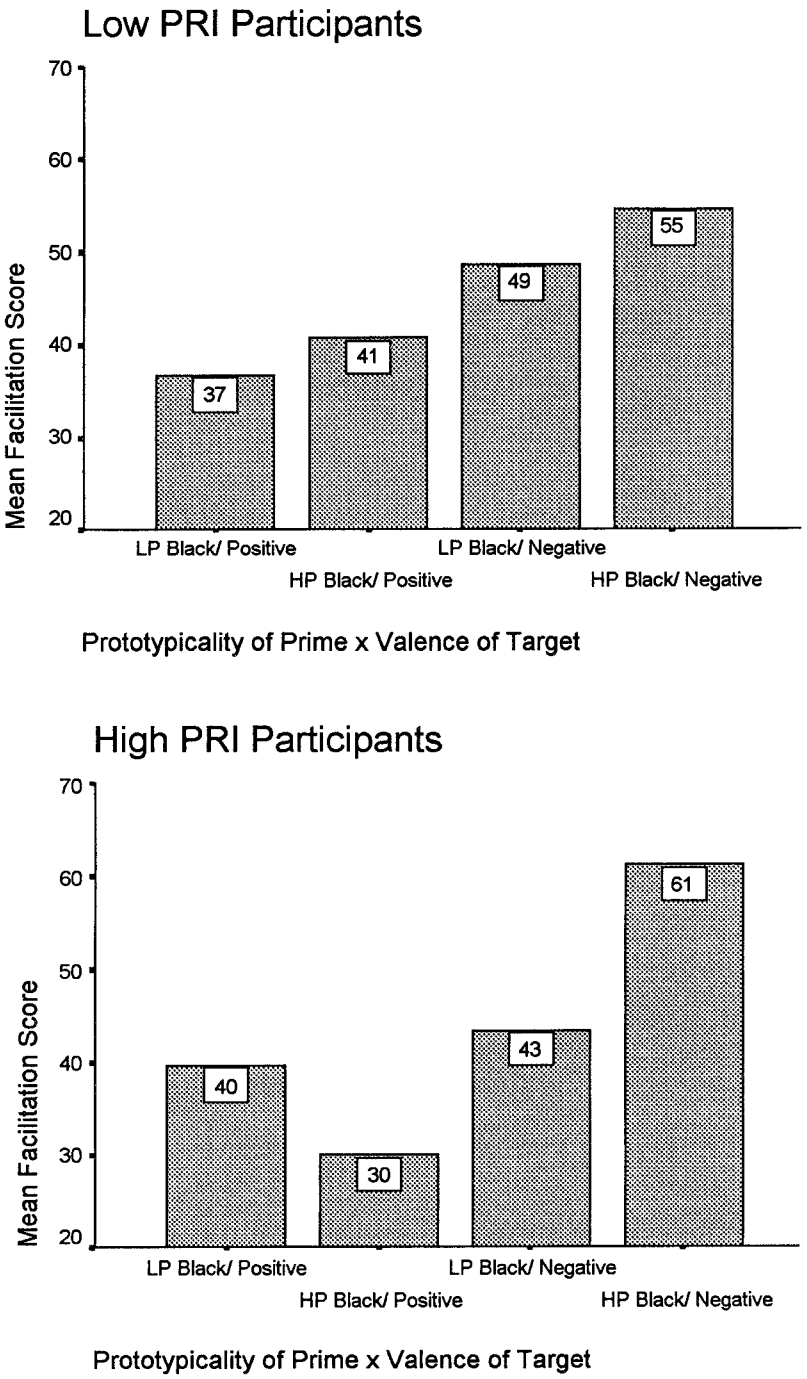

Figure 3. Experiment 5: Facilitation scores for low and high Perceptual Reliance Index (PRI) participants. HP = high prototypic; LP = low prototypic.

processing are independent and (b) affective associations to stimulus features are processed by the perceptual-based system.

Although our findings show repeatedly and consistently that automatic responses to HP faces elicit more negative affect than automatic responses to LP faces, they do not explain why this effect occurs or what consequences this differential cue-based affect might have in everyday interactions with African Americans. Although definitive answers to these questions are clearly beyond the scope of this particular research endeavor, we provide some discussion on possible antecedents and consequences of this effect and directions for future research.

Origins of Perceptual Prejudices

Demonstrating that racially prototypic cues may have affective associations independent of attitudes toward the racial category

itself does not address the question of how such evaluative meaning is acquired. One possible origin of perceptual prejudice is early emotional conditioning, with conditioned responses being strongest in the presence of strong cues and tapering off as the signal gets weaker or more dissimilar to the prototype. This explanation suggests that perceptual prejudices are initially derived from strong prejudices against the category the prototypes represent but may over time become dissociated from beliefs and attitudes about the category as a whole. Consistent with the conditioning explanation, it is certainly the case that most negative media images of African Americans depict HP Blacks, whereas positive images tend to depict LP Blacks (Russell, Wilson, & Hall, 1992). Consistent with the notion of "ambivalence amplification" (Katz & Hass, 1988), the positivity associated with LP Blacks and negativity associated with HP Blacks may become exaggerated, leading to the results of Experiments 1 and 4 (exposure condition). In some cases, the physical prototypicality of real people may be altered to become consistent with these images, such as O. J. Simpson's phenotypic alteration on the cover of *Newsweek* after having been charged for a double murder.

A second possibility is that high levels of perceptual prejudice may simply reflect lack of familiarity with targets, with unfamiliarity eliciting strong anxiety and discomfort (especially for social stimuli) among otherwise well-intentioned, egalitarian individuals. In this case, perceptual prejudice would be expected to decrease as a function of mere exposure (Zajonc, 1968). It can be argued that most White Americans, living in relatively racially segregated environments, are more frequently exposed to images of lighter skinned, less prototypic Blacks (relative to dark-skinned, prototypic Blacks) through media, entertainment, and similar channels. This effect would be further exacerbated if physical appearance and skin color affect individual mobility and opportunity for achievement (Blackwell, 1985; Russell et al., 1992). In that case, the Blacks whom Whites encounter in business, professional, and higher education settings are more likely to be LP rather than HP category members. The result would be that initial perceptual prejudice leads to differential exposure and familiarity (as well as differential associations), which, in turn, perpetuates perceptual prejudice. Although mere exposure provides one explanation for rapid, unconscious negative affect toward high-prototypic Blacks, it does not readily account for the degree of positivity elicited by low-prototypic Blacks.

Interestingly, however, our findings of high positive affect toward LP Blacks is far from novel, historically speaking. In 19th-century Louisiana, prominent White women of New Orleans society rallied together to ban the famed quadroon balls, in which their husbands were alleged to court and sometimes wed the exotic light-skinned Negro women in attendance (Degler, 1971). Moreover, Gilberto Freyre, a famous Brazilian writer and social critic, once wrote that "Brazil is hell for Negroes, purgatory for whites, and heaven for the mulatto" (Freyre, 1986). Freyre was referring not to economic or political privilege, which Whites still clearly enjoy, but rather the physical veneration and interpersonal attraction that LP Blacks command from both Whites and HP Blacks. Moreover, empirical research has shown that averaged faces tend to be more physically appealing than extreme faces (Langlois & Roggman, 1990). To the extent that LP Blacks represent a sort of "averaging" of races, this might explain their higher desirability. In addition, LP faces could be seen by Whites as being "exotic but not too exotic," thereby increasing their desirability (Bem, 1996).

Consequences of Perceptual Prejudice

Basically, what we have demonstrated in the current experiments is that physical features associated with race not only serve as cues to racial category membership but also have affective significance in their own right. According to Allport, facial features may evoke "sensory aversions" or "reflex feelings of dislike or repugnance . . . [many of which] have to do with human physical traits" (Allport, 1954, p. 137). These affective responses to visual facial features occur at early automatic stages of processing, independent of elicitation of category stereotypes. With more extended processing, differences based on physical prototypicality may disappear and category-based beliefs and stereotypes dominate conscious judgments. What is of particular interest is what happens when affective responses based on perceptual features of a particular individual conflict with evaluative responses based on that individual's category membership (as is the case for our Scandinavian acquaintance in her responses to "Mediterranean" men). Do conscious category-based judgments override preconscious evaluations, or does the cue-based affect linger and influence later processing?

In the anecdote described at the beginning of this article, the perceiver was consciously aware of her conflicting responses and of the different sources of her positive and negative feelings. However, in most situations involving perceptions of out-group members, the perceiver will not be aware that affective reactions may arise from different sources. In this case, the negative affect may be misattributed to stereotypic beliefs when it is actually the lingering effect of perceptual prejudice. Conflict between conceptual and perceptual sources of prejudice may be particularly problematic for individuals who are nonprejudiced in terms of their conceptual beliefs about the category of African Americans but still cannot get over the negative affect they experience in the physical presence of specific category members. This is nicely exemplified in a quote from Pettigrew (1987; as cited in Devine, 1989, and Devine, Monteith, Zuwerink, & Elliot, 1991):

> Many Southerners have confessed to me, for instance, that even though in their minds they no longer feel prejudice toward Blacks, they still feel squeamish when they shake hands with a Black. These feelings are left over from what they learned in their families as children. (p. 20)

This dissociation between conceptual-level attitudes toward African Americans and negative affective responses to a physical African American stimulus illustrates how perceptual-based prejudice may exist even for perceivers who hold egalitarian attitudes and consciously disavow prejudice. This may have implications for programs of prejudice reduction that focus almost exclusively on amending negative stereotypic beliefs. Moreover, affective reactions that are appearance based may work against motivation to reduce conceptual prejudice where it does exist. Perceivers may experience less "compunction" in their transgressions toward high-prototypic Black targets if they can rely on positive reactions toward less prototypic Blacks as evidence that they are not prejudiced against African Americans as a group (Devine et al., 1991; Monteith, 1993; Monteith, Devine, & Zuwerink, 1993).

In summary, our results suggest the existence of a dissociation between cue-based and category-based attitudes toward members of the same racial out-group. In addition, these findings begin to address the question of what exactly some automatic priming

paradigms are tapping and how these measures qualitatively differ from other measures of prejudice.

References

Allport, G. W. (1954). *The nature of prejudice.* Reading, MA: Addison-Wesley.

Bargh, J. A. (1989). Conditional automaticity: Varieties of automatic influence in social perception and cognition. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 3–51). New York: Guilford Press.

Bargh, J. A. (1997). The automaticity of everyday life. In R. S. Wyer, Jr. (Ed.), *Advances in social cognition* (Vol. 10, pp. 1–61). Mahwah, NJ: Erlbaum.

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology, 71,* 230–244.

Bem, D. J. (1996). Exotic becomes erotic: A developmental theory of sexual orientation. *Psychological Review, 103,* 320–335.

Blackwell, J. E. (1985). *The black community: Diversity and unity.* New York: Dodd, Mead.

Brewer, M. B. (1988). A dual-process model of impression formation. In T. Srull & R. Wyer (Eds.), *Advances in social cognition* (Vol. 1, pp. 1–36). Hillsdale, NJ: Erlbaum.

Brewer, M. B., & Feinstein, A. (1999). Dual processes in the cognitive representation of persons and social categories. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 255–270). New York: Guilford Press.

Bruner, J. S. (1957). On perceptual readiness. *Psychological Review, 64,* 123–152.

Cameron, J. A., Alvarez, J. M., & Bargh, J. A. (2000). *Examining the validity of implicit and explicit measures of prejudice: Is there really a bona fide pipeline?* Paper presented at the 1st Annual Meeting of the Society of Personality and Social Psychology, Nashville, TN.

Chen, M., & Bargh, J. A. (1997). Nonconscious behavioral confirmation processes: The self-fulfilling consequences of automatic stereotype activation. *Journal of Experimental Social Psychology, 33,* 541–560.

Degler, C. (1971). *Neither black nor white.* New York: Macmillan.

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.

Dovidio, J. F., Evans, N., & Tyler, R. B. (1986). Racial stereotypes: The contents of their cognitive representations. *Journal of Experimental Social Psychology, 22,* 22–37.

Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice Automatic and controlled processes. *Journal of Experimental Social Psychology, 33,* 510–540.

Fazio, R. H., & Dunton, B. C. (1997). Categorization by race: The impact of automatic and controlled components of racial prejudice. *Journal of Experimental Social Psychology, 33,* 451–470.

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

Freyre, G. (1986). The Afro-Brazilian experiment. *UNESCO Courier,* p. 24.

Giner-Sorolla, R., Garcia, M. T., & Bargh, J. A. (1999). The automatic evaluation of pictures. *Social Cognition, 17,* 76–96.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102,* 4–27.

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.

Katz, I., & Hass, R. G. (1988). Racial ambivalence and American value conflict: Correlational and priming studies of dual cognitive structures. *Journal of Personality and Social Psychology, 55,* 893–905.

Kawakami, K., Dion, K. L., & Dovidio, J. F. (1998). Racial prejudice and stereotype activation. *Personality and Social Psychology Bulletin, 24,* 407–416.

Klauer, K. C. (1998). Affective priming. In W. Stoebe & M. Hewstone (Eds.), *European review of social psychology* (pp. 67–103). London: Wiley.

Langlois, J. H., & Roggman, L. A. (1990). Attractive faces are only average. *Psychological Science, 1,* 115–121.

Lepore, L., & Brown, R. (1997). Category and stereotype activation: Is prejudice inevitable? *Journal of Personality and Social Psychology, 72,* 275–287.

Livingston, R. W. (2001). What you see is what you get: Systematic variability in perceptual based social judgment. *Personality and Social Psychology Bulletin, 27,* 1086–1096.

Locke, V., MacLeod, C., & Walker, I. (1994). Automatic and controlled activation of stereotypes: Individual differences associated with prejudice. *British Journal of Social Psychology, 33,* 29–46.

Macrae, C. N., Bodenhausen, G. V., Milne, A. B., Thorn, T. M., & Castelli, L. (1997). On the activation of social stereotypes: The moderating role of processing objectives. *Journal of Experimental Social Psychology, 33,* 471–489.

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

Merikle, P. M. (1992). Perception without awareness. *American Psychologist, 47,* 792–795

Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology, 32,* 89–115.

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., Devine, P. G., & Zuwerink, J. R. (1993). Self-directed vs. other-directed affect as a consequence of prejudice-related discrepancies. *Journal of Personality and Social Psychology, 64,* 198–210.

Pettigrew, T. (1987, May 12). "Useful" modes of thought contribute to prejudice. *New York Times,* pp. 17, 20.

Roediger, H. L. (1990). Implicit memory: Retention without remembering. *American Psychologist, 45,* 1043–1056.

Russell, K., Wilson, M., & Hall, R. (1992). *The color complex: The politics of skin color among African Americans.* New York: Harcourt Brace Jovanovich.

Smith, E. R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology. Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review, 4,* 108–131.

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.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology, 2,* 1–28.

Zarate, M. A., Sanders, J. D., & Garza, A. A. (2000). Neurological dissociations of social perception processes. *Social Cognition, 18,* 223–251.

Received April 17, 2000
Revision received February 15, 2001
Accepted February 28, 2001 ■