

Social and Cognitive Factors Affecting the Own-Race Bias in Whites

Ashlyn E. Slone, John C. Brigham, and Christian A. Meissner

*Department of Psychology
Florida State University*

This study investigated factors associated with the commonly found own-race bias (ORB) in face recognition. We utilized several measures of general face-recognition memory, visual perception and memory, general cognitive functioning, racial attitudes, and cross-race experience in an attempt to distinguish those individuals more likely to demonstrate the effect. White respondents ($N = 129$) were presented two facial-recognition tests (immediate and delayed) involving Black and White faces of both genders. The resulting ORB stemmed largely from a bias to respond “seen before” to Black faces, and produced an effect that was reliable across a 2-day period. An own-sex bias in accuracy was also found. The Benton Facial Recognition test and the Rey–Osterrieth Complex Figure test, 2 central measures of visual memory, were related to ability to recognize White faces. Self-reported amount of recent cross-race experiences was also correlated with overall accuracy on Black and White faces.

Concern about the accuracy of eyewitness identifications has been a central focus of many researchers who have attempted to apply psychological theory and research methodology to the intersection of psychology and law. This concern was first expressed in the psychological literature early in this century (e.g., Feingold, 1914; Munsterberg, 1908; Whipple, 1909, 1912, 1918), but the major research findings have emerged in the past 2 decades, as reviewed by Cutler and Penrod (1995); Huff, Rattner, and Sagarin (1996); Loftus (1979); Ross, Read, and Toglia (1994); Sporer, Malpass, and Koehnken (1996); Wells and Loftus (1984); and Yarmey (1979).

The own-race bias (ORB) has been of particular concern to theorists and researchers. This effect reflects the findings that recognition memory tends to be better for faces of one’s own race than for faces of other races. During the past 25 years a host of studies have investigated this phenomenon, generally in relation to Blacks and Whites. Summarizing this research, a meta-analysis of 14 samples of Whites and Blacks found evidence of a reliable ORB effect, accounting for 11% of the variance in face-recognition performance of Blacks and 10% of the variance of face-recognition performance in Whites (Bothwell, Brigham, & Malpass, 1989). A significant

degree of ORB was visible in roughly 80% of the samples we reviewed. A subsequent meta-analysis by Anthony, Copper, and Mullen (1992) of 15 studies that involved 22 separate hypothesis tests found evidence of a significant weak to moderate tendency for participants to show the ORB. There was a trend for the effect to be stronger among White respondents than among Black respondents.

In a recent review, Chance and Goldstein (1996) asserted that “the number of studies that have replicated the other-race effect is impressive. Few psychological findings are so easy to duplicate” (p. 171). In surveys, most researchers in this area have endorsed the importance and reliability of this effect (Kassin, Ellsworth, & Smith, 1989; Yarmey & Jones, 1983). It is also widely cited in courtroom expert testimony regarding disputed cross-race identifications (Brigham, 1989; Chance & Goldstein, 1996; Leippe, 1995).

Several theoretical bases have been postulated to account for the ORB effect. Over a decade ago, Brigham and Malpass (1985) outlined four reasons that had been put forth to account for the ORB effect: (a) possible group differences in the inherent recognizability of faces of particular races, (b) the influence of ethnocentric attitudes, (c) the experience or contact hypothesis, and (d) differential cognitive processes. They noted that there was no empirical evidence supporting the first point. Concerning the second possibility, that ethnocentric attitudes may cause the ORB, one could speculate that negative attitudes toward another race could lead people to cease processing a face once it has been catego-

rized as belonging to the disliked group (Brigham & Malpass, 1985). Furthermore, one could speculate that negative intergroup attitudes could motivate one to avoid contact with members of the disliked group, or to limit contact to very superficial interactions, thereby constraining the opportunity to develop expertise in distinguishing among other-race faces (Chance & Goldstein, 1996). However, research to date has failed to find any substantial relation between facial-recognition accuracy and attitudes toward the group whose faces are identified (Brigham & Barkowitz, 1978; Lavrakas, Buri, & Mayzner, 1976; Platz & Hosch, 1988; Swope, 1994; Yarmey, 1979). Studies in our lab have, however, found the predicted relation between negative attitudes toward a group and less self-reported contact with members of the disliked group (Brigham, 1993; Swope, 1994). In such studies, more prejudiced Whites and Blacks generally reported less interracial contact both in the past and in present-day interactions than did less prejudiced persons. The final two reasons enumerated by Brigham and Malpass (1985), experience or contact and differential cognitive processing, are discussed next.

It has been widely assumed that one's degree of experience with other-race persons can affect the likelihood that the ORB will occur. Back in 1914, Feingold asserted that it is "well known that, all other things being equal, individuals of a given race are distinguishable from each other in proportion to our familiarity, to our contact with the race as a whole" (p. 50). Many contemporary researchers have also endorsed this race-specific perceptual expertise hypothesis (e.g., Brigham & Malpass, 1985; Chance & Goldstein, 1996; Ng & Lindsay, 1994). However, until very recently, there has been little theoretical attention to why the ORB should be related to cross-race experience, perceptual expertise, or both.

The prediction that experience will affect the ORB is suggested by empirical work on the beneficial attitudinal effects of equal-status contact (e.g., Allport, 1954; Amir, 1969; Cook, 1970). Research findings have led some observers to hypothesize that people who have a greater degree of interracial contact would be less likely to show an ORB in face recognition (Brigham & Malpass, 1985). However, studies have provided only mixed empirical support for the hypothesis. Some studies have found that people who report greater levels of other-race contact are less likely to show an ORB in face recognition (Brigham, Maass, Snyder, & Spaulding, 1982; Carroo, 1986, 1987; Cross, Cross, & Daly, 1971; Lavrakas et al., 1976), although other studies have found no apparent relation (Brigham & Barkowitz, 1978; Malpass & Kravitz, 1969; Ng & Lindsay, 1994; Swope, 1994). Still others have found mixed results within the same study (Chiroro & Valentine, 1996; Platz & Hosch, 1988). It seems apparent that either the theoretical grounds for predicting a relation between contact or experience and face-recognition accuracy are not adequately specified, or that our methods for empirically assessing the quantity, quality, or both of contact or ex-

perience with other-race persons are not adequate (Brigham & Malpass, 1985; Shepherd, 1981).

In addition, we know little about the cognitive mechanisms through which a reduction in ORB due to social experience or perceptual expertise may occur. At a general level, it could be hypothesized that increased contact leads to more experience in processing other-race faces, thereby increasing one's expertise in making such discriminations. One could also speculate that increased contact may reduce the perceived complexity of previously unfamiliar classes of stimuli (Goldstein & Chance, 1971), or that it may convince people that stereotypical responses to other-race faces are not useful and thus instigate the search for more individuating characteristics (Shepherd, 1981). Alternatively, one could predict that increased levels of contact make it more likely that one's social rewards and punishments are dependent on correctly distinguishing among other-race persons with whom one has contact, presumably increasing one's motivation to remember faces of other-race persons accurately (Malpass, 1990).

An early candidate for a cognitive process that may affect the ORB was differential depth of processing (Craik & Lockhart, 1972) for faces of own and other races. Researchers proposed that same-race faces would be cognitively processed at a deeper level, leading to better subsequent recognition (Goldstein & Chance, 1981). However, research findings have not generally supported this hypothesis. Several studies that attempted to manipulate depth of processing via instructions to make superficial (e.g., size of facial features, racial classification) or deep (e.g., friendliness or intelligence) judgments did not find that deeper processing instructions significantly affected the ORB (Devine & Malpass, 1985; Goldstein & Chance, 1981; Sporer, 1991). The most consistent finding appears to be that shallow-processing instructions impair memory for all faces, regardless of race (Chance & Goldstein, 1996).

Another issue involves the dependent measure utilized. Most researchers have employed summary measures of recognition (d' or A') that combine hit and false alarm data. However, this procedure may obscure a potentially important distinction, as there is some evidence that the ORB effect may be largely due to a bias in responding "seen before" to other-race faces. Such a bias would be evident in the pattern of false alarm responses, or in a composite measure of response bias (i.e., β or B''_D). Although most published studies report recognition data only in terms of the d' or A' scores that combine hits and false alarm rates, we were able to find several studies that reported hit and false alarm rates (Carroo, 1987; Cross et al., 1971; Jalbert & Getting, 1992; Malpass & Kravitz, 1969; Ng & Lindsay, 1994; Shepherd, Derogowski, & Ellis, 1974). If one adopts an arbitrary criterion that a "substantial" ORB effect is one in which the ratio of the hit rate for own-race faces to the hit rate for other-race faces is 1.2 or greater, 20% (2:10) of the comparisons yielded a substantial ORB effect for hits. Using the same ratio for false alarms, 70% (7:10) of the comparisons yielded a substantial ORB ef-

fect on false alarms. Similarly, Barkowitz and Brigham (1982) also noted a significant difference in participants' bias to respond, as measured by β , when presented same- and other-race faces. Due to the potential importance of this distinction, we looked at our data in terms a combined measure of accuracy (A') and of response bias (B''_D , as presented by Donaldson, 1992), as well as separate hit and false alarm scores.

In an attempt to shed some light on these issues, we investigated whether, at an individual-difference level, presence of the ORB (in hits, false alarms, A' , or B''_D scores) would be related to other measures of perception, of memory, or of general cognitive functioning. Accordingly, in addition to a standard face-recognition task, individuals were given another measure of face-recognition ability, the Benton Face Recognition Test (BFRT; Benton, 1980), and three measures of visual perception and memory: the Rey–Osterrieth Complex Figure Test (Osterrieth, 1944; Rey, 1942), a digit span test of memory, and a measure of memory for the previous experimental session. General cognitive functioning was assessed by students' grade point average (GPA) and their verbal and quantitative scores on the Scholastic Aptitude Test (SAT). We hypothesized that participants' performance on the face-recognition task, for faces of both races, would be related to the general face-recognition measure (the BFRT), and, to a lesser degree, to the three measures of visual perception and memory. There appears to be no theoretical basis from which to predict whether the general memory measures would correlate with the tendency to show the ORB in face recognition.

Factors that seem, at a theoretical level, more pertinent to the own-race bias, are interracial experiences and racial attitudes. Accordingly, we delved more deeply into the issues of interracial contact and experience (quality, quantity, and variability) and racial attitudes. A 69-item questionnaire assessed the amount of interracial contact that respondents reported in the past and in present everyday encounters, as well as the pleasantness of the experiences. We also attempted to assess the variation in outcomes experienced, based on the ideas of Malpass (1990), who expanded on the social exchange theory of Thibaut and Kelley (1959) with a utilitarian analysis of the ORB. Malpass noted that all social interactions have positive outcomes (benefits) and negative outcomes (costs). He also proposed that people would average the value of all their experiences, and consequently, that they would gauge their future interactions based on this average. For example, individuals who have had repeated positive (beneficial) experiences with a certain person or group of people are likely to pursue further relations with that person or group in the future. He referred to this value distribution as *subjective expected utility* (SEU).

The variability of the central tendency of SEU is the *social utility variance* (SUV). SUV is considered to be large if individuals (e.g., Whites) are exposed to a large range of outcomes with members of another group (e.g., Blacks), ranging

from pleasant to unpleasant, or beneficial to costly. Such large variance should create a strong motivation to distinguish between out-group (Black) faces. Conversely, SUV is defined as small if White individuals experience a small range of outcomes with Blacks and therefore may perceive little need to distinguish between individual Blacks. Malpass (1990) proposed that people would improve in the area of cross-race recognition when they experienced a large SUV with members of the other group. In an initial test of this theory, however, Swope (1994) found no support for this theory, as the measure of SUV of Whites' reported cross-race experiences did not predict participants' ability to identify faces of Blacks. We decided to give this possibility another test in this study. Finally, we administered a measure of racial attitudes to assess whether attitudes were related to degree of ORB, to the other measures of cognitive functioning, or to reported SUV.

Although it was not a central focus of the study, we also looked at gender of participant and gender of stimulus face identified. Some previous research has found evidence of an own-sex bias in face recognition (e.g., Cross et al., 1971; Jalbert & Getting, 1992; Mason, 1986; Shapiro & Penrod, 1986), and that female faces may be better remembered than male faces (Howells, 1938; Lipton, 1977; Shapiro & Penrod, 1986). This study attempted to replicate these findings.

METHOD

Participants

Participants were White undergraduate students ($N = 129$) at Florida State University who were enrolled in an introductory psychology course and received experimental credit for their participation. There were 94 women and 35 men. Participants took part in two sessions, 2 days apart, in one of four groups at each session; group sizes ranged from 22 to 49 individuals. Participants were not permitted to interact with each other and were seated far enough apart such that they could not see one another's written responses. We requested permission from each student to obtain his or her SAT scores and GPAs from the University's Registrar. Most participants (87.5%; $n = 113$) agreed to the release of their SAT scores and GPAs. However, the Registrar's office was able to supply GPAs for only 106 students.

Instruments: Measures of General Memory

Rey–Osterrieth Complex Figure Test. The Rey–Osterrieth Complex Figure Test was devised by Rey (1942) and standardized by Osterrieth (1944) to investigate perceptual organization and visual memory in persons with brain damage. The figure resembles a space ship and

is made of multiple geometric shapes. The test consists of a copy session and delayed recall. In this study, the delayed recall form was scored by two independent raters who utilized the scoring guidelines designed by Osterrieth (1944) and adapted by Taylor (1959). A score of either ½, 1, or 2 was awarded to each of the 18 units into which the figure was divided. Two points were awarded for a correct reproduction; 1 point if misplaced, incomplete, or distorted; and ½ point if recognizable but incomplete and misplaced or distorted. Total scores could range from 0 to 36. Across all items, interrater reliability was acceptable, $r(128) = .89$, $p < .001$. For subsequent analyses, the average of the two raters' scores was used.

Digit span test. Digit spans are a standard part of the Wechsler Adult Intelligence Scale–Revised Test and are widely considered to be an indication of general memory ability. The 10-digit span items were constructed by randomly selecting numbers to build two sets of five to nine digit strings. The numbers were read at a moderate pace and an auditory cue was given after the conclusion of the sequence. Participants, on average, correctly recalled 4.96 (1.63) of the 10-digit sequences on the first administration and 4.96 (1.91) on the second administration. The correlation between the two administrations was significant, $r(129) = .48$, $p < .001$. Scores for the two administrations were summed for later correlational analysis.

Description of experimenters. In Session 2, a 15-item questionnaire inquired about the names and physical appearances of the two Session 1 experimenters, who were seen 2 days earlier, as another measure of general memory. Items were included such as hair color, type of clothing, and accessories worn. Each item had a separate scoring criterion and was assigned a score of 0, 1, or 2. Objective standards for the criteria were carefully developed prior to scoring the responses. Preliminary analyses found 100% agreement between two raters, so a single rater did the final scoring. An item received a 0 score if the individual failed to respond or if their response was completely incorrect, a 1 if the answer was partially correct, and a 2 if the response met the criteria for a correct response. A perfect score would be indicated by a score of 30 for the 15 questions.

Instruments: Face-Recognition Measures

Facial slides of Blacks and Whites. To measure facial memory, we obtained 80 head and shoulder yearbook pictures of 20 Black men, 20 Black women, 20 White men, and 20 White women. The photographs were devoid of eyeglasses, facial hair, distracting jewelry, or any other unusual features. Ten pictures were randomly selected from each race and gender category (Black men, Black women, White men,

and White women) and displayed via slide projector. Slides were initially shown at 3-sec intervals with a 1-sec interstimulus interval. Participants were told they would be asked at a later time to differentiate these 40 slides (referred to as “old”) from 40 facial slides not previously seen (referred to as “new”). At the recognition tests, slides were presented at 5-sec intervals.

Facial-recognition accuracy was initially measured in terms of hits and false alarms. Each participant's recognition performance was thereafter transformed into A' values (see Grier, 1971), a nonparametric analogue of d' (itself a standard measure in signal detection theory; e.g., Banks, 1970), and B''_D values, a nonparametric measure of response bias suggested by Donaldson (1992). Values of A' and B''_D were calculated from the computational formulas provided by Donaldson.¹ Finally, A' and B''_D difference scores (i.e., the participant's score for White faces minus his or her score for Black faces) were calculated, similar to Barkowitz and Brigham (1982) and Brigham and Barkowitz (1978).

BFRT. This test was originally designed to detect possible brain damage to the right cerebral hemisphere as a portion of a neuropsychological assessment battery (Benton & Trauel, 1980). It has been shown in several studies (Benton, 1985; Dricker, Butters, Berman, Samuels, & Carey, 1978; Warrington & James, 1967) that recognition for unfamiliar faces may be impaired by damage to the right hemispheric areas of the brain. Although we are not using it for this purpose, it provides an alternate measure of facial-recognition ability. In the first six items, students were asked to study a single White face at the top of the screen and find that face among six choices below it. For the remaining items, they were asked to study the face on top and find it in three of the six choices below it. All the photos were black and white, and each set varied as to the angle at which each photo was taken and the lighting and shading of the faces.

The original test was in booklet form for individual administration. We presented the sets of photos via slide projector for group administration. A perfect score would be indicated by a score of 54; scores of 37 or less indicate possible brain damage. (Twelve of the participants [9.3%] scored at this level or below; however, it is important to remember that the BFRT norms are based on the individual administration, not the slide projected version used in this article.) Hosch, Bothwell, Sporer, and Saucedo (1990) found that

¹ A' was computed using Equation 1:

$$A' = \frac{1}{2} + \frac{[H - FA](1 + H - FA)}{4H(1 - FA)} \quad (1)$$

B''_D was computed using Equation 2:

$$B''_D = \frac{[(1 - H)(1 - FA) - HFA]}{[(1 - H)(1 - FA) + HFA]} \quad (2)$$

In both equations, H represents the proportion of “hits” and FA represents the proportion of false alarms. Formulas were obtained from Donaldson (1992).

scores on the individually administered BFRT were associated with false alarm rates, $r(31) = .32, p < .05$, but not with hit rates, $r(31) = .09, p > .05$, on a facial recognition task (see also Hosch, 1994).

Instruments: Racial Attitude and Social Experience

Attitude toward Blacks. The Attitude Toward Blacks (ATB) scale was developed by Brigham (1993) from a factor analysis of 90 racial attitude items given to several groups of White college students. To be included in the final scale, items were required to have:

1. A response mean that was not extreme (i.e., not close to 1 or 7).
2. A reasonably large variance ($SD > 1.0$).
3. A substantial factor loading (above .40) on at least one of the first 10 factors extracted.
4. A substantial correlation with the sum of two direct self-evaluative items.

The ATB scale includes 10 positively worded and 10 negatively worded items, and has demonstrated solid reliability (Cronbach's $\alpha = .88$). Individuals responded to questions on a 7-point scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*).

Social Experience Questionnaire (SEQ). The SEQ questionnaire, which was modified from measures developed by Brigham (1993), Malpass (1990), and Swope (1994), contained 56 questions about the individual's exposure to Blacks in business settings, personal-intimate settings, personal-nonintimate settings, and public settings. Participants were also asked about the number of Blacks that attended their schools and grew up in their neighborhoods. The responses to these questions allowed us to look at the extent and quality of participants' cross-race experience. An additional four questions asked about the range of outcomes of interracial interactions in each of the four settings. The seven response categories included: (a) mostly pleasant; (b) mostly neutral; (c) mostly negative; (d) some pleasant-some neutral; (e) some pleasant-some unpleasant; (f) some neutral-some unpleasant; and (g) about the same number of pleasant, neutral, and unpleasant.

Procedure

Session 1. Sessions were 2 days apart and each session lasted approximately 1 hr. At Session 1, students were first shown 40 slides of faces and were told they would be asked to differentiate these from 40 new slides at a later time. After viewing the 40 slides, participants were shown the Rey-Osterrieth Complex Figure on an overhead projector. The students were asked to copy the figure from the overhead

as accurately as possible and were allotted approximately 6 min to do so. The figure remained on the overhead as they worked. Participants then responded to a digit span test of five to nine digits. The auditory cue was given immediately following the conclusion of the sequence, and the importance of not responding early was stressed.

Students were then shown 80 facial slides; the 40 slides seen before (old) were randomly mixed with the 40 slides that had not been previously displayed (new). These 80 slides were presented at 5-sec intervals with a 1-sec interstimulus interval, and students were asked to circle the correct response of *old* or *new*. Following the facial-recognition test, participants were asked to recall the Rey-Osterrieth Complex Figure that they had previously copied from the overhead. They were given approximately the same amount of time (6 min) to reconstruct the figure from memory. It is important to note that the students had not been previously told that they would have to remember the figure, and that they were not permitted to see the figure again before drawing it.

Session 2. At the start of Session 2, which was held 2 days later, a different experimenter administered the 15-item experimenter description survey. Participants had not been informed on the first day that they would be asked to describe their experimenters. Individuals were then asked to identify the 40 facial slides they had originally been shown (old) at the beginning of Session 1. Their task was to distinguish these 40 old faces again from the 40 new faces (as in the recognition phase of Session 1). It should be noted that at this point, all 80 of the slides had been seen before during the testing in Session 1, making the differentiation between old and new more difficult. The slides were presented in the same order, and participants were again asked to circle *old* or *new*.

Participants were then asked to respond to another digit span test. The format for the sequences was identical to the first test, but different numbers were randomly selected. Individuals were next asked to respond to the BFRT, which was presented via slide projector. Each of the items was displayed at approximately 15-sec intervals. The final two measures were the 69-item SEQ and the 20-item ATB survey. Instructions were read as to the importance of their complete honesty, and assurance of anonymity was given. Participants were permitted about 20 min to complete the two surveys. Finally, we requested permission to obtain individual SAT and GPA scores and possibly to contact them in the future, after which they were debriefed about the purpose of the study and thanked for their participation.

RESULTS

Face-Recognition Performance

Face-recognition performance was evaluated first by assessing the proportion of hits and false alarms each participant

committed, after which measures of overall accuracy (A') and response bias (B''_D) were calculated. Performance was then statistically assessed for each measure across the various task manipulations.

Overall accuracy (A' measure). A 2 (Gender of Participant) \times 2 (Gender of Face) \times 2 (Race of Face) \times 2 (Session) repeated measures analysis of variance (ANOVA) was used to investigate differences in face recognition as a function of participants' recognition scores (A'). Table 1 provides the pattern of A' means across the various conditions. A significant four-way interaction was obtained, $F(1, 127) = 4.10, p < .05$ ($MSE = .01, \eta^2 = .03$), in addition to a three-way interaction involving Session \times Race of Face \times Gender of Face, $F(1, 127) = 6.12, p < .05$ ($MSE = .01, \eta^2 = .05$). The four-way interaction was further decomposed by testing the mean accuracy for White versus Black faces at each combination of levels of Gender of Face, Gender of Participant, and Session. Results indicated that female participants were more likely to demonstrate the ORB, with significantly better performance on White male faces at both Sessions 1 and 2, $t(93) > 2.85, ps < .01$, and on White female faces at delayed testing, $t(93) = 2.81, p < .01$. The performance difference for female faces at Session 1 was not significant, $t(93) = 1.58, p > .05$. Male participants did not show any performance differences as a function of the race of the stimulus face, $t(34) < 1.70, ps > .05$.

Several two-way interactions were also significant in the overall analysis: Race of Face \times Gender of Face, $F(1, 127) = 6.65, p < .05$ ($MSE = .02, \eta^2 = .05$); Gender of Face \times Gender of Participant, $F(1, 127) = 9.24, p < .05$ ($MSE = .02, \eta^2 = .07$); and Race of Face \times Gender of Participant, $F(1, 127) = 10.25, p < .05$ ($MSE = .02, \eta^2 = .08$). The Gender of Face \times Gender of Participant interaction replicated the own-sex bias for women found previously by some researchers (Cross et al., 1971; Jalbert & Getting, 1992; Mason, 1986). As displayed in Figure 1, female participants performed significantly better than male participants on female faces, $t(127) = 3.43, p < .001$, but no different on male faces, $t(127) = .17, p > .05$.

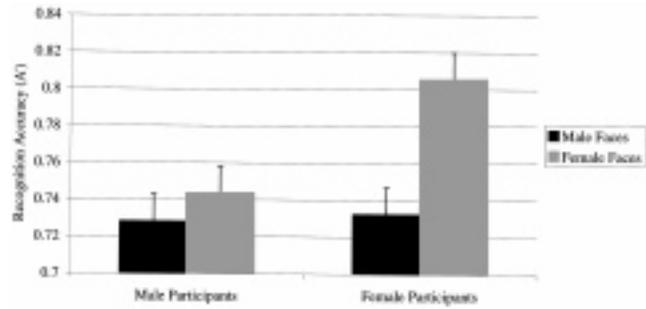


FIGURE 1 Recognition accuracy (A') as a function of gender of participant and gender of stimulus face.

Several main effects were also observed in the overall analysis. Performance was significantly better on the immediate versus the delayed testing session, $F(1, 127) = 116.50, p < .001$ ($MSE = .02, \eta^2 = .48$; Session 1: $M = .80$; Session 2: $M = .70$). In addition, female faces were recognized significantly better than male faces, $F(1, 127) = 21.49, p < .001$ ($MSE = .02, \eta^2 = .15$; male faces: $M = .73$; female faces: $M = .77$).

Response bias (B''_D). We also examined differences in participants' response bias during the facial recognition task. We utilized B''_D as a measure of bias (see Donaldson, 1992) in which negative numbers represent liberal bias (i.e., a loose criterion in which participants tend to respond with "seen before" or *old*) and positive numbers represent conservative bias (i.e., a strict criterion in which participants tend to respond with "never seen before" or *new*). Paralleling the previous analysis, a 2 (Gender of Participant) \times 2 (Gender of Face) \times 2 (Race of Face) \times 2 (Session) repeated measures ANOVA was employed. Table 2 provides the pattern of B''_D means across the various conditions. Results indicated a significant Race of Face \times Gender of Face interaction, $F(1, 127) = 14.91, p < .001$ ($MSE = .30, \eta^2 = .11$). As displayed in Figure 2, participants held a significantly more conservative criterion for Black female faces compared with Black male faces, $t(128) = 5.30, p < .001$.

TABLE 1 Accuracy in Face Recognition: Mean A' Scores as a Function of Gender of Participant, Gender of Face, Race of Face, and Session

		Male Faces				Female Faces			
		White		Black		White		Black	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Session 1	Male participants ^a	.78	.13	.79	.14	.80	.12	.81	.13
	Female participants ^b	.80	.13	.75	.14	.86	.08	.84	.12
Session 2	Male participants ^a	.69	.19	.65	.19	.64	.30	.73	.14
	Female participants ^b	.73	.13	.65	.20	.79	.12	.74	.15

^a $n = 35$. ^b $n = 94$.

TABLE 2
Response Bias in Face Recognition: Mean B''_D Scores as a Function of Gender of Participant, Gender of Face, Race of Face, and Session

	Male Faces				Female Faces			
	White		Black		White		Black	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Session 1 Male participants ^a	.31	.47	.07	.46	.42	.50	.50	.47
Female participants ^b	.34	.46	.08	.57	.48	.52	.40	.49
Session 2 Male participants ^a	.20	.60	-.06	.47	.15	.59	.23	.56
Female participants ^b	.26	.53	-.13	.55	.28	.51	.25	.62

^a*n* = 35. ^b*n* = 94.

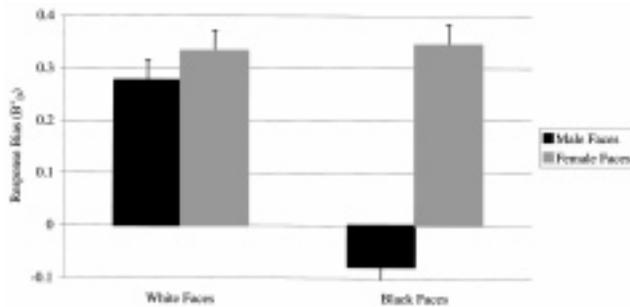


FIGURE 2 Response bias (B''_D) as a function of gender and race of stimulus face.

White male and female faces, however, did not differ in the level of response criterion, $t(128) = .02, p > .05$.

In addition, several main effects were obtained, including Race of Face, $F(1, 127) = 10.87, p < .001$ ($MSE = .35, \eta^2 = .08$); Gender of Face, $F(1, 127) = 27.95, p < .001$ ($MSE = .30, \eta^2 = .18$); and Session, $F(1, 127) = 30.72, p < .001$ ($MSE = .21, \eta^2 = .20$). Participants maintained a more conservative response criterion when judging White faces compared with Black faces ($M_s = .31$ and $.17$, respectively), thus replicating the ORB for response criterion found previously by others. Additionally, a more conservative criterion was used when judging female faces compared with male faces ($M_s = .34$ and $.13$, respectively), and when being tested immediately compared with a delayed session ($M_s = .33$ and $.15$, respectively).

Hits and false alarms. Previous research has suggested that the ORB may stem largely from false alarm responses. Therefore, an aggregate measure of accuracy, such as A' , which utilizes both hit and false alarm rates, may obscure the effect. This appeared possible in this study, as the ORB effect did not reach significance in the A' analyses, $F(1, 127) = 2.53, p < .11$ ($MSE = .02, \eta^2 = .02$; White faces: $M = .78$; Black faces: $M = .76$). To assess whether the ORB occurred in the false alarm responses only, we analyzed the pro-

portion of hits and false alarms committed by participants. As predicted, a significant race of face main effect was found for false alarms, $F(1, 127) = 15.09, p < .001$ ($MSE = .03, \eta^2 = .11$), as participants generated significantly fewer false alarms when judging White faces compared with Black faces ($M_s = .24$ and $.29$, respectively). Analysis of the proportion of hits failed to yield a significant main effect for race of face, $F(1, 127) = 3.30, p < .07$ ($MSE = .03, \eta^2 = .03$; White faces: $M = .65$; Black faces: $M = .63$). Additionally, the own-sex bias was significant in the analyses of both hits, $F(1, 127) = 4.65, p < .05$ ($MSE = .03, \eta^2 = .03$; male participants–male faces: $M = .64$; female faces: $M = .60$; female participants–male faces: $M = .64$; female faces: $M = .65$), and false alarm responses, $F(1, 127) = 4.98, p < .05$ ($MSE = .03, \eta^2 = .04$; male participants–male faces: $M = .30$; female faces: $M = .24$; female participants–male faces: $M = .31$; female faces: $M = .20$).

Overall, the analysis of participants' identification accuracy yielded several interesting findings. First, an own-sex bias for females was found for measures of accuracy (A' and proportion of hits) and on one measure of response bias (proportion of false alarms). This finding replicates that of earlier studies within the facial memory literature. Second, although the ORB was not significant across the overall accuracy measures (A' and proportion of hits), it was significant in both measures of response bias (B''_D and proportion of false alarms), replicating the results of previous studies (cf. Barkowitz & Brigham, 1982; Shapiro & Penrod, 1986). It appears that the ORB may be a function of individuals' tendency to respond "seen before" to other-race faces, rather than a pure accuracy-based phenomenon. This is of particular importance, as false alarm responses are precisely those that concern the possibility of an individual being falsely convicted on the basis of erroneous eyewitness identification.

Reliability of Effects

We also assessed the reliability of recognition scores across Sessions 1 and 2, an issue that has not been previously ad-

dressed. Overall face-recognition performance scores (A') were substantially related across Sessions 1 and 2, $r(127) = .61, p < .001$. A' scores for White faces were slightly but not significantly ($Z = .98, p > .05$) more reliable across sessions, $r(127) = .56, p < .001$, than were A' scores for Black faces, $r(127) = .44, p < .001$. The magnitude of the ORB (A' for White faces – A' for Black faces) was somewhat reliable across the 2-day interval, $r(127) = .21, p < .05$. However, this was significantly lower than the reliability of the total A' scores ($Z = 3.25, p < .001$), the reliability of A' for White faces ($Z = 3.03, p < .001$), and the reliability of A' for Black faces ($Z = 1.90, p < .06$; see Steiger, 1980, for modified Pearson–Filon Test for equality of two dependent correlations).

Overall response bias scores (B''_D) significantly correlated across the two sessions, $r(127) = .44, p < .001$. B''_D scores for Black faces were significantly more reliable, $r(127) = .52, p < .001$, than were B''_D scores for White faces, $r(127) = .25, p < .001$ ($Z = 2.15, p < .05$). Finally, the magnitude of the ORB in response bias (B''_D for White faces – B''_D for Black faces) was also reliable across the 2-day interval, $r(127) = .34, p < .001$.

We also examined the relation between each measure of accuracy (A') and response bias (B''_D) aggregated across Sessions 1 and 2 (see Table 3). Participants' scores on White faces and Black faces were significantly related to overall scores, with correlations ranging in magnitude from $r(129) > .22$ to $r(129) < .88, ps < .05$. Additionally, participants' A' scores on Black faces were significantly related to B''_D scores on Black faces, as well as to overall and ORB B''_D scores, all $rs(129) > .18, ps < .05$.

Face Recognition and Measures of Memory and Cognitive Functioning

Correlational analysis of the memory and cognitive measures indicated that scores on the measures designed to assess academic performance (i.e., SAT and GPA) were related to each other. GPA was significantly related to SAT scores, including SAT verbal, $r(127) = .31, p < .01$, SAT quantitative, $r(127) = .30, p < .01$, and SAT total score, $r(127) = .38, p < .01$. Due to the related nature of these measures, this relation was not surprising. However, almost none of the other memory measures significantly correlated with one another (see Table 4).

Next, we assessed the relation of these memory measures to the overall accuracy (A') and response bias (B''_D) measures on the facial recognition task (averaged across the two sessions for each participant). The two measures that dealt most directly with visual memory, the BFRT and the Rey–Osterrieth Complex Figure Test, showed some relation to facial recognition performance (see Table 5). Respondents who performed well on the BFRT also tended to do well when identifying faces: A' for white faces, $r(127) = .36, p < .01$; A' for overall performance, $r(127) = .28, p < .01$; and A' difference score (A' for White faces – A' for Black faces),

$r(127) = .23, p < .01$. Interestingly, the BFRT did not correlate significantly with individuals' performance on Black faces. Participants' response bias scores (B''_D) were also related to their performance on the BFRT: for White faces, $r(127) = .18, p < .05$; for Black faces, $r(127) = .18, p < .05$; and for overall performance, $r(127) = .23, p < .05$. Thus, higher scores on the BFRT were associated with a more conservative response criterion when responding in the facial recognition task.

Scores on the Rey–Osterrieth Complex Figure Test were also significantly correlated with recognition accuracy (A'): for White faces, $r(128) = .26, p < .01$; for Black faces, $r(128) = .19, p < .05$; and for overall performance, $r(128) = .26, p < .01$. The Rey–Osterrieth Complex Figure Test, however, was not related to participants' response bias on the face-recognition task. The only other measure that related to overall accuracy was the accuracy of description for the (White) experimenters, $r(127) = .23, p < .01$.²

Face Recognition, Racial Attitudes, and Social Experiences

The 56 SEQ items were subjected to an orthogonal factor analysis (varimax rotation) that yielded six meaningful factors. The first factor, which we labeled *present contacts*, contained 11 items (factor loadings of .40 or above) that asked about the amount of everyday contact with Blacks on campus, in recreational activities, in dorms or apartment complexes, in stores, as a friend or dating partner, and so forth (eigenvalue = 12.32, 22.0% of total variance accounted for). The second factor, labeled *public- and personal-nonintimate settings*, contained eight items that asked how pleasant and how costly their past interactions were with Blacks in public settings and personal-nonintimate settings, as well as expectations for future interactions in these settings (eigenvalue = 5.64, 10.1% of variance). The third factor, *past contact*, contained six items that asked about the extent of past contact with Blacks in schools and in the neighborhood (eigenvalue = 3.05, 5.5% of variance). The fourth factor, labeled *business settings*, contained five items about the pleasantness and cost of past interactions with Blacks in business settings (eigenvalue = 2.55, 4.5% of variance). The fifth factor, *intimate settings*, contained four questions on the pleasantness and cost of past interactions, and expected future interactions, in intimate settings (eigenvalue = 2.00, 3.6% of variance). The sixth factor, *past*

²We also assessed the relation between participants' gender and measures of memory and general cognitive ability. Results indicated that women tended to outperform men on the experimenter description task, $t(126) = -3.91, p < .001$ ($M_s = 15.85$ and 12.56 ; $SD_s = 4.38$ and 3.66 , respectively), whereas men performed better on the quantitative portion of the SAT, $t(98) = 3.27, p < .001$ ($M_s = 583.85$ and 530.81 ; $SD_s = 68.36$ and 71.95 , respectively), and on the SAT total score, $t(101) = 2.01, p < .05$ ($M_s = 1130.38$ and 1074.94 ; $SD_s = 128.95$ and 119.39 , respectively).

TABLE 3
Correlations Between Measures of Face Recognition Accuracy (A') and Response Bias (B''_D)

	1	2	3	4	5	6	7	8
1. A' (White)	—	.54***	.87***	.46***	.12	.09	.13	.01
2. A' (Black)		—	.88***	-.50***	.08	.27**	.24**	-.18*
3. A' (Total)			—	-.04	.11	.21*	.21*	-.10
4. A' (White-Black)				—	.03	-.20*	-.12	.20*
5. B''_D (White)					—	.22*	.73***	.52***
6. B''_D (Black)						—	.83***	-.72***
7. B''_D (Total)							—	-.20*
8. B''_D (White-Black)								—

* $p < .05$. ** $p < .01$. *** $p < .001$.

TABLE 4
Correlations Between Measures of Memory and Cognitive Functioning

	1	2	3	4	5	6	7	8
1. Benton Facial Recognition Test	—	-.04	.14	.07	.03	-.08	-.03	.09
2. Rey-Osterrieth Complex Figure Test		—	-.01	.10	.00	.08	.00	-.07
3. Description of experimenters			—	-.16	.08	-.10	.02	-.01
4. Grade point average				—	.31**	.34***	.38***	.08
5. Scholastic Aptitude Test-Verbal					—	.51***	.84***	.22*
6. Scholastic Aptitude Test-Quantitative						—	.85***	.11
7. Scholastic Aptitude Test-Total							—	.17
8. Digit span test								—

Note. $ns = 95-128$.

* $p < .05$. ** $p < .01$. *** $p < .001$.

TABLE 5
Correlations Between Measures of Memory and Cognitive Functioning and Recognition Performance

	Benton Facial Recognition Test	Rey-Osterrieth Complex Figure Test	Digit Span Test	Description of Experimenters	Grade Point Average	Scholastic Aptitude Test-Verbal	Scholastic Aptitude Test-Quantitative	Scholastic Aptitude Test-Total
A' (White)	.36***	.26**	-.05	.10	.16	.05	-.14	-.03
A' (Black)	.13	.19*	.04	-.12	.13	-.01	.00	.03
A' (Total)	.28**	.26**	-.01	-.01	.17	.02	-.08	.00
A' (White-Black)	.23**	.06	-.09	.23**	.02	.06	-.14	-.05
B''_D (White)	.18*	-.02	-.09	-.13	.14	.05	.13	.10
B''_D (Black)	.18*	-.09	.08	-.09	-.02	-.02	-.01	.02
B''_D (Total)	.23*	-.08	.00	-.14	.07	.01	.07	.06
B''_D (White-Black)	-.03	.07	-.13	-.02	.12	.06	.11	.06

Note. $ns = 100-128$.

* $p < .05$. ** $p < .01$. *** $p < .001$.

friends, contained four items about the number of Black friends in the school and the neighborhood (eigenvalue = 2.00, 3.6% of variance).

SUV scores were calculated from responses to the four questions in the SEQ that asked about the range of outcomes in the four settings (public, business, personal–intimate, and personal–nonintimate). Responses of *mostly pleasant*, *mostly neutral*, or *mostly unpleasant* experiences were scored as 1 (low SUV). Responses of *some pleasant–some neutral*, *some pleasant–some unpleasant*, or *some unpleasant–some neutral* outcomes were scored as 2 (moderate SUV). The final response, *about the same amount of pleasant, neutral, or unpleasant* interactions was scored as 3 (high SUV). Thus, total scores across the four questions could range from 4 (low SUV) to 12 (high SUV). Intimate personal settings received the largest number of *mostly pleasant* responses (48%) and public settings received the largest number of *mostly unpleasant* responses (10%). The modal response across all four settings was *mostly pleasant* (35%). The high SUV response (coded as 3) was chosen only 7.5% of the time.

Scores on each of the six SEQ factors, created by summing responses to the items contained in that factor, were related to racial attitudes (ATB) to a substantial and very similar degree, $r_s = .34$ to $.48$, all $p_s < .01$ (see Table 6). Total SEQ scores were also strongly correlated with ATB scores, $r(63) = .61$, $p < .01$. Hence, participants with relatively low prejudice (positive racial attitudes) reported more past contact with Blacks, more pleasant and less costly past experiences, and anticipation that their future contact would be more pleasant and less costly as well. The SUV score, in contrast, related only to the *public– and personal–nonintimate settings*, $r(86) = -.22$, $p < .01$, and *business settings*, $r(86) = -.34$, $p < .01$, factors of the SEQ, and was not related significantly to the racial attitudes measure (ATB). The negative correlations indicate that those individuals who reported less variance in their interracial experiences (low SUV) also reported more positive encounters with Blacks in business settings, and public and personal–nonintimate settings.

We also assessed whether any of the racial experience and attitude measures were correlated with individuals' recognition accuracy (A') and response bias (B''_D) measures (combined across administrations). The main finding involved the significant correlation between the *present contacts* factor of the SEQ and individuals' A' scores for Black faces, $r(120) = .18$, $p < .05$, for White faces, $r(120) = .21$, $p < .05$, and for overall performance, $r(120) = .22$, $p < .05$. Individuals who reported a higher level of recent contact with Blacks performed better on the facial recognition task for Black faces and for White faces as well. The other significant finding was that individuals who reported a high degree of variance in the quality of their cross-race social interactions (high SUV) held a more conservative response bias when responding on the recognition task (B''_D total), $r(90) = .23$, $p < .05$. Table 7 provides the pattern of correla-

tions between the measures of racial experience and attitudes, and recognition performance.³

Metamemory Judgments

We also attempted to gather more in-depth information regarding the potential knowledge that skilled individuals may have of their own memory abilities. Individuals ($n = 14$) who performed very well (upper quartile) on Black faces in the recognition task were recontacted. In a telephone interview, we asked these individuals if they could think of any reasons why they were better at recognizing Blacks than most students were, or if they had ever been in a position (e.g., job or group membership) where it was important that they be able to distinguish accurately between different Black people. We also asked them if there was anything else they could think of that might explain their excellent performance. Unfortunately, the responses were not particularly informative. The most common responses were “I have a good memory” or “I was really paying attention,” and many individuals expressed surprise when they were informed that they performed very well. Several students speculated that their job experiences, such as working in a restaurant, bar, or retail store, may have aided them in recognizing faces of Blacks.

DISCUSSION

Overall Facial Memory

These results replicate the central findings of an ORB in facial recognition that stems largely from a response bias to other-race faces. Consistent with past research, participants recognized other-race faces at about the same rate as own-race faces (similar A' and hit rates; Cohen's $d = .22$), but were more likely to say they had previously seen an other-race face when they had not (more liberal B''_D scores and higher false alarm rate; Cohen's $d = .40$). This result was a moderate to large effect, indicating a degree of practical significance. Indeed, from a psycholegal perspective, the increased potential for false identification of a cross-race suspect has dramatic implications. It suggests that both law enforcement and judicial officials should remain cautious when examining identifications made under cross-racial condi-

³The relation between participants' gender and measures of social experience was assessed. Results indicated no significant differences in individuals' responses to the SEQ, SUV, or ATB measures as a function of gender, all $t_s < .86$, $p_s > .05$ ($\eta < .03$). Finally, we examined the pattern of correlations between the measures of memory and cognitive functioning and the measures of racial experience and attitudes. No significant relations were found, with correlations ranging in magnitude ($r_s > .01$ – $r_s < .18$, with all $p_s > .05$).

TABLE 6
Correlations Between Measures of Racial Experience and Attitudes

	1	2	3	4	5	6	7	8	9
1. Present	—	.30***	.45***	.27**	.31**	.54***	.86***	.13	.47***
2. Public		—	.24**	.59***	.43***	.32***	.71***	-.22*	.48***
3. Contact			—	.24*	.33**	.72***	.75***	.04	.34***
4. Business				—	.43***	.30**	.55***	-.34***	.42***
5. Intimate					—	.36***	.60***	-.16	.42***
6. Friends						—	.77***	.01	.38***
7. SEQ							—	-.04	.61***
8. SUV								—	-.07
9. ATB									—

Note. $n_s = 76-121$. SEQ = Social Experiences Questionnaire; SUV = Social Utility Variance; ATB = Attitudes Toward Blacks. Present, public, contact, business, intimate, and friends represent six social experience factors.

* $p < .05$. ** $p < .01$. *** $p < .001$.

TABLE 7
Correlations Between Measures of Racial Experience and Attitudes, and Recognition Performance

	Present	Public	Contact	Business	Intimate	Friends	SEQ	SUV	ATB
A' (White)	.21*	-.11	.02	-.10	-.02	.01	.08	-.07	-.04
A' (Black)	.18*	-.08	.07	-.10	-.04	.06	.07	.07	-.12
A' (Total)	.22*	-.11	.05	-.13	-.03	.04	.09	.00	-.09
A' (White-Black)	.02	-.02	-.05	.04	.02	-.05	-.01	-.14	.09
B''_D (White)	.03	-.08	.04	.01	-.03	-.08	.02	.19	.02
B''_D (Black)	.13	-.03	.01	-.13	-.11	-.07	.01	.16	-.16
B''_D (Total)	.10	-.07	.03	-.08	-.10	-.09	.02	.23*	-.10
B''_D (White-Black)	-.09	-.03	.02	.12	.08	.01	.01	-.01	.15

Note. SEQ = Social Experiences Questionnaire; SUV = Social Utility Variance; ATB = Attitudes Toward Blacks.

* $p < .05$.

tions and seek to substantiate the eyewitness's claim with additional forensic evidence.

The reliability of the ORB effect was also assessed for the first time in a published study. It was found that individuals' ORB score (A' for White faces – A' for Black faces) was, in fact, reliable across the two testing sessions. However, this degree of reliability was significantly smaller than were the reliabilities of A' (White) scores, A' (Black) scores, or A' (total) scores across Sessions 1 and 2. Individuals' response bias scores (B''_D) were also reliable across the delay, including the response bias difference score (B''_D for White faces – B''_D for Black faces). Again, practically speaking, the presence of the ORB across a 2-day delay has important implications for the assessment of eyewitness memory in the field setting.

In addition to the ORB effect, an own-sex bias in accuracy was found, emanating largely from the responses of female participants. This finding has been demonstrated previously

within the facial memory literature (Cross et al., 1971; Jalbert & Getting, 1992; Mason, 1986). In general, female faces were also better remembered than male faces across Sessions 1 and 2, also replicating several previous studies (Howells, 1938; Lipton, 1977; Shapiro & Penrod, 1986). This effect was obtained on measures of overall accuracy (Cohen's $d = .60$) and response bias (Cohen's $d = .65$), both fairly large effects.

As a word of caution, it should be noted that the same photos were employed at encoding and at test. Thus, it is possible that some recognition judgments were made on the basis of cues in the photo, rather than in the faces per se. However, all but one (Chiroro & Valentine, 1996) of the previous studies of the ORB reviewed in this article also used the same photos at encoding and at test. Regardless, it will be valuable in future studies to use different photos of the same individuals, so that this potential confound is eliminated.

Individual Differences in Facial Memory

The focus of this study was to examine whether several measures of cognitive ability or social experience may predict the occurrence of the ORB, or overall facial memory ability in general. Individuals' responses to the two visual memory tests, the BFRT and the Rey–Osterrieth Complex Figure Test, were significantly related to the recognition of White faces, and to the recognition of all faces. Scores on the BFRT also related to the degree of response bias that individuals demonstrated for all faces, and for the magnitude of the ORB. This pattern of results raises the question of whether the processing method that Whites use to encode or remember faces of Blacks is somehow different from that used for faces of Whites, or for other complex figures. For example, there is some evidence that expert-level stimuli (i.e., own-race faces) may be processed in a more holistic or configural manner (Diamond & Carey, 1986; Rhodes, Brake, Taylor, & Tan, 1989).

A previous study (Hosch et al., 1990) had found that scores on the BFRT were predictive of eyewitness identification accuracy in a photographic lineup following live exposure to the target. In a similar vein, scores on the BFRT in this study predicted overall accuracy to White faces and response bias to Black faces. Such results further demonstrate the influence of the ORB, as associated with individuals' perceptual expertise with own-race faces. It may provide an interesting insight if a standardized BFRT using Black faces were also available.

The popular, but weakly supported, differential experience hypothesis again received rather weak support in this study. Only the *present contacts* factor of the SEQ was related to participants' overall accuracy on cross-race faces. Those who reported more recent contact with members of the other race also demonstrated increased recognition accuracy for cross-race faces; however, this relation was not obtained for the ORB score, or for any of the response bias measures. Using an analogous measure of recent contact with other-race persons, Byatt and Rhodes (1998) also found a significant correlation ($r = .48$) with other-race face-recognition accuracy. Interestingly, we were unable to find that SUV was a useful predictor of performance on facial recognition tasks (see also Swope, 1994). Furthermore, although racial attitudes were not predictive of facial recognition accuracy, they were again related to self-reported contact with other-race individuals (cf. Brigham, 1993; Swope, 1994).

A recent study by Dunning, Li, and Malpass (1998) indicated that differential experience may yet play a role in memory for cross-race faces, but at a more qualitative level. Dunning and colleagues found that White individuals who had more interest and experience in National Basketball Association basketball players were better at cross-race identification than were individuals with less knowledge or interest in basketball. The authors concluded that the ability to recognize other-race faces may have to do with the more qualita-

tive aspects of prior experience, especially the ability to individuate members of the other race. The *present contacts* factor of the SEQ appears to assess similar aspects of recent experience and individuation. Future research may inquire more qualitatively about sets of concrete experiences or memorable events that took place with members of the other race. Furthermore, researchers may attempt to obtain information about past experiences that can be independently corroborated, and hence not rely on individuals' self-reported memories of the events.

Future Directions: Cognitive and Perceptual Factors

Although previous studies and these results suggest that attention to "social" aspects of interracial contacts (e.g., amount and quality of interracial experiences) may provide a minimal degree of predictive validity, such a focus may not be the most fruitful avenue for explaining the ORB. Rather, analyses in terms of cognitive and perceptual factors may prove a more useful path. For example, the finding that the ORB largely results from a response bias to novel cross-race stimuli raises some interesting issues.

First, this finding seems to contradict the *mirror effect*, the widely found tendency for memory performance on new items (not seen previously) to mirror performance on old (seen-before) items. Glanzer and Adams (1985) reviewed over 80 studies that showed this effect, and further explicated a theoretical model (Glanzer & Adams, 1990). Based on their theory, we would expect hit rates to be lower and false alarm rates higher, to about the same degree, for other-race faces. However, this was not the case in this study.

Second, the response bias effect in the ORB may implicate the mediating role of facial typicality in eliciting increased false alarm responses to novel stimuli (see Hosie & Milne, 1996). For example, O'Toole, Deffenbacher, Valentin, and Abdi (1994) applied Vokey and Read's (1992) concepts of facial *familiarity* and perceived *memorability* to the cross-race paradigm. Their results indicated that high typicality of a face hindered performance more for cross-race faces than for same-race faces. Similar to that of Valentine's purely exemplar-based model (Valentine, 1991; Valentine & Endo, 1992), the findings of O'Toole et al. (1994) can be applied as a more comprehensive approach to modeling cross-race identification.

To illustrate, individuals' representational systems may be defined or clustered (via perceptual attributes of familiarity or memorability) based on their prior perceptual and social experiences with other-race faces. Thus, individuals with a limited range of cross-racial experiences would maintain an ill-defined, dense clustering of cross-race faces within their representational system (see Valentine, 1991; Valentine & Endo, 1992). As a result of this clustering effect, such indi-

viduals would not only have difficulty encoding cross-race faces, but would also maintain more of a bias to respond "seen before" to novel cross-race stimuli due to their high familiarity (e.g., "They all look alike"), but low memorability (typical appearance or lack of distinctive attributes in memory). When combined with the notion of configural and featural aspects of the face, and perceptual expertise (see Diamond & Carey, 1986), such a comprehensive model of facial memory could add significantly to our ability in both predicting cross-race face memory and applying our understanding to the realm of eyewitness memory.

In closing, these results indicate that the ORB resides largely in individuals' false alarm responses to novel cross-race stimuli. This effect was found to be reliable across a 2-day delay. These results carry important implications for caution in the assessment of an eyewitness's cross-racial identification. Furthermore, individual difference measures yielded minimal support for the influence of social experience, but more so for the predictive validity of individuals' perceptual memory abilities. Future research should examine the more qualitative aspects of cross-racial social experience (e.g., individuation) and general stimulus characteristics (e.g., typicality) to develop a more comprehensive model of facial memory and cross-racial identification accuracy.

ACKNOWLEDGMENTS

This study was conducted as an undergraduate honors thesis by Ashlyn E. Slone under the supervision of John C. Brigham.

Portions of the research were presented at the International Conference on Memory in Padua, Italy, July 1996.

We appreciate the comments of Daniel Cole, William Doerner, James Meyer, Don Read, and Adina Wasserman on earlier versions of this article. We also appreciate the help of David Hill and Wendy Young during the data-gathering phase.

REFERENCES

- Allport, G. W. (1954). *The nature of prejudice*. New York: Doubleday.
- Amir, Y. (1969). Contact hypothesis in ethnic relations. *Psychological Bulletin*, *71*, 319–341.
- Anthony, T., Copper, C., & Mullen, B. (1992). Cross-racial facial identification: A social-cognition integration. *Personality and Social Psychology Bulletin*, *18*, 296–301.
- Banks, W. M. (1970). The changing attitudes of Black students. *Personnel & Guidance Journal*, *48*, 739–745.
- Barkowitz, P., & Brigham, J. C. (1982). Recognition of faces: Own-race bias, incentive, and time delay. *Journal of Applied Social Psychology*, *12*, 255–268.
- Benton, A. L. (1980). The neuropsychology of facial recognition. *American Psychologist*, *35*, 176–186.
- Benton, A. L., & Trauel, D. (1985). Visuo-perceptual, visuo-spatial, and visuo-constructive disorders. In K. M. Heilman & E. Valenstein (Eds.), *Clinical neuropsychology* (2nd ed., pp. 165–213). New York: Oxford University Press.
- Bothwell, R. K., Brigham, J. C., & Malpass, R. S. (1989). Cross-racial identifications. *Personality and Social Psychology Bulletin*, *15*, 19–25.
- Brigham, J. C. (1989). Disputed eyewitness identifications: Can experts help? *The Champion*, *8*(5), 10–18.
- Brigham, J. C. (1993). College students' racial attitudes. *Journal of Applied Social Psychology*, *23*, 1933–1967.
- Brigham, J. C., & Barkowitz, P. (1978). Do "they all look alike?" The effect of race, sex, experience, and attitudes on the ability to recognize faces. *Journal of Applied Social Psychology*, *8*, 306–318.
- Brigham, J. C., Maass, A., Snyder, L. D., & Spaulding, K. (1982). The accuracy of eyewitness identifications in a field setting. *Journal of Personality and Social Psychology*, *42*, 673–681.
- Brigham, J. C., & Malpass, R. S. (1985). The role of experience and contact in the recognition of faces of own- and other-race persons. *Journal of Social Issues*, *41*, 139–155.
- Byatt, G., & Rhodes, G. (1998). Recognition of own-race and other-race caricatures: Implications for models of face recognition. *Vision Research*, *38*, 2455–2468.
- Carroo, A. W. (1986). Other race recognition: A comparison of Black American and African subjects. *Perceptual and Motor Skills*, *62*, 135–138.
- Carroo, A. W. (1987). Recognition of faces as a function of race, attitudes, and reported cross-racial friendships. *Perceptual and Motor Skills*, *64*, 319–325.
- Chance, J. E., & Goldstein, A. G. (1996). The other-race effect and eyewitness identification. In S. L. Sporer, R. S. Malpass, & G. Koehnken (Eds.), *Psychological issues in eyewitness identification* (pp. 153–176). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Chiroro, P., & Valentine, T. (1996). An investigation of the contact hypothesis of the own-race bias in face recognition. *Quarterly Journal of Experimental Psychology*, *48A*, 879–894.
- Cook, S. W. (1970). Motives in a conceptual analysis of attitude-related behavior. In W. J. Arnold & D. Levine (Eds.), *Nebraska symposium on motivation, 1969* (pp. 199–231). Lincoln: University of Nebraska Press.
- Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*, 671–684.
- Cross, J. F., Cross, J., & Daly, J. (1971). Sex, race, age, and beauty as factors in recognition of faces. *Perception and Psychophysics*, *10*, 393–396.
- Cutler, B. L., & Penrod, S. D. (1995). *Mistaken identification: The eyewitness, psychology, and the law*. Cambridge, England: Cambridge University Press.
- Devine, P. G., & Malpass, R. S. (1985). Orienting strategies in differential face recognition. *Personality and Social Psychology Bulletin*, *11*, 33–40.
- Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, *115*, 107–117.
- Donaldson, W. (1992). Measuring recognition memory. *Journal of Experimental Psychology*, *121*, 275–277.
- Dricker, J., Butters, N., Berman, G., Samuels, I., & Carey, S. (1978). The recognition and encoding of faces in alcoholic Korsakoff and right hemispheric patients. *Neuropsychologia*, *16*, 683–695.
- Dunning, D., Li, J., & Malpass, R. (1998, March). *Basketball fandom and cross-race identification among European-Americans: Another look at the contract hypothesis*. Paper presented at the American Psychology-Law Society Biennial Conference, Redondo Beach, CA.
- Feingold, G. A. (1914). The influence of environment on identification of persons and things. *Journal of Criminal Law and Police Science*, *5*, 39–51.
- Glanzer, M., & Adams, J. K. (1985). The mirror effect in recognition memory. *Memory & Cognition*, *13*, 8–20.
- Glanzer, M., & Adams, J. K. (1990). The mirror effect in recognition memory: Data and theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 5–16.
- Goldstein, A. G., & Chance, J. E. (1971). Visual recognition of complex configurations. *Perception and Psychophysics*, *9*, 237–241.
- Goldstein, A. G., & Chance, J. E. (1981). Laboratory studies of face recognition. In G. Davies, H. Ellis, & J. Shepherd (Eds.), *Perceiving and remembering faces* (pp. 81–104). London: Academic.

- Grier, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychological Bulletin*, 75, 424–429.
- Hosch, H. M. (1994). Individual differences in personality and eyewitness identification. In D. F. Ross, J. D. Read, & M. P. Toglia (Eds.), *Adult eyewitness testimony: Current trends and developments* (pp. 328–347). New York: Cambridge University Press.
- Hosch, H. M., Bothwell, R. K., Sporer, S. L., & Saucedo, C. (1990). *The accuracy of a witness's choice: Facial recognition ability, confidence, decision-time and person descriptions*. Unpublished manuscript, University of Texas at El Paso.
- Hosie, J. A., & Milne, A. B. (1996). The effect of experimental design on memory for typical and distinctive faces. *Memory*, 4, 175–197.
- Howells, T. H. (1938). A study of ability to recognize faces. *Journal of Abnormal and Social Psychology*, 33, 124–127.
- Huff, R., Rattner, A., & Sagarin, E. (1996). *Convicted but innocent: Wrongful conviction and public policy*. Thousand Oaks, CA: Sage.
- Jalbert, N. L., & Getting, J. (1992). Racial and gender issues in facial recognition. In F. Losel, D. Bender, & T. Bliesener (Eds.), *Psychology and law: International perspectives* (pp. 309–316). Berlin: deGruyter.
- Kassin, S. M., Ellsworth, P. C., & Smith, V. L. (1989). The “general acceptance” of psychological research on eyewitness testimony: A survey of the experts. *American Psychologist*, 44, 1089–1098.
- Lavrakas, P. J., Buri, J. R., & Mayzner, M. S. (1976). A perspective on the recognition of other race. *Perception and Psychophysics*, 20, 475–481.
- Leippe, M. R. (1995). The case for expert testimony about eyewitness memory. *Psychology, Public Policy, and Law*, 1, 909–959.
- Lipton, J. P. (1977). On the psychology of eyewitness testimony. *Journal of Applied Psychology*, 66, 79–89.
- Loftus, E. F. (1979). *Eyewitness testimony*. Cambridge, MA: Harvard University Press.
- Malpass, R. S. (1990). An excursion into utilitarian analyses, with side trips. *Behavioral Science Research*, 24, 1–15.
- Malpass, R. S., & Kravitz, J. (1969). Recognition for faces of own- and other-race faces. *Journal of Personality and Social Psychology*, 13, 330–334.
- Mason, S. E. (1986). Age and gender as factors in facial recognition and identification. *Experimental Aging Research*, 12, 151–154.
- Munsterberg, H. (1908). *On the witness stand: Essays on psychology and crime*. New York: Doubleday.
- Ng, W., & Lindsay, R. C. L. (1994). Cross-race facial recognition: Failure of the contact hypothesis. *Journal of Cross-Cultural Psychology*, 25, 217–232.
- Osterrieth, P. A. (1944). Le test du copie d'une figure complexe [Test of copying a complex figure]. *Archives of Psychology*, 28, 206–356.
- O'Toole, A. J., Deffenbacher, K. A., Valentin, D., & Abdi, H. (1994). Structural aspects of face recognition and the other-race effect. *Memory and Cognition*, 22, 208–224.
- Platz, S. J., & Hosch, H. M. (1988). Cross-racial/ethnic eyewitness identification: A field study. *Journal of Applied Social Psychology*, 18, 972–984.
- Rey, A. (1942). L'examen psychologique dans tes d'encephalopathie traumatique [The psychological examination in cases of traumatic encephalopathy]. *Archives of Psychology*, 28, 286–340.
- Rhodes, G., Brake, S., Taylor, K., & Tan, S. (1989). Expertise and configural coding in face recognition. *British Journal of Psychology*, 80, 313–331.
- Ross, D. F., Read, J. D., & Toglia, M. P. (Eds.). (1994). *Adult eyewitness testimony: Current trends and developments*. New York: Cambridge University Press.
- Shapiro, P. N., & Penrod, S. (1986). Meta-analysis of facial identification studies. *Psychological Bulletin*, 100, 139–156.
- Shepherd, J. W. (1981). Social factors in face recognition. In G. Davies, H. Ellis, & J. Shepherd (Eds.), *Perceiving and remembering faces* (pp. 55–79). London: Academic.
- Shepherd, J. W., Derogowski, J. B., & Ellis, H. D. (1974). A cross cultural study of recognition memory for faces. *International Journal of Psychology*, 9, 205–211.
- Sporer, S. L. (1991). Deep-deeper-deepest? Encoding strategies and the recognition of human faces. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 17, 323–333.
- Sporer, S. L., Malpass, R. S., & Koehnken, G. (1996). *Psychological issues in eyewitness identification*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, 87, 245–251.
- Swope, T. (1994). *Social experience, illusory correlation, and facial recognition ability*. Unpublished master's thesis, Florida State University.
- Taylor, E. M. (1959). *Psychological appraisal of children with cerebral defects*. Cambridge, MA: Harvard University Press.
- Thibaut, J. W., & Kelley, H. H. (1959). *The social psychology of groups*. New York: Wiley.
- Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion and race on face recognition. *Quarterly Journal of Experimental Psychology*, 43A, 161–204.
- Valentine, T., & Endo, M. (1992). Towards an exemplar model of face processing: The effects of race and distinctiveness. *Quarterly Journal of Experimental Psychology*, 44A, 671–703.
- Vokey, J. R., & Read, J. D. (1992). Familiarity, memorability, and the effect of typicality on the recognition of faces. *Memory & Cognition*, 20, 291–302.
- Warrington, E. K., & James, M. (1967). An experimental investigation of facial recognition with unilateral cerebral lesions. *Cortex*, 3, 317–326.
- Wells, G. L., & Loftus, E. F. (Eds.). (1984). *Eyewitness testimony: Psychological perspectives*. Cambridge, England: Cambridge University Press.
- Whipple, G. M. (1909). The observer as reporter: A survey of the psychology of testimony. *Psychological Bulletin*, 6, 153–170.
- Whipple, G. M. (1912). Psychology of testimony and report. *Psychological Bulletin*, 12, 221–224.
- Whipple, G. M. (1918). The obtaining of information: Psychology of observation and report. *Psychological Bulletin*, 15, 217–248.
- Yarmey, A. D. (1979). *The psychology of eyewitness testimony*. New York: Free Press.
- Yarmey, A. D., & Jones, H. P. T. (1983). Is the psychology of eyewitness identification a matter of common sense? In S. Lloyd-Bostock & B. R. Clifford (Eds.), *Evaluating witness evidence: Recent psychological research and new perspectives* (pp. 13–40). Chichester, England: Wiley.

Copyright of Basic & Applied Social Psychology is the property of Lawrence Erlbaum Associates and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.