

They Don't All Look Alike: Individuated Impressions of Other Racial Groups

Leslie A. Zebrowitz, Joann M. Montepare, and Hoon Koo Lee

Reliability, content, and homogeneity of own- and other-race impressions were assessed: U.S. White, U.S. Black, and Korean students rated faces of White, Black, or Korean men. High intraracial reliabilities revealed that people of 1 race showed equally high agreement regarding the traits of own- and other-race faces. Racially universal appearance stereotypes—the attractiveness halo effect and the babyface overgeneralization effect—contributed substantially to interracial agreement, which was only marginally lower than intraracial agreement. Moreover, similar attention to variations in appearance yielded similar degrees of own- and other-race trait differentiation. When own- and other-race differences in the differentiation of faces on babyfacedness were statistically controlled, differences in trait differentiation were eliminated. Despite the individuated impressions of other-race faces, certain racial stereotypes persisted.

In the fall of 1990, seven white police officers with guns drawn ordered a Black U.S. basketball player, Dee Brown, from his parked car in an affluent and predominately White Boston suburb. Brown had been mistaken for another Black man who had recently committed a bank robbery. This incident seems to support the adage “they all look alike,” and consistent with the unfortunate experience of Dee Brown, considerable research has shown that people are indeed better at recognizing faces of people from their own racial group than those from a different racial group (see Anthony, Copper, & Mullen, 1992; Brigham & Malpass, 1985; Shapiro & Penrod, 1986; and Shepherd, 1981, for reviews of this literature). Moreover, there is evidence to indicate that this “other-race” effect is quite pervasive, extending to judgments other than recognition. Specifically, there is also support for the homily “they all are alike”: People perceive out-group members, including other racial groups, as more homogeneous in their traits and behavior than in-group members (see Linville, Salovey, & Fischer, 1986; and Quattrone, 1986, for reviews of this literature).

Leslie A. Zebrowitz, Department of Psychology, Brandeis University; Joann M. Montepare, Department of Psychology, Tufts University; Hoon Koo Lee, Department of Psychology, Yonsei University, Seoul, Korea.

Portions of this research were presented at the Sixth International Conference on Event Perception and Action, Amsterdam, August 1991, and the European Association of Experimental Social Psychology, Small Meeting on Changing Stereotypes, Paris, September 1991.

This research was supported by National Institute of Mental Health Grant BSR 5 ROI MH42684 to Leslie A. Zebrowitz.

Thanks are extended to Sheila Brownlow, Natalie Greenberg, Karen Hoffman, and Lynn Steiner for their help in the data collection and Ranjana Dutta for her help with the analyses. We also thank Roy Malpass for sharing his photographs of Black people's faces with us and three anonymous reviewers for their helpful comments on a draft of this article.

Correspondence concerning this article should be addressed to Leslie A. Zebrowitz, Department of Psychology, Brandeis University, Waltham, Massachusetts 02254.

Despite the considerable evidence that we fail to distinguish among people of a race different from our own, there is reason to believe that we can make such distinctions. The present study addressed the question of how well people distinguish the traits of another racial group when their attention is focused on facial information. Before considering this question, a note is in order concerning the concept of race.

Anthropologists question the validity of race as a scientific concept, noting the lack of a precise definition. However, race is nevertheless a widely accepted concept in folk psychology, and social conceptions of race can have a significant impact on social perceptions. (See Zuckerman, 1990, for a discussion of this issue.) Thus, although we acknowledge that clear decisions on category membership are often difficult, in this article we nevertheless use the “fuzzy” category system of racial groups and the terms *White*, *Black*, and *Korean* to denote the physical appearance of target faces and the group identification of perceivers.

In contrast to the foregoing evidence for an other-race homogeneity effect, research suggests that when perceivers are explicitly requested to compare faces on physical dimensions, they make similar distinctions among the faces of targets of their own and another race. Specifically, Whites can discriminate among Japanese people's faces as well as they can among White people's faces (Goldstein & Chance, 1976, 1978). Also, although the features used to judge the similarity of Black people's faces are different from those used with White people's faces, these differences hold true for both Black and White perceivers (Shepherd & Deregowski, 1981). Thus, people of different races make similar distinctions among faces even though the facial features they use to distinguish among faces of their own race may differ from those they use for another. People of different races also perceive similar variability in appearance among faces of their own and another race, and they show crossrace agreement in appearance judgments (Bernstein, Lin, & McClellan, 1982; Cunningham, 1986; Maret, 1983; Maret & Harling, 1985; McArthur & Berry, 1987; Wagatsuma & Kleinke, 1979).

That people can detect similarities and differences in the appearance of members of another race when making on-line judgments has implications for homogeneity in crossrace trait ascriptions. According to the ecological theory of social perception (McArthur & Baron, 1983), social stimulus information, such as that provided in facial appearance, can communicate intentions, emotional states, and traits. Thus, perceivers who attend to variations in the appearance of targets of another race should show individuated judgments about the psychological attributes of those targets. Consistent with this reasoning, there is considerable evidence that perceived variations in physical appearance do produce variations in perceived traits (see Berry & McArthur, 1986; Bull & Rumsey, 1988; and McArthur, 1982, for reviews of pertinent literature). Moreover, the studies that have consistently demonstrated an out-group homogeneity effect in trait judgments have not provided appearance information. Rather, they have used methods that foster category-based responding, that is, responding to members of the out-group in terms of their social category membership rather than in terms of their unique individual characteristics. More specifically, subjects in these studies have been asked to imagine a particular group of people and to rate their attributes (Jones, Wood, & Quattrone, 1981; Linville, 1982; Linville, Fischer, & Salovey, 1989; Linville & Jones, 1980; Linville, Salovey, & Fischer, 1986; Park & Judd, 1990; Park & Rothbart, 1982; Quattrone & Jones, 1980). Whereas the other-race impressions elicited by these procedures are relatively homogeneous, reflecting schematic processing in which the out-group is represented either by a prototype or exemplars, impressions elicited by actual faces may be more individuated. Indeed, Smith and Zarate (1992) recently argued that "There is a real need for the development of assessment techniques that can reveal what exemplars or other knowledge structures are accessed when a real person is the cue" (p. 17).

Consistent with the foregoing argument, some studies have established that people can make reliable trait judgments about people of another race when asked to rate specific faces (Keating, Mazur, & Segall, 1981; McArthur & Berry, 1987). However, researchers have not systematically compared the reliability of own-race versus other-race judgments nor have they assessed homogeneity in other-race trait judgments using such a methodology. Reason to believe that this method would eliminate the other-race homogeneity effect is provided by research using the minimal-group paradigm (Tajfel, Billig, Bundy, & Flament, 1971), which found that although the out-group as a whole is perceived as less variable than the in-group, ratings of specific out-group members were not less variable than ratings of specific in-group members (Judd & Park, 1988).

The present study was designed to investigate the extent to which White, Black, and Korean perceivers make physical and psychological distinctions when judging individual members of their own versus the other two races and the similarity of the distinctions they make. Six questions were addressed. First, is there a shared standard by which people of one race make judgments about faces? To answer this question, intraracial reliabilities of appearance and trait judgments were compared for own-race and other-race targets. In keeping with the ecological theory assumption that facial information communicates cer-

tain traits, it was predicted that intraracial reliabilities would be uniformly high, indicating that people of one race all perceive the same trait information in faces and that there is as much agreement regarding the traits conveyed by faces of another race as one's own.

The second question concerned whether the standards of judgment differ across perceiver race, that is, do perceivers of all races make the same trait judgments on the basis of facial information? White raters might show as much agreement about the honesty of various Black people's faces as Black raters do, whereas at the same time White and Black raters might not agree as to which Blacks look the most honest. To address this question, the extent of between-race versus within-race agreement was compared. To the extent that individuals of all races are attuned to the same facial information, it was predicted that both within- and between-race agreement would be highly significant; that is, that faces will be ordered on appearance and trait dimensions in roughly the same way by perceivers of their own race and perceivers of another race. It was further predicted that within- and between-race agreement would be equally high for judgments of babyfacedness, as maturational changes in craniofacial appearance are universal (Lorenz, 1943; Todd, Mark, Shaw, & Pittenger, 1980). On the other hand, greater within- than between-race agreement was expected for judgments of attractiveness, which may be subject to cultural influences (e.g., Hatfield & Sprecher, 1986).

A third question concerned the degree of differentiation among own-race and other-race targets. Agreement between perceivers of different races could be obtained even if those who were of a different race from the targets used only two extreme points on a rating scale (low differentiation), whereas those from the targets' own race used the whole range of scale values (high differentiation). To investigate this possibility, we used a measure proposed by Linville, Salovey, and Fischer (1986). In accordance with the ecological theory tenet that facial information communicates certain traits, similar degrees of own-race and other-race differentiation were predicted for trait ratings to the extent that there was similar differentiation in facial appearance ratings. Following the ecological theory assumption that perceptual experience can influence the perceiver's ability to detect various qualities, we expected differentiation to be higher for faces of a familiar race than a less familiar one.

A fourth question concerned the content of racial stereotypes. Like the research investigating out-group homogeneity effects, past research investigating specific racial stereotypes has engaged category-based judgments, asking raters to judge the qualities of an abstract group of individuals. This research has found in-group favoritism effects on evaluatively loaded dimensions as well as specific racial stereotypes. Most pertinent to the present study are the results of an investigation that examined stereotypes of various racial and ethnic groups held by perceivers from various groups. A wide range of perceiver groups agreed in rating an abstract group of Blacks as strong and naive, Asians as shrewd, and Whites as dominant (Vinacke, 1949). Although more recent stereotyping research has not systematically varied perceiver and target race, a number of studies investigating Whites' stereotypes of Blacks have corroborated

rated Vinacke's earlier findings. Blacks are still more apt to be viewed as ignorant and naive and less apt to be viewed as smart. Also, Whites respond more negatively to a Black supervisor than to a Black subordinate and they are less apt to see Blacks as ambitious, which suggests that Blacks are also still seen as less dominant than Whites (Dovidio & Gaertner, 1981; Gaertner & Dovidio, 1986; Gaertner & McLaughlin, 1983; Karlins, Coffman, & Walters, 1969).

In the present investigation, we sought to determine whether the in-group favoritism effect and specific racial stereotypes would be manifested when raters judged pictures of individual targets. Reason to believe that this method would eliminate racial stereotypes was provided by Quattrone and Jones (1980), who found that out-group stereotypes did not affect ratings of a specific out-group member who had been viewed on videotape, whereas they did affect ratings of the group as a whole. Similarly, research has shown that when subjects are given specific information concerning the behaviors of targets who vary in age or sex, their ratings of these targets do not reflect the age and sex stereotypes that are manifested when targets' age or sex category is the only information provided (Braithwaite, 1986; Locksley, Borgida, Brekke, & Hepburn, 1980; Locksley, Hepburn, & Ortiz, 1982; Pratto & Bargh, 1991). Whereas the foregoing studies indicate that group stereotypes do not affect judgments about individuated group members, other research has found that the perceived variability of group members, that is, the extent to which they are individuated, can be independent of the extent to which they are perceived favorably or as fitting a group stereotype (Linville, Fischer, & Salovey, 1989; Park & Judd, 1990). Thus, no firm predictions could be made concerning whether racial stereotypes would co-occur with high individuation of other race targets on the other measures.

The fifth question addressed concerned the crossrace generality of facial appearance stereotypes. According to the ecological theory, such stereotypes may result from the overgeneralization of perceptions that are usually adaptive. Thus, the perception of particular traits in people with particular facial configurations may reflect the resemblance of these faces to others that do accurately communicate the traits. In particular, Zebrowitz and her colleagues have proposed that because it is so important to perceive the dependency and need for nurturance that is communicated by a baby's facial appearance, these qualities may be overdetected in babyfaced adults, who physically resemble the young (Berry & McArthur, 1985, 1986; McArthur & Apatow, 1983-84; McArthur & Berry, 1987; Zebrowitz-McArthur & Montepare, 1989; Zebrowitz & Montepare, 1992). Such impressions of babyfaced adults should be racially universal inasmuch as both the appearance and the psychological attributes of babies are universal. Thus, it was predicted that there would be crossrace generality for the babyfacedness overgeneralization effect, whereby adults with more neotenus facial structures are perceived to have more childlike traits.

Another well-established appearance stereotype is the attractiveness halo effect, whereby those who are more physically attractive are perceived to have more favorable traits. This effect is most pronounced in the domain of social competence, but can also be observed in other domains (See Adams, 1977; Ber-

scheid & Walster, 1974; Eagly, Ashmore, Makhijani, & Longo, 1991; Feingold, 1992, for reviews). Unlike the babyfacedness overgeneralization effect, the origins of the attractiveness halo effect are not well understood, and it is therefore difficult to predict whether crossrace generality will occur. If, as some have proposed, the attribution of positive qualities to attractive faces is a product of cultural influences, then individuals from a culture very different from our own may not show this effect. Alternatively, if the attractiveness halo derives from universally positive affective responses to aesthetically pleasing faces, then the effect should be universal.

The crossrace commonalities in attunement to facial qualities that give rise to the babyfacedness overgeneralization effect and the attractiveness halo effect could contribute to the predicted interracial agreement in trait ratings. To test the mediational role of these appearance stereotypes, interracial agreement was assessed with judgments of facial attractiveness or babyfacedness statistically controlled. It was predicted that the agreement reflected in zero-order correlations would be diminished in these partial correlation analyses.

The final question addressed in the present investigation concerned crossrace generality in the facial features specifying babyfacedness and attractiveness. Given the universality of maturational changes in craniofacial appearance, it was predicted that perceivers of all races would show high babyfacedness ratings for faces of all races with features that have yielded high babyfacedness ratings for White people's faces in past research: a round face, a small chin, a small nosebridge, large eyes, and thin, high eyebrows (Berry & McArthur, 1985; Keating, 1985; McArthur & Apatow, 1983-84; Zebrowitz-McArthur & Montepare, 1989; Zebrowitz & Montepare, 1992). Because there is no clear consensus regarding the facial components of attractiveness (e.g., Cunningham, 1986, 1990; Keating, 1985; Langlois & Roggman, 1990; Zebrowitz & Montepare, 1992), no predictions were made concerning the specific features that would yield high attractiveness ratings. However, past evidence of crossrace agreement in attractiveness judgments suggested that whatever features predicted attractiveness would do so regardless of the perceiver's race (Bernstein, Lin, & McClellan, 1982; Cunningham, 1986; Maret, 1983; Maret & Harling, 1985; McArthur & Berry, 1987; Wagatsuma & Kleinke, 1979).

It should be reiterated that, in addressing each of the foregoing questions, the purpose of this study was not to determine what people most typically do when making crossrace trait judgments but rather to determine whether attention to facial information enables them to make intraracially reliable, interracially consensual, differentiated, and theoretically predictable judgments and whether such attention eliminates racial stereotypes.

Method

Subjects

For their participation in the study, 124 U.S. White college undergraduates received partial course credit, 72 U.S. Black college undergraduates received \$10, and 102 Korean college undergraduates received 1000 won (approximately \$1.25). Approximately equal numbers of men and women from each subject group were randomly assigned to

rate male faces representing one of three racial groups. Subjects participated in nine conditions: Whites rating Whites ($n = 80$), Whites rating Blacks ($n = 24$), Whites rating Koreans ($n = 20$), Blacks rating Whites ($n = 24$), Blacks rating Blacks ($n = 24$), Blacks rating Koreans ($n = 24$), Koreans rating Whites ($n = 48$), Koreans rating Blacks ($n = 30$), and Koreans rating Koreans ($n = 30$).¹ All subjects rated faces of only one race, with one exception: The Black subjects who rated Black targets also rated the Korean targets. The ratings were counterbalanced and made at a 1-week interval. For each group of targets, subjects viewed one of two orders of faces and completed ratings in one of two orders.

Dependent Measures

Impression ratings. Faces were rated on 7-point bipolar trait scales reflecting dimensions that had been influenced by variations in facial attractiveness and babyfacedness in past research with young adult faces and that could be used to examine own-race favoritism and racial stereotypes (e.g., Berry & McArthur, 1986; Eagly et al., 1991; Feingold, 1992; Vinacke, 1949). The trait scales were *cold-warm*, *physically weak-physically strong*, *dishonest-honest*, *submissive-dominant*, and *naive-shrewd*. In addition to trait ratings, subjects rated how *babyfaced-maturefaced*, and how *unattractive-attractive* each face was on 7-point bipolar scales. Ratings were coded so that higher scores represented greater submissiveness, physical weakness, naïveté, warmth, honesty, babyfacedness, or attractiveness. Subjects also estimated the age, in years, of each face. All rating scales for Korean perceivers were initially translated into Korean by a Korean visiting scholar at Brandeis University and then translated back into English by a Korean professor at Seoul National University, with a resultant change in the original Korean translation of one word: *naive*.

Whereas all five traits tapped the babyfacedness stereotype, only warmth, honesty, and dominance tapped the attractiveness halo. Warmth was considered to be a component of Feingold's (1992) sociability dimension, which included traits akin to warmth, such as sociability and friendliness, and which showed a strong attractiveness halo in western samples, as did judgments of sexual warmth. Warmth can also be viewed as a component of the social competence dimension that Eagly et al. (1991) found to yield an attractiveness halo, and our own research has consistently shown an attractiveness halo on warmth judgments (e.g., McArthur & Berry, 1987; Zebrowitz & Montepare, 1992). This work also has shown a reliable attractiveness halo on honesty judgments, although judgments of character and integrity were only weak correlates of attractiveness in the Feingold (1992) and Eagly et al. (1991) meta-analyses. Finally, both meta-analyses revealed a moderate effect of attractiveness on potency and dominance judgments (Eagly et al., 1991; Feingold, 1992), although our own research has yielded no such effects (e.g., Zebrowitz & Montepare, 1992).

Facial characteristics. Five age-related facial characteristics were determined from facial measurements made independently by two judges who projected each slide onto a wall and measured the facial features depicted in Figure 1 with a ruler. The specific facial characteristics included cranium height, eyebrow height, eye area, nose area, and chin area. These features were selected because all had been shown in previous research to predict ratings of babyfacedness or attractiveness.

In addition to the foregoing linear facial measures, three qualitative ratings were made of other facial features that had been shown to be associated with babyfacedness and attractiveness. Specifically, two independent judges rated the following facial characteristics on 7-point bipolar scales: *thick eyebrows-thin eyebrows*, *small nosebridge-large nosebridge*, and *round face-angular face*. Nosebridges were coded as *small* to the extent that the area between the eyes was flat rather than protruding. Faces were coded as *round* to the extent that face length

and width appeared equal, and the jaw was rounded. They were coded as *angular* to the extent that face length appeared greater than width, and the jaw was angular. A *neutral* rating was given to faces that appeared equal in length and width and had an angular jaw or to faces that had greater length than width and a rounded jaw. Facial roundness was distinct from measured chin area, with which it was uncorrelated. On the other hand, nosebridge size was significantly correlated with nose area for some racial groups. To avoid redundant predictors, nose area was replaced with one of its components, measured nose width, which was uncorrelated with nosebridge ratings.

Pearson product-moment correlation coefficients between the two judges' measurements of facial features revealed an average reliability across racial groups equal to .95 for cranium height, .84 for eye area, .87 for eyebrow height, .81 for eyebrow thickness, .81 for chin area, .81 for facial roundness, .85 for nosebridge size, and .87 for nose width. The reliable facial measurements and ratings made by the two judges were averaged to yield a single score for each face for use in subsequent analyses.

Facial Stimuli

The facial stimuli consisted of black-and-white slides depicting the faces of White ($n = 20$), Black ($n = 24$), and Korean ($n = 24$) young adult men, all of whom had a serious facial expression, no facial hair, and no glasses or other distinguishing marks. Two slide orders were generated for each group of faces. One order of slides was a random sequence, and the other was the reverse of the first. The slides of the White and Korean men were made from yearbook photos of graduating seniors at a U.S. and a Korean college. The slides of the Black men were created from photographs of men at four colleges that had been used in research that has yielded an other-race homogeneity effect in facial recognition (Malpass & Kravitz, 1969). The 52 Black people's faces used by Malpass and Kravitz were obtained by asking students on Black college campuses who walked past a researcher whether they would consent to having their photograph taken. Of these faces, only 32 met the "no facial hair" selection criterion in the present study, and 24 of them were used.

Comparability of faces across race. Slides within each racial group were selected with the goal of achieving a range of attractiveness and babyfacedness that was comparable for faces of all races. Achieving such comparability was important to address questions concerning the crossrace universality of facial stereotypes. To assess the level of comparability, babyfacedness and attractiveness ratings were averaged across subjects to create mean ratings of faces of each race by each group of raters. (As will be seen below in the section on Intra-racial Reliability of Judgments, reliability of ratings across judges supported an analysis of their mean ratings.) These mean appearance ratings, which are shown in Table 1 along with the standard deviations, were subjected to 3 (target race) \times 3 (perceiver race) analyses of variance (ANOVAs). Neither the babyfacedness nor the attractiveness of the targets varied with target race (both $F_s < 1$), and simple effects analyses similarly revealed no significant effects of target race within any of the three perceiver groups (all $p_s > .10$). Tests for heterogeneity of variance also revealed no significant differences across the three target races, $F_{\max}(3, 67) = 1.68$ and 1.42, for babyfacedness and attractiveness, respectively, both $p_s > .05$.

¹ The varying number of subjects in each group reflects the historical order in which each set of data was collected. The data for White people's faces, collected in earlier studies, included more subjects than the data for Black people's and Korean faces, which were collected in the present investigation.

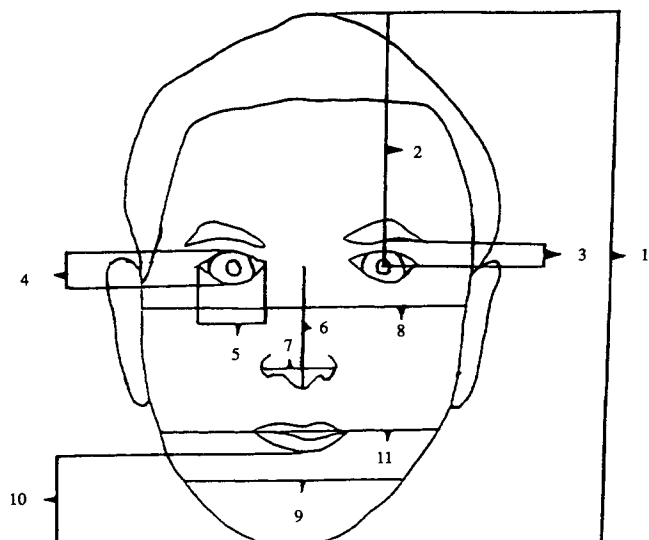


Figure 1. Facial measurements—The methods for calculating physiognomic measures were as follows: *Eye area* was the product of (4) eye height (distance between top and bottom of eye at center) and (5) eye width (distance between inner and outer corners of the eye); *eyebrow height* was (3) the distance from the pupil center to the eyebrow; *chin area* was the product of (9) the width at the point halfway between the bottom lip and the base of the chin and (10) the distance from the lower lip center to the base of the chin; *cranium height* was (2) the distance from the middle of the pupil to the top of the head minus $\frac{1}{2}$ of hair length; *nose area* was the product of (6) nose length (distance from the bridge at middle of the eye to the tip of the nose) and (7) nose width (distance between the left and right sides of the nose at the tip). (To correct for variations in face size, all vertical measures were divided by face length [1]; Horizontal Measures 5 and 7 were divided by Face Width 8; and Horizontal Measure 9 was divided by Face Width 11.)

Representativeness of faces. Equating the babyfacedness and attractiveness of faces of each race, as was necessary to investigate facial stereotypes, raises a question regarding the representativeness of the faces that could limit the generalizability of the differentiation results. In particular, if equal own-race and other-race trait differentiation is found when facial information is used, this could conceivably reflect the selection of faces that were more differentiated in appearance than a random set of faces would be. In evaluating this concern, it should be noted that research using physical measures of large samples of faces has yielded no evidence for racial differences in facial-feature variability (Goldstein, 1979). This strongly suggests that the equal variability in attractiveness and babyfacedness of the White people's, Black people's, and Korean's faces used in this study would also be obtained with other samples of faces.

To provide more direct evidence for the representativeness of the faces used in the present study, 24 faces were randomly selected from the entire set of 52 Black people's faces originally used by Roy Malpass. A group of 17 White perceivers rated these faces on the same scales used in the main study, and another group of 20 White perceivers rated the set of 24 faces that originally had been selected. The reliability of attractiveness and babyfacedness ratings were equally high for each group of slides, with an average standardized alpha of .95 across all ratings. Ratings of each slide were therefore averaged across subjects,

Table 1
Babyfacedness and Attractiveness of the Three Target Groups

Rating and raters	Target race					
	White		Black		Korean	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Attractiveness						
All	3.55	0.72	3.46	0.76	3.61	0.85
White	3.54	0.76	3.77	0.73	3.69	0.86
Black	3.44	0.48	3.36	0.75	3.53	0.78
Korean	3.67	0.86	3.24	0.71	3.62	0.94
Babyfacedness						
All	3.84	1.08	3.81	1.24	3.62	0.96
White	3.92	0.97	3.75	1.11	3.78	0.97
Black	3.82	0.99	3.78	1.21	3.26	0.68
Korean	3.77	1.31	3.90	1.44	3.82	1.12

and the mean ratings for slides in each group were subjected to one-way ANOVAs. The results revealed no significant differences in mean attractiveness ratings ($M_s = 2.96$ and 2.93 for old and new slides, respectively, $F < 1$) or mean babyfacedness ratings ($M_s = 3.98$ and 3.86 for old and new slides, respectively, $F < 1$). There were also no significant differences in the variability of attractiveness ratings ($SD_s = 1.10$ and 0.83 for old and new slides, respectively), $F_{\max}(2, 23) = 1.75$, $p > .05$, or in the variability of babyfacedness ratings ($SD_s = 1.45$ and 1.31 for old and new slides, respectively), $F_{\max}(2, 23) = 1.22$, $p > .05$. Thus, the babyfacedness and attractiveness of the Black people's faces used in this study are no different from the range and levels found in a sample of Black people's faces that has been selected without regard to particular appearance criteria.

Procedure

Subjects, run in noninteracting groups of 2–8, were informed that the study dealt with how people's physical appearance influences perceivers' impressions of them and that their task was to rate faces of different people on a series of scales. Each face was presented for 8 s with a 2-s interslide interval, during which subjects made their ratings on one scale. After the first rating of all of the faces was completed, subjects viewed the set of faces a second time and completed the second rating. This procedure was repeated until all of the ratings had been made.

Two orders of rating scales were generated for each facial group. Half of the subjects rated the faces on a random sequence of the trait scales, and this rating order was reversed for the remaining subjects. Ratings of attractiveness, babyfacedness, and age were made after all trait ratings had been completed, with half of the subjects rating babyfacedness first and half rating attractiveness first. These ratings were made last so that subjects would not be explicitly aware of variations along these appearance dimensions when making their trait impression ratings.

Results

Intraracial Reliability of Judgments

Cronbach's coefficients computed across ratings made by a racially homogeneous group of subjects revealed high reliability for faces of each race (see Table 2). The only exception was

Table 2
Reliability of Facial and Trait Ratings

Target race and raters	Rating							
	Attractive	Babyface	Age	Submissive	Naive	Weak	Warm	Honest
White								
White	.96	.98	.98	.91	.96	.96	.98	.97
Black	.54	.89	.93	.91	.87	.96	.93	.87
Korean	.94	.98	.98	.95	.95	.97	.97	.94
Black								
White	.89	.95	.96	.88	.78	.94	.88	.86
Black	.89	.96	.97	.90	.87	.93	.93	.87
Korean	.88	.98	.98	.95	.93	.96	.94	.94
Korean								
White	.88	.91	.90	.91	.87	.93	.90	.86
Black	.89	.87	.95	.83	.75	.86	.87	.85
Korean	.94	.96	.96	.89	.83	.89	.94	.90
Own race	.93	.97	.97	.90	.89	.93	.95	.91
Other race	.84	.93	.95	.90	.84	.94	.92	.89

Black subjects' ratings of the attractiveness of White people's faces. Thus, as predicted, White, Black, and Korean perceivers showed high agreement with others of their own race regarding the traits and appearance of White people's, Black people's, and Korean's faces. Moreover, with the exception of Blacks judging the attractiveness of Whites, crossrace judgments were no less reliable than within-race judgments ($M_s = 0.90$ and 0.93 , respectively).

Interracial Agreement in Judgments

Having demonstrated high intraracial agreement, whether judging members of another race or one's own, a second set of analyses investigated whether the standards of judgment differ across perceiver race. The first step in addressing this issue was to compute mean ratings for each face across subjects within each racial group. Means were initially calculated separately for male and female raters so that cross-sex agreement within each perceiver-target racial group could be determined. Correlational analyses revealed strong male-female agreement for all judgments, with the average agreement across the nine perceiver-target racial groups equal to .87 for babyfacedness, .72 for attractiveness, .89 for age, .79 for submissiveness, .75 for naïveté, .90 for strength, .85 for warmth, and .81 for honesty. Given these results, the mean ratings of each face were collapsed across sex of rater in subsequent analyses.

Between-race agreement was calculated by correlating the mean ratings of the faces provided by raters of one race with the mean ratings provided by raters of another race; for example, the mean attractiveness rating of each White person's face by White perceivers was correlated with the mean attractiveness ratings of the same faces by Black perceivers. To provide a comparable measure of within-race agreement, each sample of raters was randomly divided in half, and the mean ratings of own-race faces provided by one half of the group was correlated with the mean ratings provided by the second half.

As predicted, there was highly significant between-race

agreement for all pairs of perceiver groups judging targets of all three races (see Table 3). Moreover, a 2 (between vs. within race) $\times 8$ (rating dimension) ANOVA on the level of agreement revealed that the average between-race agreement was only slightly lower ($M = 0.75$) than the average agreement within races ($M = 0.84$). $F(1, 4) = 4.38, p = .10$. As predicted, between- and within-race agreement differed more for attractiveness ($M_s = 0.64$ and 0.78 , respectively) than for babyfacedness ($M_s = 0.87$ and 0.91 , respectively). However, neither difference was significant because of the small sample size (both $p_s > .05$).²

Differentiation of Judgments

Whereas the strong between-race agreement in judgments reveals that individuals are ordered on appearance and trait dimensions in roughly the same way by perceivers of their own race and perceivers of another race, the question remains as to whether perceivers make finer differentiations among individuals of their own race. To investigate the impact of target and perceiver race on degree of differentiation, we used a measure proposed by Linville et al. (1986). This measure reflects the probability that a perceiver will assign two randomly chosen individuals to different levels of the trait. The more levels on a rating scale that a perceiver uses when rating a series of faces, the more differentiated the perceiver's ratings are. Also, for any given number of levels used, differentiation is higher when each level is perceived to be equally likely. In studies eliciting category-based processing, subjects are asked to distribute 100 hypothetical people from a given category across seven levels of a trait. In the current study, the distribution of ratings of faces of a

² The finding that within-race agreement was only marginally higher than crossrace agreement did not reflect lower reliability in the within-race scores because the sample of raters was divided in half. When the number of raters was equated across all comparisons, within- and between-race agreement remained comparable ($M_s = 0.84$ and 0.71 , respectively).

Table 3
Interracial Agreement in Facial and Trait Ratings

Target race and agreement category	Rating							
	Attractive	Babyface	Age	Submissive	Naive	Weak	Warm	Honest
White								
White-White	.72	.91	.96	.83	.88	.95	.91	.90
White-Black	.69	.93	.90	.90	.86	.87	.90	.89
White-Korean	.47	.94	.87	.73	.61	.80	.86	.82
Black-Korean	.23 ^a	.91	.77	.79	.70	.82	.82	.70
Black								
Black-Black	.77	.91	.94	.82	.70	.80	.81	.65
White-Black	.74	.94	.91	.90	.73	.93	.70	.88
White-Korean	.69	.88	.90	.70	.48	.85	.79	.80
Black-Korean	.75	.92	.93	.65	.72	.87	.62	.76
Korean								
Korean-Korean	.85	.91	.96	.78	.73	.85	.90	.81
White-Black	.79	.86	.90	.79	.84	.85	.78	.83
White-Korean	.72	.71	.81	.33 ^a	.16 ^a	.65	.75	.42
Black-Korean	.69	.73	.78	.36 ^a	.38 ^a	.63	.68	.64
All								
Within race	.78	.91	.95	.81	.77	.87	.87	.79
Between race	.64	.87	.86	.68	.61	.81	.77	.75

Note. For White people's faces, $n = 20$; for Black people's and Korean's faces, $n = 24$. Correlations are reported for the combined ratings by male and female subjects.

^a ns. All other correlations are significant at $p < .05$.

given race served as the data for calculating the probability of differentiation (P_D) for each rater. The formula used in both cases is the same:

$$P_D = 1 - \sum_{i=1}^n P_i^2,$$

where i = level on the rating scale (1-7) and P_i = the proportion of ratings at the i th level.

Following the procedure of Linville et al. (1989), we averaged the differentiation scores for each subject across the five trait ratings to form a trait differentiation index, and we then standardized this index within each perceiver race. (As in the Linville et al., 1989, study, a factor analysis supported construction of the index.) This permitted a comparison of White, Black, and Korean perceivers' trait differentiations for own race and other races while controlling for any perceiver race differences in differentiation. Differentiation scores for ratings of attractiveness and babyfacedness were also standardized within each perceiver race.

The trait differentiation index and appearance differentiation scores were subjected to 3 (perceiver race) \times 3 (target group) between-groups ANOVAs for unweighted means. The three levels of perceiver race were White, Black, and Korean, and the three target groups were own race, the most familiar other race (Blacks for Whites, Whites for Blacks and Koreans), and the less familiar other race (Koreans for Blacks and Whites, Blacks for Koreans). A survey assessing the familiarity of Black ($n = 6$) and White ($n = 37$) Brandeis students with members of each of the three racial groups corroborated the a priori designation of Koreans as less familiar than Whites to Black students. Although Blacks reported no more Whites than Asians in their

neighborhoods when growing up, in their elementary schools, or in their high schools, they did report seeing more Whites than Asians on a typical day, interacting with more Whites, and knowing more Whites well (all $ps < .05$). They also reported more overall familiarity with Whites than Asians ($p < .01$), and when asked to rank order the three racial groups according to how much contact they had with each, all Black subjects ranked Whites above Asians.

The survey responses provided less support for the a priori designation of Koreans as less familiar than Blacks to White students. No significant differences in familiarity with Blacks and Asians were obtained on the specific questions noted above or on a composite of all questions, all $ps > .05$. Finally, when forced to rank order the three racial groups according to contact with each, half of the White subjects ranked Blacks above Asians and half ranked Asians above Blacks. Because familiarity with Asians and Blacks was similar for White subjects, all of the analyses involving target group familiarity were performed two ways: with Koreans designated as the less familiar group to White subjects and with Blacks so designated. Only the analyses treating Blacks as the more familiar group are reported, because both sets of analyses yielded the same effects.

As shown in Table 4, differentiation of attractiveness was no greater for targets of the perceivers' own race than for those of another race. However, perceivers did show marginally more differentiation in babyfacedness ratings for targets from their own race than those from a less familiar race, $t(284) = 1.67$, $p = .10$, but not than those from a familiar race, $t(284) = 1.32$, $p > .10$. This low differentiation in the babyfacedness of the least familiar targets was paralleled by low differentiation on the trait index. Planned comparisons revealed that perceivers' trait

Table 4
Own-Race Versus Other-Race Differentiation

Rating	Target group			<i>F</i> (2, 284)
	Own race	Familiar race	Less familiar race	
<i>P_D</i>				
Attractive	.026	.011	-.120	<1
Babyface	.083 _a	.260 _a	-.154 _b	3.32**
Trait index	.073 _a	-.002	-.153 _b	2.33*
Adjusted index	.066	-.064	-.084	1.34
<i>SD</i>				
Attractive	-.012	-.037	-.136	<1
Babyface	-.015	.209	-.112	2.12
Trait index	-.026	-.029	-.115	<1

Note. The familiar race is Blacks for White perceivers and Whites for Black and Korean perceivers. P_D = probability of differentiation. Means are based on differentiation scores and standard deviations standardized within each perceiver race. Row means with different subscripts differ at $p < .10$ or $.05$.

* $p < .10$. ** $p < .05$.

ratings differentiated own race targets more finely than targets from the less familiar other race, $t(284) = 2.12$, $p < .05$, albeit not than targets from the familiar other race ($t < 1$).

To test the hypothesis that low differentiation of the babyfacedness of targets from an unfamiliar race mediated the low differentiation of their traits, we performed a 3 (perceiver race) \times 3 (target group) analysis of covariance on the trait differentiation index, with babyfacedness differentiation as the covariate. The babyfacedness covariate was highly significant, $F(1, 283) = 61.25$, $p < .001$, and planned comparisons on the adjusted means revealed no more trait differentiation of own-race targets than of targets from a familiar race or a less familiar race (both $ps > .10$). Thus, when own- and other-race differences in the differentiation of facial babyishness were controlled, own- and other-race differences in trait differentiation were eliminated.

Because Korean targets were overrepresented in the unfamiliar racial group, it is possible that the foregoing effects reflected the particular groups of faces being judged rather than their familiarity to the perceivers. To examine this possibility, we subjected the trait index and appearance ratings to 3 (perceiver race) \times 3 (target race) between-groups unweighted means ANOVAs. There was no significant main effect of target race on the differentiation of babyfacedness ratings ($F < 1$), and a significant Perceiver Race \times Target Race interaction, $F(4, 284) = 4.89$, $p < .001$, revealed that whereas Blacks differentiated the babyfacedness of Koreans less than that of the other two target races ($ps < .01$), White and Korean perceivers did not (all $ps > .05$). Also, although a marginally significant effect for target race on trait differentiation, $F(2, 284) = 2.33$, $p = .10$, reflected a tendency for Korean's faces to be less finely differentiated than Black's or White's faces, neither of these comparisons was significant (both $ps > .05$). Moreover, an inspection of trait differentiation scores within each perceiver group revealed that Koreans differentiated Korean's faces more than White's or Black's faces.

Thus the low differentiation of an unfamiliar race reported above did not simply reflect low differentiation of Korean's faces. Finally, a significant effect of target race for attractiveness, $F(2, 284) = 3.38$, $p < .05$, revealed that, whereas differentiation of attractiveness did not vary with familiarity of the targets, Black targets ($M = -.281$) were differentiated less than either White targets ($M = .099$) or Korean targets ($M = .072$), both $ps < .05$.³

Because there is some evidence that the differentiation measure may yield less robust out-group homogeneity effects than the perceived variability of the group members, as indexed by the standard deviation of a perceiver's ratings (Park & Judd, 1990), the standard deviations of the five trait ratings were averaged for each subject to form a trait variability index, which was then standardized within each perceiver race. The z score standardized trait variability index and babyfacedness and attractiveness variability scores were each subjected to 3 (perceiver race) \times 3 (target group) between-groups unweighted means ANOVAs. As shown in Table 4, the results revealed no evidence of an out-group homogeneity effect.

Own-Race Favoritism

To examine own-race favoritism effects, subjects' mean ratings of all targets on the evaluatively loaded dimensions of attractiveness, honesty, and warmth were subjected to 3 (perceiver race) \times 3 (target group) unweighted means ANOVAs in which the three levels of perceiver race were White, Black, and Korean and the three target groups were own race, most familiar other race, and less familiar other race. As shown in Table 5, perceivers did not rate targets of their own race higher than those of a different race on any of these scales. In fact, perceivers rated targets of their own race significantly lower in honesty than targets from another race. A significant Perceiver Race \times Target Group interaction effect, $F(4, 288) = 5.16$, $p < .001$, revealed that this "other-race favoritism" effect held true for White and Black perceivers but not for Koreans.

The Perceiver Race \times Target Group interaction effects were also significant for ratings of warmth and attractiveness, $F_s(4, 288) = 3.05$ and 2.36 , $ps \leq .05$, respectively. An inspection of the warmth means revealed that Black perceivers showed an other-race favoritism effect, just as they did for honesty ratings; White perceivers showed an "own-race favoritism" effect, particularly in contrast with the familiar Black out-group; and Korean perceivers gave equal warmth ratings to all target groups. The attractiveness means revealed that ratings of White and Black perceivers did not differ significantly across target groups, whereas Korean perceivers rated the less familiar Black targets as less attractive than the more familiar White targets or targets from their own group.

³ To eliminate the possibility that the observed differentiation effects were a by-product of differences in the central tendency of distributions for own-race versus other-race targets, analyses of covariance were conducted using the mean differentiation score as a covariate. The results were unchanged.

Table 5
Own-Race Versus Other-Race Favoritism

Rating	Target group						<i>F</i> (2, 288)
	Own race		Familiar race		Less familiar race		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Attractive	3.50	.57	3.63	.63	3.49	.69	1.23
Warm	4.04	.48	3.98	.52	4.13	.65	1.71
Honest	3.96 _a	.51	4.23 _b	.57	4.20 _b	.46	8.06*

Note. The familiar race is Blacks for White perceivers and Whites for Black and Korean perceivers. Row means with different subscripts differ at $p < .05$.

* $p < .001$.

Racial Stereotypes

To examine specific racial stereotypes, subjects' mean ratings of all targets on the dimensions of strength, dominance, and shrewdness were subjected to 3 (perceiver race) \times 3 (target race) unweighted means ANOVAs (see Table 6). Consistent with past research, a main effect for target race revealed that Blacks were rated as stronger than Whites. However, a qualifying interaction, $F(4, 222) = 4.59, p < .001$, revealed that this overall effect was entirely due to White perceivers (see Figure 2). The *dominant-submissive* scale yielded no significant effect for target race, thus failing to replicate past findings that Blacks are stereotyped as submissive, whereas Whites are stereotyped as dominant. However, a significant Perceiver Race \times Target Race interaction, $F(4, 222) = 9.06, p < .001$, revealed that Blacks were rated as more submissive than the two other target groups by both Blacks and Koreans, whereas White perceivers rated Blacks as more dominant than the other two target groups (see Figure 3). Finally, the *naive-shrewd* scale yielded no significant effect for target race, thus failing to replicate past findings that Asians are stereotyped as shrewd, whereas Blacks are stereotyped as ignorant. However, a significant Perceiver Race \times Target Race interaction, $F(4, 288) = 6.00, p < .001$, revealed that both Black and Korean perceivers tended to judge Blacks as

Table 6
Racial Stereotypes

Rating	Target race						<i>F</i>
	White		Black		Korean		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Physically strong	4.03 _a	.39	4.30 _b	.47	4.18 _a	.40	7.79*
Dominant	4.14	.39	4.04	.49	4.10	.41	1.09
Shrewd	4.09	.42	4.08	.56	4.14	.44	<1

Note. Row means with different subscripts differ at $p < .05$ or $.10$. *dfs* = 2 and 222 for physically strong and dominant; *dfs* = 2 and 288 for shrewd.

* $p < .001$.

more naive than the other two target groups, whereas White perceivers judged Black targets to be shrewder than other targets (see Figure 4).⁴

Crossrace Generality of Appearance Stereotypes

The between-race agreement in appearance and trait ratings that were documented above suggest that the latter may reflect crossrace commonalities in two well-documented appearance stereotypes: the babyfacedness overgeneralization effect, whereby those with more neotenous facial structures are perceived to have more childlike traits, and the attractiveness halo effect, whereby those who are physically attractive are perceived to have more favorable traits. To assess the crossrace generality of appearance stereotypes, Pearson product-moment correlation coefficients were computed between the mean babyfacedness or attractiveness ratings and the mean trait ratings of faces within each perceiver-target group. These coefficients, which appear in Table 7, provide strong evidence for crossrace generality in impressions elicited by a babyface and by an attractive face.

With few exceptions, babyfaced men of all races were perceived by subjects of all races as more submissive, naive, physically weak, warm, and honest than their more mature-looking peers. More attractive men of all races were also perceived by subjects of all races as more honest. Although White perceivers rated more attractive faces as warmer, regardless of their race, Black perceivers showed the attractiveness halo effect in their judgments of warmth only for Black faces, and Korean perceivers showed it only in their judgments of White faces. It should be noted that these effects cannot be attributed to higher variability in attractiveness judgments in these instances. (See Table 1.) Consistent with past research, attractiveness was not significantly related to ratings of naïveté or physical weakness, which are less evaluatively loaded than honesty or warmth judgments, and should thus be less susceptible to an attractiveness halo. Attractiveness also failed to predict ratings of dominance, a finding which is inconsistent with the meta-analytic findings of a moderate effect of attractiveness on potency and dominance judgments (Eagly et al., 1991; Feingold, 1992), albeit consistent with other research on dominance judgments (e.g., Zebrowitz & Montepare, 1992).⁵

The evidence for crossrace generality in appearance stereotypes suggests that between-race agreement in trait impressions may be mediated by judgments of facial babyishness or attractiveness. To investigate this possibility, agreement in trait

⁴ The degrees of freedom are lower for strength and dominance than other ratings because of missing data in the White perceiver-White target cell. Covariance analyses revealed that differences in the differentiation of attractiveness and babyfacedness could not account for the racial stereotypes. Neither could differences in the mean ratings of attractiveness and babyfacedness.

⁵ Regression analyses in which babyfacedness and attractiveness judgments were simultaneously entered as predictors of trait impressions together with perceived age revealed that virtually all of the effects of facial babyishness and facial attractiveness on impressions held true independently of each other and perceived age.

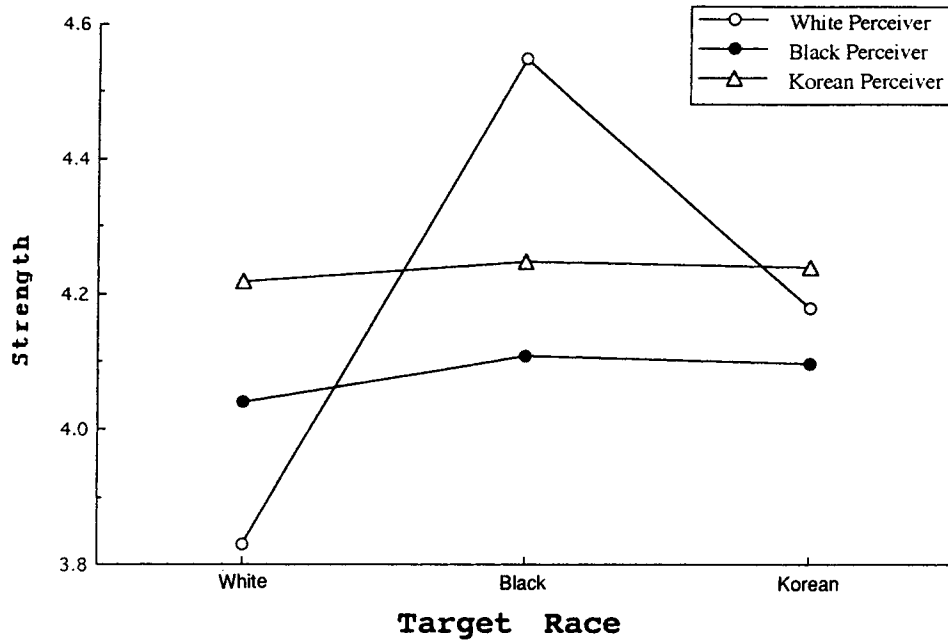


Figure 2. Ratings of physical strength as a function of perceiver and target race.

ratings was assessed with babyfacedness or attractiveness ratings partialled out. Analyses partialing out perceived age were also conducted, because this was another possible mediator of the shared trait impressions. The babyfacedness, attractiveness, and age ratings for each face averaged across raters of all three races were used as the covariates in these analyses.

As shown in Table 8, the results revealed a consistent decrease in interracial agreement when facial babyishness was controlled, and the average agreement across all groups lost significance for ratings of submissiveness and naïveté. Controlling perceived age did not diminish agreement in warmth and honesty ratings, whereas it did decrease agreement for ratings

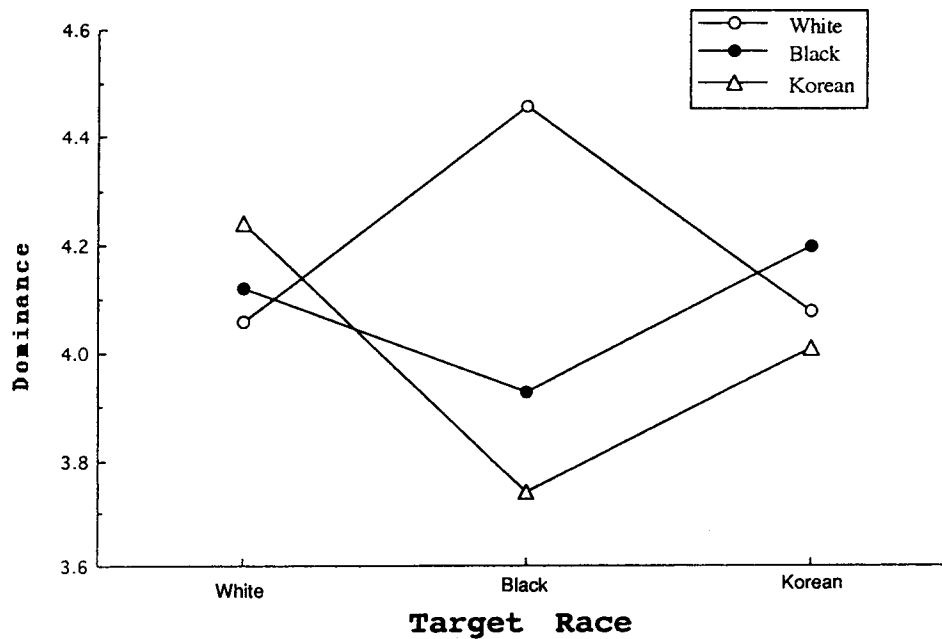


Figure 3. Ratings of dominance as a function of perceiver and target race.

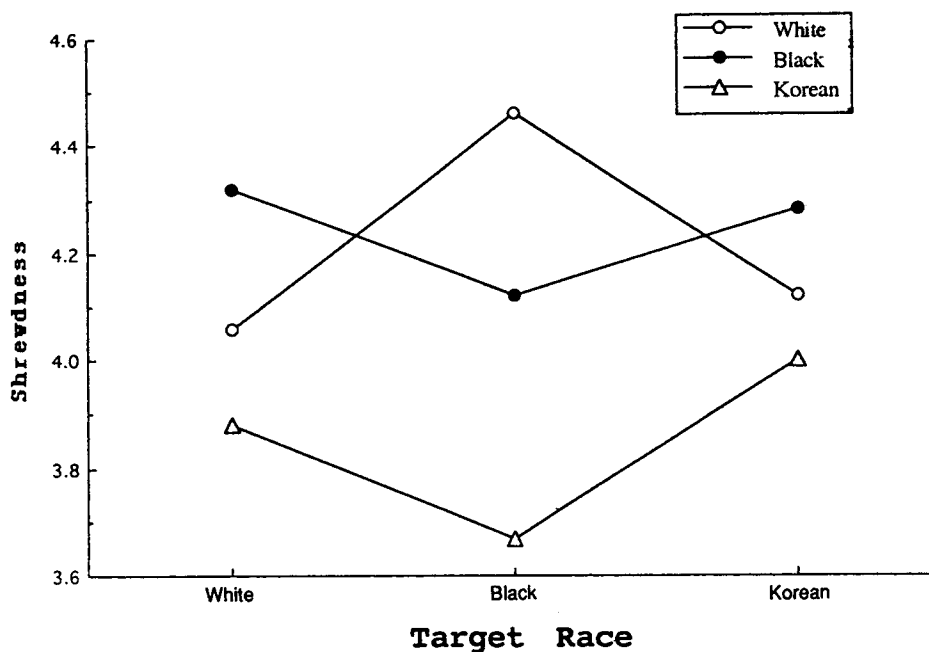


Figure 4. Ratings of shrewdness as a function of perceiver and target race.

of submissiveness, naïveté, and weakness, albeit not so much as did controlling babyfacedness. When facial attractiveness was controlled, there was a slight decline in agreement only for ratings of warmth and honesty.

Crossrace Generality in the Structural Components of Babyfacedness and Attractiveness

The strong interracial agreement in judgments of babyfacedness and attractiveness suggests that there is consistency across perceiver and target race in the facial features that specify these qualities of appearance. Because babyfacedness and attractiveness are configural qualities that cannot be captured by any single structural feature, regression analyses were performed to determine the predictive power of a constellation of features. Specifically, the eight facial structure measures were entered as predictors of perceived babyfacedness and perceived attractiveness in a series of stepwise regression analyses, the results of which are reported in Tables 9 and 10.

Structural babyfacedness. Consistent with past research, a round face, thin eyebrows, and large eyes predicted the babyfacedness of White people's faces across all perceiver groups. The facial characteristics predicting the babyfacedness of Black people's faces were also highly similar across perceiver groups, as shown in Table 9. For White and Black perceivers, the significant predictors were a round face, a wide nose, and large eyes. A wide nose was also a significant predictor for Korean perceivers, together with a small nosebridge. For Korean's faces, two facial characteristics predicted babyfacedness ratings by perceivers of all races: a round face and large eyes. Thin eyebrows also predicted babyfacedness ratings by White and Black perceivers, whereas a wide nose and a small nosebridge also

predicted ratings by White and Korean perceivers. In summary, perceivers of all races showed higher babyfacedness ratings in response to a round face, large eyes, thin eyebrows, and a wide nose. A composite of these features was created by converting the values of each facial characteristic to z scores within each target racial group and then computing the average of the transformed scores. For all perceiver-target groups, this structural babyfacedness composite was unrelated to attractiveness ratings (average $r = .14$) and significantly correlated with babyfacedness ratings (average $r = .69$).

Structural attractiveness. The regression equations revealed that the measured facial characteristics generally accounted for much less of the adjusted variance in attractiveness ratings ($M = 24\%$) than in babyfacedness ratings ($M = 52\%$). Moreover, as shown in Table 10, there was less consistency in the facial characteristics predicting attractiveness ratings across the various perceiver-target combinations. A wide nose predicted the attractiveness of Korean's faces for all perceiver groups, and a small nosebridge predicted the attractiveness of Black people's faces for White and Black perceivers. Large eyes also had relatively consistent effects, predicting the attractiveness of Black people's faces for Black and Korean perceivers and the attractiveness of Korean's faces for White and Black perceivers. On the other hand, small eyes predicted the attractiveness of White people's faces for Korean perceivers. Thick eyebrows was another recurrent predictor, contributing to the attractiveness of Korean's faces for White and Korean perceivers and the attractiveness of Black people's faces for Korean perceivers. However, White people's faces with thin eyebrows were seen as more attractive by Blacks. In sum, the particular features that predicted attractiveness ratings showed little consistency across different groups of faces.

Table 7
Correlations of Babyfacedness and Attractiveness With Trait Impressions

Impression and perceiver race	Babyfacedness			Attractiveness		
	White targets	Black targets	Korean targets	White targets	Black targets	Korean targets
Submissiveness						
White	.79	.80	.90	.22	.32	-.08
Black	.84	.85	.89	.08	.36	.07
Korean	.72	.70	.37	.12	-.06	-.36
Naïveté						
White	.81	.74	.87	.16	.03	-.04
Black	.90	.83	.84	.08	.51	.23
Korean	.49	.63	.48	.32	.30	.38
Physical weakness						
White	.81	.81	.86	.19	.02	-.35
Black	.86	.84	.75	.13	.24	.03
Korean	.68	.75	.38	.12	.11	-.38
Warmth						
White	.79	.54	.32	.58	.76	.54
Black	.61	.57	.59	.27	.54	.24
Korean	.49	.63	.44	.56	.18	.29
Honesty						
White	.70	.58	.66	.60	.64	.35
Black	.70	.71	.67	.34	.54	.62
Korean	.41	.58	.29	.67	.43	.75
Attractiveness						
White	.37	.41	-.05			
Black	.44	.56	.14			
Korean	.01	.36	.17			

Note. For White targets, $n = 20$; for Black and Korean targets, $n = 24$. All correlations in bold type are significant at $p < .05$ or better.

Discussion

The present findings do not support the maxims "they all look alike" and "they all are alike." People did make distinctions regarding the appearance and traits of members of another race when they were asked to judge the faces of specific individuals. First, in keeping with the ecological theory assumption that facial information communicates certain traits, raters showed consensual judgments about the traits of those of another race, as revealed in the finding of high interrater reliability among judges who were all of the same race. Second, evidence that individuals of all races make the same trait dis-

tinctions on the basis of facial information was provided by the finding that interrater agreement tended to be as high when raters were of a different race as when they were of the same race. Thus, for example, a Black man who looked highly honest or dominant to Black perceivers also looked highly honest or dominant to White and Korean perceivers. Similarly, a White or Korean man tended to be judged similarly on trait dimensions by perceivers of all races.

Although the preponderance of the data revealed strong cross-race agreement, there was also some indication of racial differences. Agreement between Koreans and other perceivers was not significant when the submissiveness and naïveté of Korean faces was being judged. Thus, whereas Koreans' judgments of the naïveté and submissiveness of other-race faces reflected a culturally universal response to facial information, their judgments of these attributes in men of their own race was culturally idiosyncratic. Another indication of racial differences in face perception was a tendency for within-race agreement to be higher than between-race agreement for judgments of attractiveness, whereas within- and between-race agreement was equal for judgments of babyfacedness. These trends are consistent with the greater opportunity for culture to influence attractiveness than babyfacedness, which reflects culturally universal maturational changes (Lorenz, 1943; Todd et al., 1980).

Further evidence for the universality of the facial information specifying a babyface was provided by the finding that perceivers of all races showed higher babyface ratings in re-

Table 8
Average Crossrace Agreement in Trait Ratings Partialing out Babyfacedness, Perceived Age, or Attractiveness

Trait	Zero-order correlation	Partialing babyfacedness	Partialing age	Partialing attractiveness
Submissive	.68	.37	.58	.69
Naive	.61	.30	.51	.62
Weak	.81	.61	.72	.82
Warm	.77	.68	.74	.71
Honest	.75	.66	.71	.64

Note. Correlations are averaged across correlations between all pairs of perceiver groups rating all three face groups.

Table 9
Facial Characteristics Predicting Babyfaceness

Perceiver race	Target Race					
	White		Black		Korean	
	Feature	<i>pr</i>	Feature	<i>pr</i>	Feature	<i>pr</i>
White	Round face	.59	Round face	.52	Round face	.74
	Large eyes	.67	Large eyes	.63	Large eyes	.56
	Thin eyebrows	.49	Wide nose	.54	Thin eyebrows	.48
					Wide nose	.45
				Small nosebridge	.54	
				High eyebrows	.41	
	Adjusted $R^2 = .49$ $F(3, 16) = 7.13,$ $p = .003$		Adjusted $R^2 = .51$ $F(3, 20) = 9.08,$ $p < .001$		Adjusted $R^2 = .60$ $F(6, 17) = 6.80,$ $p < .001$	
Black	Round face	.70	Round face	.56	Round face	.51
	Large eyes	.72	Large eyes	.58	Large eyes	.37
	Thin eyebrows	.51	Wide nose	.50	Thin eyebrows	.57
	Adjusted $R^2 = .60$ $F(3, 16) = 10.34,$ $p < .001$		Adjusted $R^2 = .47$ $F(3, 20) = 7.74,$ $p = .001$		Adjusted $R^2 = .50$ $F(3, 20) = 8.83,$ $p < .001$	
Korean	Round face	.57	Wide nose	.58	Round face	.74
	Large eyes	.71	Small nosebridge	.61	Large eyes	.50
	Thin eyebrows	.48			Wide nose	.57
					Small nosebridge	.64
				Small cranium	.62	
	Adjusted $R^2 = .52$ $F(3, 16) = 7.82,$ $p = .002$		Adjusted $R^2 = .45$ $F(2, 21) = 10.45,$ $p < .001$		Adjusted $R^2 = .52$ $F(5, 18) = 6.09,$ $p = .002$	

Note. *pr* = partial correlation.

sponse to a round face, large eyes, a wide nose, and thin eyebrows. In contrast to the universal specification of a babyface, there were no consistent facial feature predictors of attractiveness. Nevertheless, despite some cultural variations in attractiveness judgments, the absolute level of interrater agreement was generally very high. This indicates that attractiveness does have racially universal objective referents in certain facial features, although our measures failed to adequately capture these external referents.

A fourth set of findings indicated that the interracial agreement in trait impressions reflected cross-race commonalities in appearance stereotypes. Not only did perceivers of all races tend to agree on the attractiveness and babyfaceness of faces of all races, but also they showed an attractiveness halo effect and a babyfaceness overgeneralization effect in their impressions of these faces. Moreover, interracial agreement in trait impressions was diminished when judgments of facial attractiveness and babyfaceness were statistically controlled.

The universality of the tendency to perceive babyfaced men as possessing childlike traits—submissiveness, physical weakness, naivete, honesty, and warmth—is consistent with the fact that the psychological attributes as well as the appearance of babies are universal, and it extends past research documenting a babyfaceness overgeneralization effect in judgments of White men by Whites and Koreans (McArthur & Berry, 1987). The

finding that attractive men were perceived as possessing socially good traits—honesty and warmth—extends past research documenting an attractiveness halo effect in crossrace judgments of infants and children by Whites, Blacks, and Hispanics (Langlois & Stephan, 1977; Stephan & Langlois, 1984). However, it is interesting to note that the halo effect for warmth was not consistently manifested across all perceiver–target race groups. Whereas Whites perceived more attractive targets as warmer, regardless of their race, Black perceivers showed this halo effect only for Black targets, and Koreans showed it only for White targets. Thus, certain meanings of attractiveness seem to be culturally variable across perceiver–target racial groups, as it does not always signify warmth. On the other hand, the present findings indicate that attractiveness does universally convey honesty. It may be that the attractiveness halo effect is universal but that the traits that are highly esteemed are not. Positive valuation of honesty may be more universal than positive regard of warmth. Indeed, some evidence for cultural differences in the valuation of warmth is provided by the finding that extraverted people are more highly regarded in western than in eastern cultures (Bond & Forgas, 1984).

In interpreting the consistency of facial stereotypes across the three subject groups, one may suggest that it reflects exposure to media images, because even the Korean subjects are exposed to western faces on television and in the movies. How-

Table 10
Facial Characteristics Predicting Attractiveness

Perceiver race	Target race					
	White		Black		Korean	
	Feature	<i>pr</i>	Feature	<i>pr</i>	Feature	<i>pr</i>
White	No solution		Small nosebridge	.45	Wide nose	.72
			Low eyebrows	.43	Large eyes	.61
					Thick eyebrows	.70
					Round face	.40
				Small cranium	.56	
		Adjusted $R^2 = .16$ $F(2, 21) = 3.28,$ $p = .06$		Adjusted $R^2 = .73$ $F(5, 18) = 13.25,$ $p < .001$		
Black	Thin eyebrows	.46	Small nosebridge	.44	Wide nose	.39
			Large eyes	.40	Large eyes	.50
	Adjusted $R^2 = .17$ $F(1, 18) = 4.89,$ $p = .04$		Adjusted $R^2 = .32$ $F(2, 21) = 6.41,$ $p < .01$		Adjusted $R^2 = .24$ $F(2, 21) = 4.69,$ $p = .02$	
Korean	High eyebrows	.40	Thick eyebrows	.41	Wide nose	.38
	Large eyes	.49	Large eyes	.45	Thick eyebrows	.39
	Adjusted $R^2 = .18$ $F(2, 17) = 3.10,$ $p = .07$		Adjusted $R^2 = .21$ $F(2, 21) = 4.06,$ $p = .03$		Adjusted $R^2 = .18$ $F(2, 21) = 3.59,$ $p < .05$	

Note. *pr* = partial correlation.

ever, such media exposure cannot satisfactorily account for the obtained results. Whereas Koreans may have learned from movies what kinds of men westerners perceive as attractive and how western men with such faces behave, it seems more far-fetched to suggest that they have also learned what kinds of men Westerners perceive as babyfaced and how these people behave. Moreover, a media explanation would need to posit that Korean media images of babyfaced men and their behavior parallel Western images. Even if this were true, to attribute the babyfacedness overgeneralization effect to media portrayals begs the question of its origin. Ecological theory provides a viable explanation for the origin of these appearance stereotypes: Crossrace consistencies in perceptions of men with relatively large eyes and round faces derive from analogous perceptions of babies, who have this facial configuration in all ethnic groups.

A fifth finding indicated that not only did appearance stereotypes yield a similar ordering of individuals on trait dimensions by perceivers of their own and another race, but also that own-race and other-race individuals were often rated with equally fine differentiations. In keeping with the ecological theory tenet that facial information communicates certain traits, similar degrees of own-race and other-race differentiation were obtained for trait ratings to the extent that there was similar differentiation in facial appearance ratings. Individuals of one's own race and a familiar other race were equally differentiated on appearance and trait dimensions. On the other hand, individuals of an unfamiliar other race were less differentiated on trait dimensions than individuals of one's own race. There was

also low differentiation of the babyfacedness of individuals of an unfamiliar other race, which is consistent with the ecological theory assumption that perceptual experience can influence the perceiver's ability to detect various qualities. It is also consistent with Linville et al.'s (1989) model, which holds that out-group homogeneity should be greater the less familiar the out-group. Finally, when own- versus other-race differences in appearance differentiation were statistically controlled, the differences in trait differentiation were eliminated.

The equal trait differentiation of individuals of one's own race and those of a familiar other race observed in the present study should not be attributed to the particular trait impressions that were assessed, because past research has found more homogeneous impressions of out-groups on stereotypic and nonstereotypic traits as well as on positive and negative traits (Linville et al., 1986; Park & Rothbart, 1982). This finding also cannot be attributed to insensitivity in the differentiation measure: Subjects did show more own-race differentiation in comparison with the least familiar racial group. Equal differentiation of own race and familiar race targets also cannot be explained by the number of faces that were judged, because research revealing an other-race effect in face recognition has typically involved showing subjects about 20 target faces, which is comparable to the numbers shown in the present study (e.g., Chance & Goldstein, 1981; Malpass, 1979). Finally, it cannot be attributed to subjects' judging faces of only one race. Past research documenting homogeneity effects in category-based judgments has also required subjects to estimate the distribution of only one category of targets, because this procedure

reduces the demand to differentiate both groups equally to avoid appearing prejudiced (Linville, 1982, Experiment 1; Linville & Jones, 1980, Experiment 3; Linville et al., 1986; Park & Rothbart, 1982; Quattrone & Jones, 1980). Whereas out-group homogeneity effects do not depend on requiring subjects to make direct comparisons between the in-group and the out-group, they may depend on the absence of comparisons among out-group members. Such comparisons are precisely what the present method fostered. By making salient appearance variations among individuals from another racial group, this method yielded the perception of differences in their traits, an effect that is consistent with evidence that attention to subgroups increases outgroup differentiation (Park, Ryan, & Judd, 1992).

Although perceiving differences in appearance among individuals of another race fostered differentiation of their traits, a final result was that it did not eliminate the perception of trait differences between racial groups. This finding parallels results reported by Linville, Fischer, and Salovey (1989), and it indicates that the perceived dispersion of group members can be independent of the extent to which they are seen as fitting a racial stereotype (Park & Judd, 1990). It should be noted that the variation in racial stereotypes across perceiver race indicates that they are not simply an artifact of the particular faces chosen as exemplars of each racial category, but rather reflect racially biased inferences about the faces.

The perceiver race effects on racial stereotypes were intriguing. Both Black and Korean perceivers endorsed the classic Black stereotype, rating Blacks as more submissive and naïve than they rated other racial groups, an effect that may reflect common exposure to the images of Blacks in western media. White perceivers showed a diametrically opposed stereotype of Blacks, rating them as more dominant and shrewd than they rated other racial groups. Coupled with the high rating Whites gave to Blacks on the dimension of physical strength, these findings suggest that Whites see Blacks as very powerful. This stereotype may reflect the fear that White Americans have of Black people in these times of rising violence in American inner cities.

Three questions may be raised concerning the generalizability of the present findings. The first concerns extending the conclusions to other male faces. As noted earlier, there is good reason to believe that the faces used in this study were fairly representative. A second question concerns generalizability to female faces. Because the attractiveness halo and the babyfacedness overgeneralization effect, which contributed substantially to the interracial agreement in impressions in the present study, have been documented for women as well as men (e.g., Eagly et al., 1991; Zebrowitz & Montepare, 1992), there is also good reason to believe that the present findings would generalize to judgments of women. Moreover, research comparing male and female faces has revealed that both male and female judges showed more refined distinctions, more variability, and more consensus in ratings of female faces (Schulman & Hoskins, 1986), which suggests that the present findings would, if anything, be stronger for female than for male faces.

A third question concerns generalizability to real-world interracial judgments. Whereas the present research examined

the intra- and interracial impressions that are elicited by individuals' faces, prior research has generally investigated impressions that are based on a racial category label. It has been argued that such category-based processes are characteristic of many real-life social settings (Brewer, 1988; Fiske & Neuberg, 1990). However, according to the ecological theory of social perception (McArthur & Baron, 1983), the conditions that yield such processing in experimental settings are rare in real life encounters, which provide individuating, perceptual information about various stimulus persons. Regardless of which judgment situation is most common, the present findings suggest that in a face-to-face interracial interaction some of the reactions that people encounter will not be foreign to them. Specifically, people in interracial interactions may confront the same expectations regarding their traits as they are accustomed to find in interactions with people of their own race.

Finally, in considering the generalizability question, it should be recalled that the purpose of the present study was not to determine what people most typically do when making other-race trait judgments but rather to determine whether attention to facial information produces reliable, interracially consensual, differentiated, and theoretically predictable judgments. That facial information does have these effects is consistent with the ecological theory assumption that people's perceptible attributes communicate certain behavioral propensities. It is also consistent with models of impression formation, which hold that the reason outgroup members generally appear more similar to each other is because their individuating attributes receive less attention (Brewer, 1988; Fiske & Neuberg, 1990; Smith & Zarate, 1992; Turner, 1987).

That other-race individuation is fostered when people merely notice global appearance differences among faces of another race may have important implications for combatting prejudice and racism. According to models proposed by Brewer and Miller (1984) and Fiske and Neuberg (1990), the first step in reducing prejudice is for people to adopt a more individuated, personalized approach to processing information about others, because this reduces the tendency to view them merely as representative of a particular racial category. Some research suggests that increased individuation of out-group members does in fact reduce out-group discrimination (Wilder, 1978). However, other research indicates that to have such effects, out-group individuation must also weaken category boundaries such that the perceived similarity among group members is no greater within than between groups (Miller & Brewer, 1986).

Whereas forming impressions of individuals on the basis of facial information tended to equalize the perceived similarity among individuals within each racial group, some differences were nevertheless perceived between groups. Although an in-group favoritism effect was not observed, certain racial stereotypes were, indicating that in some cases the average variability perceived between groups remained higher than the average within-group variability. Thus, although perceivers learned that there are some people who look shrewd and others who look naïve, some who look weak and others who look strong, these differentiations sometimes co-occurred with the perception of differences in the central tendencies for different racial groups. Still, the recognition of variability in appearance

within another racial group could mitigate racist responses to particular individuals within that group. Perceivers who attend to such variability, even if they are White, may be less likely than the aforementioned police officers to treat a Black basketball player as they would a bank robber.

References

- Adams, G. R. (1977). Physical attractiveness research. Toward a developmental social psychology of beauty. *Human Development, 20*, 217–239.
- Anthony, T., Copper, C., & Mullen, B. (1992). Cross-racial facial identification: A social cognitive integration. *Personality and Social Psychology Bulletin, 18*, 296–301.
- Bernstein, I. H., Lin, T. D., & McClellan, P. (1982). Cross vs. within-racial judgments of attractiveness. *Perception and Psychophysics, 32*, 495–503.
- Berry, D. S., & McArthur, L. Z. (1985). Some components and consequences of a babyface. *Journal of Personality and Social Psychology, 48*, 312–323.
- Berry, D. S., & McArthur, L. Z. (1986). Perceiving character in faces: The impact of age-related craniofacial changes on social perception. *Psychological Bulletin, 100*, 3–18.
- Berscheid, E., & Walster, E. (1974). Physical attractiveness. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 7, pp. 158–216). San Diego, CA: Academic Press.
- Bond, M. H., & Forgas, J. P. (1984). Linking person perception to behavior intention across cultures: The role of cultural collectivism. *Journal of Cross-Cultural Psychology, 15*, 337–352.
- Braithwaite, U. A. (1986). Old age stereotypes: Revealing contradictions. *Journal of Gerontology, 41*, 353–360.
- Brewer, M. (1988). A dual process model of impression formation. In T. K. Srull & R. S. Wyer, Jr. (Eds.), *Advances in Social Cognition* (Vol. 1, pp. 1–36). Hillsdale, NJ: Erlbaum.
- Brewer, M. B., & Miller, N. (1984). Beyond the contact hypothesis: Theoretical perspectives on desegregation. In N. Miller & M. B. Brewer (Eds.), *Groups in contact: The psychology of desegregation* (pp. 281–302). San Diego, CA: Academic Press.
- 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.
- Bull, R., & Rumsey, N. (1988). *The social psychology of facial appearance*. New York: Springer-Verlag.
- Chance, J. E., & Goldstein, A. G. (1981). Depth of processing in response to own- and other-race faces. *Personality and Social Psychology Bulletin, 7*, 475–480.
- Cunningham, M. R. (1986). Measuring the physical in physical attractiveness: Quasi-experiments on the sociobiology of female facial beauty. *Journal of Personality and Social Psychology, 50*, 925–935.
- Cunningham, M. R. (1990). What do women want? Facialmetric assessment of multiple motives in the perception of male facial physical attractiveness. *Journal of Personality and Social Psychology, 59*, 61–72.
- Dovidio, J. F., & Gaertner, S. L. (1981). The effects of race, status, and ability on helping behavior. *Social Psychology Quarterly, 44*, 192–203.
- Eagly, A. H., Ashmore, R. D., Makhijani, M. G., & Longo, L. C. (1991). What is beautiful is good, but . . . : A meta-analytic review of research on the physical attractiveness stereotype. *Psychological Bulletin, 110*, 109–128.
- Feingold, A. (1992). Good-looking people are not what we think. *Psychological Bulletin, 111*, 304–341.
- Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 23, pp. 1–74). San Diego, CA: Academic Press.
- Gaertner, S. L., & Dovidio, J. F. (1986). The aversive form of racism. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism*. (pp. 61–89). San Diego, CA: Academic Press.
- Gaertner, S. L., & McLaughlin, J. P. (1983). Racial stereotypes: Associations and ascriptions of positive and negative characteristics. *Social Psychology Quarterly, 46*, 23–30.
- Goldstein, A. G. (1979). Race-related variation of facial features: Anthropometric data I. *Bulletin of the Psychonomic Society, 13*, 187–190.
- Goldstein, A. G., & Chance, J. (1976). Measuring psychological similarity of faces. *Bulletin of the Psychonomic Society, 7*, 407–408.
- Goldstein, A. G., & Chance, J. (1978). Judging face similarity in own and other races. *The Journal of Psychology, 98*, 185–193.
- Hatfield, E., & Sprecher, S. (1986). *Mirror, mirror . . . The importance of looks in everyday life*. Albany, NY: State University of New York Press.
- Jones, E. E., Wood, G. C., & Quattrone, G. A. (1981). Perceived variability of personal characteristics in in-groups and out-groups: The role of knowledge and evaluation. *Personality and Social Psychology Bulletin, 7*, 523–528.
- Judd, C. M., & Park, B. (1988). Out-group homogeneity: Judgments of variability at the individual and group levels. *Journal of Personality and Social Psychology, 54*, 778–788.
- Karlins, M., Coffman, T. L., & Walters, G. (1969). On the fading of social stereotypes: Studies in three generations of college students. *Journal of Personality and Social Psychology, 13*, 1–16.
- Keating, C. F. (1985). Gender and the physiognomy of dominance and attractiveness. *Social Psychology Quarterly, 48*, 61–70.
- Keating, C. F., Mazur, A., & Segall, M. H. (1981). A cross-cultural exploration of physiognomic traits of dominance and happiness. *Ethology and Sociobiology, 2*, 41–48.
- Langlois, J. H., & Roggman, L. A. (1990). Attractive faces are only average. *Psychological Science, 1*, 115–121.
- Langlois, J. H., & Stephan, C. (1977). The effects of physical attractiveness and ethnicity on children's behavioral attributions and peer preferences. *Child Development, 48*, 1694–1698.
- Linville, P. W. (1982). The complexity–extremity effect and age-based stereotyping. *Journal of Personality and Social Psychology, 42*, 193–211.
- Linville, P. W., Fischer, G. W., & Salovey, P. (1989). Perceived distributions of characteristics of in-group and out-group members: Empirical evidence and a computer simulation. *Journal of Personality and Social Psychology, 57*, 165–188.
- Linville, P. W., & Jones, E. E. (1980). Polarized appraisals of outgroup members. *Journal of Personality and Social Psychology, 38*, 689–703.
- Linville, P. W., Salovey, P., & Fischer, G. W. (1986). Stereotyping and perceived distributions of social characteristics: An application to in-group–outgroup perception. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 165–208). San Diego, CA: Academic Press.
- Locksley, A., Borgida, E., Brekke, N., & Hepburn, C. (1980). Sex stereotypes and social judgment. *Journal of Personality and Social Psychology, 39*, 821–831.
- Locksley, A., Hepburn, C., & Ortiz, V. (1982). Social stereotypes and judgment of individuals: An instance of the base-rate fallacy. *Journal of Experimental Social Psychology, 18*, 23–42.
- Lorenz, K. (1943). Die angeborenen Formen möglicher Erfahrung [The innate forms of potential experience.] *Zeitschrift für Tierpsychologie, 5*, 234–409.

- Malpass, R. S. (1979). A cross-cultural face recognition field manual: Description and a validation study. In L. H. Eckensberger, W. J. Lonner, & Y. H. Poortinga (Eds.), *Cross-cultural contributions to psychology* (pp. 27-39). Lisse, The Netherlands: Swets & Zeitlinger.
- Malpass, R. S., & Kravitz, J. (1969). Recognition of faces of own and other race. *Journal of Personality and Social Psychology*, *13*, 330-334.
- Maret, S. M. (1983). Attractiveness ratings of photographs of blacks by Cruzans and Americans. *Journal of Psychology*, *115*, 113-116.
- Maret, S. M., & Harling, C. A. (1985). Cross-cultural perceptions of physical attractiveness: Ratings of photographs of Whites by Cruzans and Americans. *Perceptual and Motor Skills*, *60*, 163-166.
- McArthur, L. Z. (1982). Judging a book by its cover: A cognitive analysis of the relationship between physical appearance and stereotyping. In A. Hastorf & A. Isen (Eds.), *Cognitive Social Psychology* (pp. 149-211). New York: Elsevier North Holland.
- McArthur, L. Z., & Apatow, K. (1983-84). Impressions of babyfaced adults. *Social Cognition*, *2*, 315-342.
- McArthur, L. Z., & Baron, R. M. (1983). Toward an ecological theory of social perception. *Psychological Review*, *90*, 215-247.
- McArthur, L. Z., & Berry, D. S. (1987). Cross-cultural agreement in perceptions of babyfaced adults. *Journal of Cross-Cultural Psychology*, *18*, 165-192.
- Miller, N., & Brewer, M. B. (1986). Categorization effects on ingroup and outgroup perception. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 209-230). San Diego, CA: Academic Press.
- Park, B., & Judd, C. M. (1990). Measures and models of perceived group variability. *Journal of Personality and Social Psychology*, *59*, 173-191.
- Park, B., & Rothbart, M. (1982). Perception of out-group homogeneity and levels of social categorization: Memory for the subordinate attributes of in-group and out-group members. *Journal of Personality and Social Psychology*, *42*, 1051-1068.
- Park, B., Ryan, C. S., & Judd, C. M. (1992). Role of meaningful subgroups in explaining differences in perceived variability for in-groups and out-groups. *Journal of Personality and Social Psychology*, *63*, 553-567.
- Pratto, F., & Bargh, J. A. (1991). Stereotyping based on apparently individuating information: Trait and global components of sex stereotypes under attention overload. *Journal of Experimental Social Psychology*, *27*, 26-47.
- Quattrone, G. A. (1986). On the perception of a group's variability. In S. Worchel & W. Austin (Eds.), *Psychology of intergroup relations* (Vol. 2, pp. 25-48). Chicago: Nelson-Hall.
- Quattrone, G. A., & Jones, E. E. (1980). The perception of variability within ingroups and outgroups: Implications for the law of small numbers. *Journal of Personality and Social Psychology*, *38*, 141-152.
- Schulman, G. I., & Hoskins, M. (1986). Perceiving the male versus the female face. *Psychology of Women Quarterly*, *10*, 141-154.
- Shapiro, P. N., & Penrod, S. (1986). Meta-analysis of facial identification studies. *Psychological Bulletin*, *100*, 139-156.
- Shepherd, J. (1981). Social factors in face recognition. In G. Davies, H. Ellis, & J. Shepherd (Eds.), *Perceiving and remembering faces* (pp. 55-80). San Diego, CA: Academic Press.
- Shepherd, J., & Deregowski, J. B. (1981). Races and faces: A comparison of the responses of African and Europeans to faces of the same and different races. *British Journal of Social Psychology*, *20*, 125-133.
- Stephan, C., & Langlois, J. H. (1984). Baby beautiful: Adult attributions of infant competence as a function of infant attractiveness. *Child Development*, *55*, 576-585.
- Smith, E. R., & Zarate, M. A. (1992). Exemplar-based model of social judgment. *Psychological Review*, *99*, 3-21.
- Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behavior. *European Journal of Social Psychology*, *1*, 149-178.
- Todd, J. T., Mark, L. S., Shaw, R. E., & Pittenger, J. B. (1980). The perception of human growth. *Scientific American*, *242*, 106-114.
- Turner, J. C. (1987). *Rediscovering the social group: A self-categorization theory*. Oxford, England: Basil Blackwell.
- Vinacke, W. E. (1949). Stereotyping among national-racial groups in Hawaii: A study in ethnocentrism. *Journal of Social Psychology*, *30*, 265-291.
- Wagatsuma, E., & Kleinke, C. L. (1979). Ratings of facial beauty by Korean-American and Caucasian females. *Journal of Social Psychology*, *109*, 299-300.
- Wilder, D. A. (1978). Reduction of intergroup discrimination through individuation of the out-group. *Journal of Personality and Social Psychology*, *36*, 1361-1374.
- Zebrowitz, L. A., & Montepare, J. M. (1992). Impressions of babyfaced males and females across the lifespan. *Developmental Psychology*, *28*, 1143-1152.
- Zebrowitz-McArthur, L., & Montepare, J. M. (1989). Contributions of a babyface and a childlike voice to impressions of moving and talking faces. *Journal of Nonverbal Behavior*, *13*, 189-203.
- Zuckerman, M. (1990). Some dubious premises in research and theory on racial differences. *American Psychologist*, *45*, 1297-1303.

Received July 2, 1992

Accepted December 30, 1992 ■