Memory & Cognition
1994, 22 (2), 208-224

Structural aspects of face recognition and the other-race effect

ALICE J. O'TOOLE
University of Texas at Dallas, Richardson, Texas

KENNETH A. DEFFENBACHER
University of Nebraska, Omaha, Nebraska

and

DOMINIQUE VALENTIN and HERVÉ ABDI
University of Texas at Dallas, Richardson, Texas

The other-race effect was examined in a series of experiments and simulations that looked at the relationships among observer ratings of typicality, familiarity, attractiveness, memorability, and the performance variables of d' and criterion. Experiment 1 replicated the other-race effect with our Caucasian and Japanese stimuli for both Caucasian and Asian observers. In Experiment 2, we collected ratings from Caucasian observers on the faces used in the recognition task. A Varimax-rotated principal components analysis on the rating and performance data for the Caucasian faces replicated Vokey and Read's (1992) finding that typicality is composed of two orthogonal components, dissociable via their independent relationships to: (1) attractiveness and familiarity ratings and (2) memorability ratings. For Japanese faces, however, we found that typicality was related only to memorability. Where performance measures were concerned, two additional principal components dominated by criterion and by d' emerged for Caucasian faces. For the Japanese faces, however, the performance measures of d' and criterion merged into a single component that represented a second component of typicality, one orthogonal to the memorability-dominated component. A measure of *face representation quality* extracted from an autoassociative neural network trained with a majority of Caucasian faces and a minority of Japanese faces was incorporated into the principal components analysis. For both Caucasian and Japanese faces, the neural network measure related both to memorability ratings and to human accuracy measures. Combined, the human data and simulation results indicate that the memorability component of typicality may be related to small, local, distinctive features, whereas the attractiveness/familiarity component may be more related to the global, shape-based properties of the face.

For many years, it has been suspected that faces of one's own race are recognized more accurately than faces of other races (Feingold, 1914). Indeed, there is abundant empirical evidence for this *other-race* phenomenon, as two recent metaanalyses of the face recognition literature attest (Bothwell, Brigham, & Malpass, 1989; Shapiro & Penrod, 1986). In addition to the empirical support for this phenomenon, the other-race effect is widely known outside of the laboratory. Deffenbacher and Loftus (1982), for example, showed that approximately half of potential jurors believe that a recognition bias exists.

Recent advances have been made by Vokey and Read (1992) in understanding how facial characteristics (e.g., rated typicality), which are important for recognition of same-race faces, are interrelated. They examined the structure inherent in ratings of typicality, familiarity, memorability, likableness, and attractiveness, using a principal components analysis, and looked also to see how different components of this structure related to various measures of recognition.

The purpose of the present study was twofold. First, we wished to undertake a similar structural analysis for other-race face recognition. It has long been known that there is a quantitative difference in recognition accuracy for same and other-race faces, but whether or not there are qualitative differences in how other-race faces are processed is less certain. A paradigm similar to that used by Vokey and Read (1992), which uncovers the structural aspects of recognition, seemed to us to be an excellent method for comparing the qualitative aspects of same- and

Thanks are due June Chance and Al Goldstein for providing the Caucasian and Japanese faces used in the present experiments and simulations and to James C. Bartlett for helpful comments throughout the entire course of the project. We would like to thank also James I. Chumbley, John W. Shepherd, and John R. Vokey for very helpful comments on a previous version of the manuscript. We are grateful also to Barbara Edwards and to Ray Zhu for assistance in testing subjects. Requests for reprints should be sent to A. J. O'Toole, School of Human Development, GR4.1, University of Texas at Dallas, Richardson, TX 75083-0688 (e-mail: otoole@utdallas.edu).

—Accepted by previous editor, Margaret Jean Intons-Peterson

other-race face recognition. The second purpose of this study was to incorporate a more stringent control than has been used in previous work for assessing the relationship between individual or combined ratings of faces and signal detection measures of recognition performance. We have incorporated this control both for same- and for other-race faces, and we will show that, while the structural aspects of the ratings of same-race and other-race faces remain very similar to those found by Vokey and Read (1992), some qualification of the relationship between ratings and performance may be needed.

In the present study, we have looked at the effects of rated typicality, attractiveness, context-free familiarity, and memorability. We have also used Vokey and Read's (1988) measure of false recognition—an observer judgment that a face has appeared previously, when, in fact, no faces have appeared more than once. Henceforth, we shall refer to the latter measure as *repetition*. Since Vokey and Read (1992) have provided a thorough discussion of the nature of the effects for each of these variables, we shall limit ourselves to a discussion of how their results qualify the generally accepted role of typicality in predicting recognizability and how this may give us insight into understanding other-race face recognition.

Over the years, the relationship between rated typicality and recognizability of faces has been studied by a number of investigators (e.g., Cohen & Carr, 1975; Going & Read, 1974; Light, Kayra-Stuart, & Hollander, 1979). This relationship has generally been interpreted in terms of the existence of a facial prototype, with typical faces being recognized less well than unusual faces. Recent findings of Vokey and Read (1992), however, indicate that rated typicality is a more complicated concept than had been thought previously. Applying a principal components analysis followed by a Varimax rotation to a set of faces rated for typicality, memorability (i.e., "one that the observer thought would be easy to remember"), familiarity (i.e., "a face that they believe they may have seen around campus"), attractiveness, and likableness, Vokey and Read show convincingly that the rated typicality of faces is composed of two orthogonal components—one related to rated familiarity, attractiveness, and likability, and a second component related inversely to the rated memorability of a face. The assertion that rated typicality is composed of two orthogonal components means that observers' typicality ratings are a joint function of two independent aspects of a face. These aspects can be dissociated via their independent relationships to familiarity, attractiveness, and likability, which appear on one axis of the principal components analysis,[1] and to memorability, which appears on a separate axis.

The finding that rated typicality is composed of two orthogonal components is consistent with earlier work (Valentine & Bruce, 1986; Vokey & Read, 1988) indicating various measurable dissociations among typicality, measures of distinctiveness, and context-free familiarity. While these terms are somewhat confusing, the general idea is that the part of typicality due to the distinctiveness of faces affects the encoding and retrieval processes, whereas the part of typicality called *context-free* (see, e.g., Bartlett, Hurry, Thorley, 1984) or *general* familiarity is an aspect of typicality that is due to an experience or episode that fails to index a source or episode of encounter.

Vokey and Read (1992) further extend their findings by applying a regression analysis to the general familiarity and memorability components to predict recognition discrimination, criterion, and hit and false alarm rates. They reason that if the effect of typicality on recognition is a function of both general familiarity and memorability, each of these should be a significant predictor of recognition performance. To do this, they derived a regression equation composed of differential additive weightings of the general familiarity and memorability components to predict discrimination performance. In fact, they found that both familiarity and memorability were significant predictors of discrimination performance, but not of criterion. They also found a dissociation between the predictability of hit and false alarm rates by the separate components of typicality. Specifically, both the familiarity and memorability components predicted false alarms, but only the memorability component predicted hits.

One difficulty that arises in interpreting relationships between observer ratings such as typicality (or the components of typicality) and performance measures such as d', criterion, hit rate, and false alarm rate is that there are systematic covariations occurring among measures derived from signal detection theory. In particular, both hit and false alarm rate confound d' and criterion. For example, if the false alarm rate for a face is high, it may be because the face has a low d' (its "old" or "new" status is difficult to discriminate, given its value on the attribute being used to make recognition judgments) or because observers have adopted a loose criterion for categorizing the face as "old," or because of some combination of the two. Thus, certain combinations of recognition difficulty and criterion within a set of faces can artificially inflate or deflate correlations between facial characteristics and either hit or false alarm rate. Specifically, false alarm and hit rate are each composed of two independent components, d' and criterion. Since the technique of univariate correlation is not sensitive to the potentially multidimensional nature of the variables to be correlated, interpreting correlations between rating data and hit or false alarm rate is problematic, because neither hit nor false alarm rate alone is interpretable without knowing both the d' and criterion variations within a set of faces. These difficulties are even more profound when a relationship exists between d' and criterion for a set of faces (i.e., when d' and criterion are not independent). This confounding is important in the present study for Caucasian observers recognizing Japanese faces. To address this problem, we applied a principal components analysis to the rating and performance data simultaneously. Using d' and criterion as performance measures in the principal components analysis provides a simplified and more comprehensive picture of the relationships among performance measures and facial characteristic ratings.

This paper is organized as follows. The first experiment replicates the other-race effect with our stimuli. We used both Asian and Caucasian observers for this study, to ensure equal discriminability of the faces in the two categories. In Experiment 2, we used only Caucasian observers to obtain ratings of typicality, attractiveness, memorability, and familiarity on the faces used in Experiment 1. We incorporated these measures into separate principal components analyses for the Caucasian and Japanese faces. Finally, we trained an autoassociative neural network with a majority of Caucasian faces and a minority of Japanese faces. Face images are represented in the network in a parallel and distributed way that allows for relatively natural interference effects related to the similarity of the faces to each other. We (O'Toole, Deffenbacher, Abdi, & Bartlett, 1991) have shown recently that this kind of autoassociative memory can model some qualitative aspects of the other-race effect. Valentine and Ferrara (1991) have also used an autoassociative memory to simulate the effects of typicality in face identification and categorization. Although their model did not use face encodings that preserve the perceptual information in faces, their results illustrate the potential usefulness of the autoassociator as a tool for modeling psychological phenomena in face processing.

In the present study, the autoassociative network stores face images and yields a measure of *representation quality* for each face. We will define this measure precisely in the autoassociative model section. For the present purposes, suffice it to say that when the model represents new or unlearned faces, the representation quality is dependent both on the interface similarity of the learned faces and on the similarity of the new faces to the learned faces. Intuitively, this can be seen as a confounded measure of typicality, familiarity, and memorability. Since the model was able to replicate some qualitative effects of the other-race effect, we incorporated the model's representation quality measure directly into the principal components analysis of human rating and performance measures taken on the same faces used in Experiments 1 and 2. We show that this measure is related to human observer ratings and to recognition performance for both the Caucasian and the Japanese faces.

EXPERIMENT 1

The purpose of this experiment was to replicate the other-race effect with a yes/no recognition task, testing Caucasian and Asian observers with Japanese and Caucasian faces.

Method

Caucasian observers. Twenty-six Caucasian observers from the University of Texas at Dallas undergraduate population were recruited in exchange for a core psychology course research credit. They were assigned in pseudorandom fashion to one of three different random-order stimulus presentation conditions and were tested in groups of 1 to 5 members.

Asian observers. We had difficulty in obtaining a sample of Asian observers, since we generally recruit observers from the population of undergraduate psychology majors, which has only a small proportion of Asian students. Nonetheless, we were able to recruit 9 students of Asian backgrounds (Chinese, Korean, and Vietnamese), varying in the number of years they had spent in the United States (from 1 to 17 years). Despite the modest size of the sample, we were able to demonstrate the other-race effect reliably.

Stimuli. Three hundred and nineteen faces were digitized from slides to a resolution of 16 gray levels, using a Fotovix digitizer attached to a Zenith 286-based microcomputer equipped with a 16-bit TARGA board (True Vision). The faces were of young adults, with each race set consisting of approximately half male and half female faces. None of the slides pictured persons with facial hair or glasses. The images were aligned so that the eyes of all faces were at about the same height and so that the center point between the eyes was at the same place in all the photographs. The faces were not normalized explicitly for size. However, since the photographs were taken under the same general distance conditions, the faces were roughly equal in size. All experiments and computer simulations used subsets of these faces. The images were also cropped to eliminate clothing cues.

The digitized faces were transferred to videotape. Three videotapes, each consisting of a study list and a test list of faces, were created by using different random orders of faces. The study phase of each tape consisted of the presentation of 120 randomly ordered Caucasian and Japanese faces for 3 sec each, with an interstimulus interval of 5 sec. These 120 faces included 30 Caucasian female (CF), 30 Caucasian male (CM), 30 Japanese female (JF), and 30 Japanese male (JM) faces. In the test phase of each tape, 120 target (*old*) faces and 120 *new* faces were presented for 5 sec each, with an interstimulus interval of 5 sec. The new ones consisted of equal sets of 30 CM, 30 CF, 30 JM, and 30 JF faces. These test faces were blocked by race, and the order of presentation of the Japanese and Caucasian faces was counterbalanced across observers.

Procedure. Observers were instructed that they would be asked first to view a series of Japanese and Caucasian faces, after which they would be asked to participate in a recognition test. Because of the length of the study list, the observers were offered a break halfway through the list. All observers took a 5-min break between study and test lists. During the recognition test, the observers indicated their certainty that a face had previously appeared on the study list by circling the appropriate number on a response sheet. The response sheet contained a 6-point rating scale, varying from "1," *absolute certainty that the face was "old,"* to "6," *absolute certainty that it was "new."* These certainty ratings were used categorically to calculate a d', with ratings less than or equal to three considered an "old" response and ratings greater than or equal to four considered a "new" response.

Results and Discussion

We performed a signal detection analysis of the data and used the d's in a correlated t test, with each observer contributing a d' score for Caucasian and for Japanese faces. Since there were no differences as a function of study/test list order, the results were collapsed across the three lists. As predicted, the Caucasian observers were significantly better at recognizing Caucasian faces than Japanese faces. For the Caucasian observers, mean d's were 1.36 and 0.77 for Caucasian and Japanese faces, respectively [$t(25) = 3.95$, $p < .001$]. Comparable means for the Asian observers were .84 and 1.53 for the Caucasian and Japanese faces, respectively [$t(8) = 6.16$, $p < .001$]. Nonparametric versions of d' such as A' gave identical results, so d' was used as the measure of recognition accuracy in subsequent analyses.

EXPERIMENT 2

The purpose of this experiment was to obtain ratings of typicality, attractiveness, memorability, familiarity, and repetition on the Caucasian and Japanese faces used in Experiment 1. Because of the large amount of data involved, and because of the fact that Experiment 1 established the equal discriminability of the two sets of faces, we used only Caucasian raters in this experiment. Because of the possibility of overtaxing observers by asking them to rate a face for all of these qualities, and because of the possibility of interference among some combinations of ratings, we subdivided the task into parts. Different groups of observers rated faces for (1) typicality, (2) familiarity and repetition, and (3) attractiveness and memorability. A summary of the instructions for the rating tasks appears in Table 1.

Method

Observers. Caucasian volunteers were recruited from the University of Texas at Dallas undergraduate psychology program and again received a research credit for a core course as compensation. Each observer participated in only one of the rating subexperiments. Twenty-five observers rated faces for typicality, 20 observers rated familiarity and repetition, and 20 observers rated memorability and attractiveness.

Stimuli. The stimuli were the recognition test sections of the videotapes described above. Each tape contained all 240 faces, blocked by race. The order of rating Japanese versus Caucasian faces was counterbalanced across the observers in each of the subexperiments.

Procedure. Observers were tested in groups of 1 to 5 and were given response sheets to make the appropriate ratings. Again, because of the length of the list, the observers were given breaks periodically.

Results

The typicality, familiarity, memorability, attractiveness, and repetition ratings for each face were collapsed across the observers in Experiment 2. For clarity of interpretation and ease of reading, we present all the rating data so that the high numbers indicate high values of the facial characteristic in question. The proportion of "yes" responses is reported for the yes/no scales. The recognition performance measures of d' and criterion for each face were also collapsed across the Caucasian observers in Experiment 1. That is, d' and criterion scores were calculated for every face for which both hit and false alarm rates could be generated, rates based on the responses of different observers. Hit and false alarm rates could not be generated for all *faces*, since the selection of faces for the three videotapes was random. Thus, some faces did not appear as both *old* and *new* in Experiment 1. From the faces that were seen as both *old* and *new* across observers, a randomly chosen subset of 80 Caucasian faces (40 male and 40 female) and 80 Japanese faces (40 male and 40 female) was used for all subsequent analyses. Table 2 contains the means and standard deviations for all seven of these variables for both the Japanese and the Caucasian faces.

In addition to the difference seen in Experiment 1 in recognition performance for Caucasian observers with Caucasian and Japanese faces, there were statistically reli-

Table 1
Rating Tasks Summary

Rating		
Typicality*	Imagine you were to meet someone in a train station—How difficult would it be to pick the person out of a crowd? For Japanese faces, the observers were instructed to imagine that they were in a train station in Japan. 4-point rating scale	Valentine & Bruce (1986)
Familiarity	Is the face confusable with someone you know?† yes/no?	Vokey & Read (1988)
Memorability	Is the face easy to remember? 4-point rating scale	Vokey & Read (1992)
Repetition	Has the face appeared previously? yes/no?	Vokey & Read (1988)
Attractiveness	Is the face attractive? 4-point rating scale	Vokey & Read (1992)

*This is not the definition used by Vokey and Read (1992). They described a "typical" face to their subjects as one that was "average." The similarity of our results to those of Vokey and Read (1992) indicates that these definitions may be generally compatible.

†To obtain somewhat higher affirmative rates, Vokey and Read (1992) modified this definition to imply that the photographs were taken from students around campus and, hence, that they might have been seen by the observers. In the present experiment, we used the 1988 definition and replicated the results found by Vokey and Read (1992) for familiarity, thus indicating that the two definitions are capturing a similar aspect of the faces.

Table 2
Performance and Rating Measures for Caucasian and Japanese Faces by Caucasian Observers

	Faces			
	Caucasian		Japanese	
Rating	*M*	*SD*	*M*	*SD*
	Four-Point Scale			
Attractiveness	2.21	.40	2.18	.37
Memorability	2.48	.28	2.46	.21
Typicality	1.55	.36	2.20	.30
	Yes/No			
Familiarity	.10	.06	.07	.06
Repetition	.23	.16	.29	.19
	Performance			
d'	1.36	1.00	.77	.97
Criterion	−.11	.59	−.05	.64

able differences in ratings of Japanese and Caucasian faces for the attributes of typicality [$t(158) = 4.83, p < .0001$], familiarity [$t(158) = 2.45$, $p < .02$], and repetition [$t(158) = 2.03, p < .05$]. In general, Caucasian observers rated Japanese faces to be more typical and less familiar than Caucasian faces. Further, Caucasian observers mistakenly judged Japanese faces as having appeared earlier in the series more often than Caucasian faces.

Before presenting the principal components analyses, we need to define and explain some conventions that we adopt in interpreting and presenting the results of this and all subsequent analyses in this paper. Although the conventions we adopt are necessarily arbitrary, we believe that they will facilitate comparisons: (1) between the analyses for Japanese and Caucasian faces, (2) between the human and simulation data, and (3) between the present study and that of Vokey and Read (1992).

First, in any application of principal components or factor analysis, a decision must be made concerning the number of axes to present. The axes in a principal components analysis are ordered by their eigenvalues, which are directly related to the percentage of variance explained by each axis.[2] If all axes were used, 100 percent of the variance in the rating and performance measures would be explained.[3] There is no universally established rule to indicate the number of axes of a principal components analysis that should be retained in any given case, although there are many rules of thumb that give quite different recommendations. In the present study, we have chosen to retain axes that explain more than 10% of the variance.

Second, with principal components or factor analysis, it is impossible to test the statistical reliability of the loadings or differences between loadings for the individual variables. Thus, it is necessary to chose an arbitrary threshold for loadings that will be considered important in interpreting the axes. In the present paper, we will arbitrarily confine our major conclusions to be based only on loadings greater than or equal to .30. To facilitate reading the tables presenting these loadings, all loadings greater than or equal to this value will be in boldface.

Finally, since we will be presenting eight structurally similar principal components analyses, for ease of understanding and comparison, we have arranged all tables of axis loadings into a set order. The proportion of explained variance for each axis, and hence, the information about the actual order of the axes in any given analysis, appears under each of the columns. Additionally, for convenience, we include the ordered rank of the axes below the proportion of variance explained. It was possible in the present study to arrange the axes in a fixed order, since in all cases the rotated axes can be interpreted, at least in a prima facie way, as either rating dominated or performance dominated. To coordinate our discussion with that of Vokey and Read (1992), we have arranged the tables as follows. Rating-dominated axes appear in the first two columns, with an axis comparable to Vokey and Read's (1992) memorability axis in column 1 and one comparable to their general familiarity axis in column 2. For Caucasian faces, there were two performance axes, one related to criterion and one related to accuracy; they appear for all the analyses in columns 3 and 4, respectively. For Japanese faces, a single performance-dominated axis appears in column 3 of all the analyses.

Separate Varimax-rotated principal components analyses[4] were performed for the Caucasian and Japanese faces by using the z scores (calculated by race) for the performance and rating variables. The z scores were used to ensure that the principal components analysis results were not perturbed by differences in the absolute range of scores for the rating scales and performance measures.

For completeness, intercorrelations among ratings and performance measures for both Caucasian and Japanese faces appear in Table 3. These correlations must be interpreted with caution, however, for a reason analogous to the one mentioned in the introduction on interpreting correlations between facial characteristic ratings and false alarm or hit rate. In that case, we noted that false alarm and hit rate are each two-dimensional entities composed of the independent components of d' and criterion. As we shall see shortly, and as Vokey and Read (1992) have found, typicality is also composed of two orthogonal components. Thus, if typicality and a facial characteristic correlate (do not correlate), the role of the memorability and familiarity/attractiveness components in establishing (can-

Table 3
Intercorrelation Matrices of the Ratings and Performance for the Caucasian Faces (Upper Triangle) and for the Japanese Faces (Lower Triangle)

		1	2	3	4	5	6	7
1	d'		.12	−.09	.10	.19	.00	−.18
2	Criterion	.44		−.52	−.12	−.11	.02	.01
3	Repetition	−.42	−.62		−.01	.01	−.03	.10
4	Familiarity	−.12	−.34	.39		.06	.28	.17
5	Memorability	.20	.06	−.16	−.11		.03	−.71
6	Attractiveness	−.14	−.21	.14	.50	−.23		.40
7	Typicality	−.38	−.34	.60	.29	−.57	.18	

Note—An r with absolute value greater than .22 is significant with $\alpha = .05$ and a nondirectional test.

celing) this relationship is not clear from the correlation alone. The situation is even more perilous in other cases where correlations between typicality, a two-dimensional entity, and false alarm or hit rate, also two-dimensional entities, are reported.

Finally, one additional point should be made. While the direction of the correlations reported here for the Caucasian faces are identical to those found in the literature, we found somewhat smaller correlations than those reported by other authors. We attribute this difference to the fact that our data may be somewhat noisier than that of other studies, because of the use of both male and female observers and male and female faces. Further, the exposure to other-race faces during the course of the rating experiment, even with order counterbalancing of the Japanese and Caucasian face rating tasks, may have added noise to the consistency of the application of rating criteria.

Caucasian analysis. For the Caucasian faces, four axes met our principal components analysis inclusion criterion. The loadings for these components and the percentage of variance explained by each appear in Table 4A. As noted, a consistent result found throughout this and all subsequent analyses, including those in which the neural network results are incorporated, was the clear identifiability of the axes as either rating- or performance-dominated. The rating-dominated axes were structurally very similar to those found by Vokey and Read (1992). Although there is no way to compare directly the axes that emerge from the present analysis with those of Vokey and Read (1992), we will proceed by making comparisons of how the structure of the loadings is similar and different in each case. On the axis appearing in column 1 of Table 4A, memorability and typicality loaded strongly in opposing directions. Following the terminology used by Vokey and Read (1992), we will call this axis *memorability*. One difference from Vokey and Read (1992) is that memorability and typicality have roughly equal loadings in our results, whereas memorability loaded more strongly than typicality in their data.

A unidirectional manifold of attractiveness, familiarity, and to a lesser degree, typicality, defined the second rating-dominated axis (see column 2 in Table 4A). Again, following the terminology of Vokey and Read (1992), we will call this axis *general familiarity*, although we stress that this axis might be labeled just as accurately as *attractiveness*. Two points are worth noting. First, consistent with Vokey and Read (1992), the magnitudes of the attractiveness and familiarity loadings were larger on this axis than that of the typicality loading. Second, in contrast to Vokey and Read (1992), typicality clearly did not have an equal foothold on the memorability and general familiarity axes. Typicality loaded more strongly on the memorability axis than on the familiarity axis. Nonetheless, typicality appeared consistently at a level above our chosen criterion on the general familiarity axis for Caucasian faces in all subsequent analyses. A comparison with the analysis of Japanese faces will further support the contention that typicality, despite its relatively small loading

Table 4A
Human Rating and Recognition Performance for Caucasian Faces

	First Four Rotated Factors			
	Memorability	Familiarity	Indictability	Accuracy
d'	.11	.04	−.09	**.95**
Criterion	−.11	−.08	**−.86**	.09
Repetition	−.09	−.06	**.87**	−.01
Familiarity	.01	**.74**	.12	.23
Memorability	**.94**	.14	.06	.05
Attractiveness	−.11	**.83**	−.10	−.15
Typicality	**−.88**	**.34**	.05	−.11
Proportion of variance accounted for by axis	.27	.20	.22	.20
Actual order of axis	1	3	2	3

Table 4B
Human Rating and Recognition Performance for Japanese Faces

	First Three Rotated Factors		
	Memorability	Familiarity	Indictability/Accuracy
d'	.23	.05	**.69**
Criterion	−.07	−.23	**.82**
Repetition	−.16	.18	**−.83**
Familiarity	−.01	**.83**	−.27
Memorability	**.93**	−.11	.00
Attractiveness	−.17	**.86**	.00
Typicality	**−.72**	.12	**−.51**
Proportion of variance accounted for by axis	.21	.22	.31
Actual order of axis	3	2	1

Note—Boldface is used for loadings greater than or equal to .30.

in comparison with that found on the memorability axis, contributes in a substantial way to the familiarity axis for Caucasian faces.

The performance-dominated axes appear in columns 3 and 4 of Table 4A. On the axis that appears in column 3, criterion loaded strongly in opposition to repetition on the first axis. A low value of criterion[5] indicates a loose or liberal criterion and hence goes with high judgments that a face appeared earlier in the series. We will call this axis *indictability*[6] since it would seem to indicate criterion variation or the tendency of a face to evoke a recognition response regardless of its status as a learned or new face in the recognition experiment. On the axis that appears in column 4 of Table 4A, d' appears alone. This axis might be interpreted, therefore, as an *accuracy* axis.

A final point is that the performance measures did not load substantially on the rating-dominated axes, nor did ratings load substantially on the performance-dominated axes. Had any of the ratings related strongly to variations in criterion or accuracy, we would expect to see cross-loadings between the performance and rating measures. In fact, we will see such loadings only in the analysis of the Japanese faces. We will discuss this in more detail after the Japanese analysis is presented.

Japanese analysis. For the Japanese faces, only three axes met our inclusion criterion. The loadings for the principal components and the percentage of variance explained by each component appear in Table 4B. As can be seen, the Japanese faces also showed clearly identifiable rating-dominated and performance-dominated axes. The rating-dominated axes were generally similar to those seen for the Caucasian faces and to those reported by Vokey and Read (1992), but with one important difference on the general familiarity axis. As for the Caucasian faces, the axis appearing in column 1 of Table 4B can be interpreted as *memorability*, with typicality opposing memorability. Contrary to the situation seen for Caucasian faces, however, the axis appearing in column 2 of Table 4B, while showing a positive manifold of familiarity and attractiveness, shows no substantive typicality loading. We will discuss this result shortly in the context of the performance-dominated axis.

The axis appearing in column 3 can be interpreted as a performance-dominated axis, but one that indicates that d' and criterion are not independent for Caucasian observers recognizing Japanese faces. Since criterion loads in the same direction as d', it suggests that there was a systematic loosening of criterion with less discriminable Japanese faces and a tightening of criterion with more discriminable faces. A second difference from the Caucasian data on this axis is the appearance of typicality on the accuracy axis opposing d'. This suggests, in contrast to the Caucasian data, that a component of rated typicality for the other-race Japanese faces was related to performance accuracy, but was unrelated to either general familiarity or memorability. There is no completely appropriate label for this axis. However, since the performance measures are strongly represented, and since the appearance of typicality is not unique to this axis, we label this axis *indictability/accuracy*.

Discussion

Several points are worth noting. First, little or no cross-loading occurred between ratings and performance variables for the Caucasian faces. By *cross-loading*, we mean simply the loading of performance variables on the rating-dominated axes and the loading of rating variables on the performance-dominated axes. The absence of cross-loadings between the performance and rating variables for the Caucasian faces could indicate that the ratings, while internally consistent, have little to do with recognition. This would be a very disappointing result. An alternative possibility might be presented simply as follows. For both the Caucasian and Japanese faces, the principal components analysis reveals a robust structure among the various rating measures, with several common points. For Caucasian faces, the principal components analysis allows us to see that the typicality rating is related independently to memorability and to the composite of attractiveness and familiarity. For the Japanese faces, by using the principal components analysis, we can see two additional structural aspects of the space: (1) the performance measures of d' and criterion are related structurally to each other and to typicality, and (2) the component of typicality related to familiarity for Caucasian faces is not present for the Japanese faces. The fact that there is a strong structure among the rating variables does not necessarily mean that there is no relationship between the ratings and performance measures, but rather may indicate that the within-rating structure is a good deal stronger for the Caucasian faces than is the structure between the performance measures and the ratings.

To assess more directly the degree to which ratings and performance measures are related, we used canonical correlation, which is a statistical technique used to assess the relationship between two sets of variables. It computes a linear combination within each set of variables so as to maximize the correlation between the two linear combinations of variables (Kshirsagar, 1972). The advantage of using canonical correlation is that the strength of the relationship can then be tested for statistical reliability. The results of this analysis on the raw data for rating and performance showed a reliable correlation between the ratings and performance measures both for the Caucasian (canonical correlation = .54, maximum likelihood ratio test, $p < .02$) and the Japanese (canonical correlation = .74, maximum likelihood ratio test, $p < .0001$) faces.[7] This indicates that, combined, the rating measures and the performance measures are related to one another in a way that explains roughly 29% of the variance for the Caucasian faces and 55% of the variance for the Japanese faces. The difference in the magnitude of this correlation for the Caucasian and Japanese faces indicates that the relationship between ratings and performance was stronger for the other-race Japanese faces than for the same-race

Caucasian faces. This is consistent with the fact that we see the presence of rating measures like typicality and familiarity (to a lesser extent) on the performance axes for the Japanese faces but not for the Caucasian faces.

In summary, the structural analysis indicates two important sets of differences between the same-race Caucasian faces and the other-race Japanese faces. First, for same-race Caucasian faces, the present analysis replicates Vokey and Read's (1992) finding that typicality is composed of two orthogonal components, one related inversely to memorability and the other related directly to attractiveness and familiarity. For Japanese faces, typicality was also composed of two components, one indeed related inversely to memorability. The other component, however, was related to accuracy and indictability. We will address the question of the role of typicality on the accuracy axis in the context of the simulations. In addition, whereas familiarity and attractiveness remained a structural component of the space, rated typicality for the other-race Japanese faces was not related to this component. Second, in general, the relationship between ratings and performance data was stronger for the other-race Japanese faces than for the same-race Caucasian faces. This is a somewhat surprising result, in that in Experiment 1 Caucasian observers were better able to recognize the Caucasian faces than the Japanese faces. This result suggests that expertise in face recognition, which is greater for same-race faces than for other-race faces, may be inversely related to the predictive value of facial characteristic ratings for recognition performance.

AUTOASSOCIATIVE MEMORY SIMULATIONS

One simple hypothesis of the other-race effect has been put forth recently by some of us (O'Toole, Deffenbacher, et al., 1991). We modeled the other-race effect as a problem in perceptual learning. By this account, exposure to the many faces of one race allows the perceptual system to make effective use of subtle variations in the form and configuration of the facial features of the race of faces learned. Unfortunately, other-race faces are not well characterized by these highly specialized and primed features, so we are less accurate at recognizing these faces. This account of the other-race effect is not unlike what is known about learning one's own native language. With a great deal of exposure to a single language, people become adept at processing the features of the language that are most useful for distinguishing between speech sounds in that language. This occurs at the cost of losing an ability to distinguish speech sounds that are important in other languages but not in one's own language.

To model a perceptual learning account of the other-race effect, we used a linear autoassociative neural network in conjunction with a low-level visual coding of Japanese and Caucasian faces. This network implements a principal components analysis of the *face images* on which it is trained. We simulated a biased "face history" by training a neural network to recognize a large number of faces of one race, a "majority" race, and a lesser number of faces of another race, a "minority" race. The model was then used to reconstruct faces by using the principal components derived from the learned set of faces. This reconstruction will be referred to, henceforth, as the *model's representation* of the face, since it is based not only on the face itself, but on the statistical properties of the set of faces learned by the model. The results of simulations using a biased face history showed three differences in the model's treatment of majority versus minority race faces: (1) The model was more proficient at representing novel faces from the majority race than from the minority race. In other words, the model's representation of a face was physically more similar to the *actual* face for majority race faces than for minority race faces. (2) The model's representations of novel faces from the minority race were more similar to one another than the codings with which it represented novel majority faces. (3) The model was better at recognizing majority faces than minority faces in an episodic memory task.[8] Recognition was defined for the model as its ability to discriminate learned from novel faces.

While there are no psychological data concerning Effect 1, Effects 2 and 3 can be related to anecdotal beliefs and empirical results concerning other-race faces. Effect 3 is simply the frequently reported recognition accuracy difference between same- and other-race faces. Effect 2, the fact that the model's codings of the minority race faces were more similar to one another than were its codings for the majority race faces, is reminiscent of the oft-noted feeling that other-race faces "all look alike." It is also consistent with Bruce's (1988) and Shepherd's (1981) suggestion that other-race faces are less recognizable because they are perceived as more similar to one another. They suggest that the higher interface similarity for other-race faces is due to the fact that the dimensions of the similarity space are determined mostly by same-race faces.

By itself, Effect 1 would appear to have little testable relevance for human processing of other race faces. However, since the model's representation quality measure was the basis of the interface similarity effect seen in the simulation, this measure can be related to some theoretical views of typicality (e.g., Light et al., 1979). Inasmuch as the model captured some qualitative aspects of the other-race effect, the purpose of the present simulations was to examine the relevance of this representation quality measure to human facial characteristic ratings and face recognition performance. To do this, we trained the associative memory with a majority of Caucasian faces and a minority of Japanese faces. This was meant to simulate a biased face history. Next, faces that were not learned by the model were "reconstructed" by the memory, and the representation quality measure for each face was incorporated directly into the principal components analyses for the human ratings and performance measures described in Experiment 2. This yielded one more measure on each of the faces, but a measure that was based on the model's ability to represent Caucasian versus Japanese faces when its face history was biased by training it with a majority of Caucasian faces. This permitted a view of the relationship between the Caucasian-biased autoassociative model's ability to represent faces and the ratings and recognition performance of Caucasian subjects.

SIMULATION 1

The purpose of this simulation was to examine the relevance of the model's quality of representation measure to human recognition performance and rating data.

Method

Stimuli. The stimuli consisted of the digitized version of a subset of the 240 faces from Experiments 1 and 2. Each digitized face image was 151 pixels wide and 225 pixels long. To teach a face image to the autoassociative network, the face must be represented as a vector of pixel intensities. We created a vector, $\mathbf{f}$, from each face image by concatenating the rows of the digitized image to produce a vector with 33,975 pixel elements. All face vectors were normalized (i.e., $\mathbf{f}_j^T\mathbf{f}_j = 1$ or the dot/inner product of the face with itself is one) prior to the simulations.

Apparatus. The simulations were performed on a Sun SparcStation and on a Convex C-1 Vector computer.

Procedure. The autoassociative memory can be described equivalently in neural network or in numerical/statistical analysis terminology. For the present purposes, we present both descriptions and show that they are equivalent. An excellent review of linear algebra, as it applies to autoassociators and related neural networks, can be found in Jordan (1986). This review covers all of the basic concepts and terminology that we employ in this section.

An autoassociative memory was created by using 75 Caucasian and 10 Japanese faces (approximately half male and half female in each race) by summing the outer product matrices for each learned face as follows:

$$\mathbf{A} = \sum_j \mathbf{f}_j\mathbf{f}_j^T, \tag{1}$$

where $\mathbf{f}_j$ is the jth face, and where T indicates the transpose operation. Thus, $\mathbf{A}$ can be thought of as a pixel by pixel matrix that represents a kind of composite memory of the faces. Each element a_{ij} of the matrix is a measure of the covariance between the ith and jth pixel of the face vectors across all of the learned faces. A simple neural network interpretation of this linear autoassociator is that each element a_{ij} of the matrix represents the connection strength between the ith and jth neurons. This is a distributed rather than a localized code for the faces, since the representation of any given face is not found in a localized part of the memory but is distributed across the connection strengths throughout the entire memory.

The next step of the process was to reconstruct or recall faces from the memory and to measure the quality of the reconstructions so that this quality measure could be incorporated into the Varimax analysis on the human rating and performance data.

First, using a neural network description, the jth face can be recalled from the memory matrix as follows:

$$\hat{\mathbf{f}}_j = \mathbf{A}\mathbf{f}_j, \tag{2}$$

where $\hat{\mathbf{f}}_j$ is the system estimate of the jth face, $\mathbf{f}_j$. A neural network interpretation of this equation can be seen as follows. Each element i of the retrieved jth face vector $\hat{\mathbf{f}}_j$ is interpretable as the output of the ith neuron. Since this neuron is connected to all other neurons, its output is simply the sum of its inputs (the face vector), weighted by the connection strengths between itself and all the neurons to which it connects. These connection strengths are established by the statistical properties of the *set* of learned faces, via the pixel covariance structure captured by the summed outerproducts in $\mathbf{A}$.

Second, to consider the reconstruction process in terms of a numerical or statistical analysis, it is important to realize that the matrix $\mathbf{A}$, like any square, symmetric matrix, can be expressed as a weighted sum of the outer products of its eigenvectors (e.g., Jackson, 1991):

$$\mathbf{A} = \sum_i \lambda_i \mathbf{e}_i\mathbf{e}_i^T, \tag{3}$$

where λ_i is the ith eigenvalue and $\mathbf{e}_i$ is the ith eigenvector. The eigen decomposition of the autoassociative matrix $\mathbf{A}$ is indicative of the statistical structure of the stimulus set used to construct it, and thus will depend on the kinds of faces stored in the memory. Viewed in terms of eigenvectors, retrieval of a face vector from this matrix can be illustrated by rewriting Equation 2 and substituting Equation 3 for $\mathbf{A}$ as follows for the jth face:

$$\hat{\mathbf{f}}_j = \lambda_1(\mathbf{f}_j^T\mathbf{e}_1)\mathbf{e}_1 + \lambda_2(\mathbf{f}_j^T\mathbf{e}_2)\mathbf{e}_2 + \ldots + \lambda_n(\mathbf{f}_j^T\mathbf{e}_n)\mathbf{e}_n, \tag{4}$$

where $\mathbf{e}_i$ indicates the ith eigenvector, and where $(\mathbf{f}_j^T\mathbf{e}_i)$ is the dot product between the jth face and the ith eigenvector. It is clear from this equation that recalling a particular face from the autoassociative matrix is equivalent to summing together a weighted combination of eigenvectors, where the weights are the eigenvalues multiplied by the dot products between the face vector and each eigenvector. Furthermore, since the spatial position of pixels is preserved by the autoassociative matrix, "retrieved" faces and their constituent eigenvectors can be displayed as images, providing a very useful analytical tool, as we shall illustrate shortly (cf. O'Toole & Thompson, 1993).

In the present simulation, we added the Widrow-Hoff error correction procedure (delta rule) to the training of the network. This is achieved in numerical analysis terms by dropping the eigenvalues from Equation 4, leaving the weights to be only the dot products between the face vectors and the eigenvectors. As noted previously, the eigenvalues are related to the proportion of variance explained by the associated eigenvectors, so dropping the eigenvalues serves to increase the relative importance of eigenvectors that explain smaller amounts of the variance. We shall consider the psychological relevance of this operation in detail in the introduction to Simulations 2A and 2B.

It is worth pointing out, as we have done in detail elsewhere (O'Toole, Abdi, Deffenbacher, & Valentin, 1993), that the autoassociative memory implements principal components analysis on the face images on which it is trained. In other words, the autoassociative model is carrying out the same kind of analysis on a physical coding of faces, (i.e., pixel images), that we have used in analyzing the rating and performance data in Experiment 2. The matrix decomposed into its eigenstructure in Experiment 2 contained the covariation among various performance and rating measures of faces, whereas the autoassociative matrix decomposed here contained the covariation among pixels in a set of face images. A second important difference between the analysis in Experiment 2 and the one carried out here is the use to which we put the autoassociative model of face images. Our end goal in this simulation was to use this model to produce a representation of each face, which we can evaluate in terms of its quality or goodness. This representation quality measure is defined simply as the cosine or normalized correlation between the original and recalled face vectors. This is a "goodness-of-fit" measure between the original and reconstructed face, and it is defined formally as:

$$\cos(\hat{\mathbf{f}},\mathbf{f}) = \frac{\hat{\mathbf{f}}^T\mathbf{f}}{\|\hat{\mathbf{f}}\|\ \|\mathbf{f}\|}, \tag{5}$$

where $\|\mathbf{f}\|$ is the length of the face vector, defined as the square root of the dot product of the vector with itself.

In sum, reconstructing faces from an autoassociative memory containing a majority of Caucasian faces and a minority of Japanese faces is like applying a perceptual filter (i.e., the autoassociative matrix) to the face images. The properties of the filter have been determined by the face history of the matrix. Intuitively, a face can be thought of as the sum of a set of global "features" (the eigenvectors), with different faces requiring different combinations of the features (for a discussion of eigenvectors as global features, see O'Toole & Abdi, 1989). We have referred previously to the particular set of weights needed to reconstruct a particular face as its "coefficient profile" (O'Toole, Abdi, Deffenbacher, & Bartlett, 1991) and will address the uses of the weights for face categorization in Simulations 2A and 2B.

A final theoretical and a procedural point should be made before describing the results of the simulation. The difference between reconstructing a face that the model has learned versus a face that the model has not learned is important, although perhaps not obvious. The retrieval equation(s) (equivalently, Equations 2 and 4), will produce an estimate of any input face vector (and indeed of any input vector of the appropriate dimensions), regardless of whether or not the model has learned (i.e., was trained on) that vector. We know from previous studies (e.g., O'Toole, Deffenbacher, et al., 1991) that the quality of the reconstruction, as measured by the cosine between the original and reconstructed face vectors, is generally better for learned than for novel faces. In this sense the model can be said to recognize faces. We have previously tested the model's ability to recognize faces with signal detection methodology, by setting a "criterion cosine," usually the ideal observer criterion (i.e., the mean of the mean cosines for the learned and new faces). The model responds "old" for faces with cosines higher than the criterion and "new" for faces with cosines less than the criterion. This procedure gives rise to the standard measures of signal detection theory (i.e., hits, false alarms, misses, and correct rejections), and so, d' can then be computed in the standard way.

In the present study, we reconstructed only unlearned faces because we wished to consider the autoassociative memory as a biased face *history*. We were interested more in the model's characteristics as perceptual filter for same- versus other-race faces than in its characteristics as a memory device for these faces. We have considered the properties of the autoassociator as a memory device for majority and minority race faces in detail in O'Toole, Deffenbacher, et al. (1991). As a final procedural point, since we wished to maximize the number of faces available for the principal components analysis, and, since we wished these faces to be only "unlearned" faces (i.e., those not used in creating the face history matrix **A**), we actually created two face histories with nonoverlapping sets of learned faces. This increased the number of faces available for the analysis. Hence, the eigenvectors were then extracted from each matrix, and reconstructions for each face were made from the set of eigenvectors extracted from the matrix that did not learn the particular face.

Results and Discussion

The quality of the model's reconstruction for each of 160 unlearned faces (gathered from the two face history matrices) was calculated as the cosine between the original face vector and the vector reconstructed by the model using Equation 2. A cosine of 1.0 indicates perfect reconstruction: The higher the cosine, the more similar the reconstruction to the original face.

Caucasian analysis. The cosine measure was incorporated into the principal components analyses of Experiment 2 for Caucasian and Japanese faces separately. The loadings on the first four axes appear in Tables 5A and 5B. As can be seen, the structure of the space for variables other than the cosine remains generally similar to that previously discussed. Several points are worth noting about the presence of the cosine in the space. First, beginning with the performance-dominated axes, the cosine loaded strongly on the accuracy axis in the direction opposite to d', but did not load on the indictability axis. The direction of the loading of the cosine on the accuracy axis in opposition to d' makes sense because we expect the reconstructions to be best for unlearned faces that share the most in common with the set of faces learned—in other words, those most similar to the average face. For our observers, however, conventional wisdom would indicate that faces least similar to the average face should be recognized best. The fact that the cosine loads more strongly on this axis than did any of the rating measures indicates that the autoassociative memory model's quality of representation measure is more strongly related to human recognition performance than any of the human rating measures are.

For the rating-dominated axes, the cosine loaded moderately on the memorability axis, but not on the familiarity axis to a degree that met our criterion of .30. This indicates that the cosine is capturing aspects of the human rating data related to memorability. However, the loading of the cosine on this axis is in a counterintuitive direction in comparison with what is seen on the accuracy axis. On the accuracy axis, the cosine opposes d'. On the memorability axis, however, the cosine loads in the same direction as do memorability and d', and opposing typicality. Hence, the orthogonal components of the cosine appear paradoxical.

Since we found this result puzzling, we examined the contributions of each face to the memorability axis. The contribution of a face to a particular axis is simply the standardized projection (z score of the projection) of the face on the axis. Specifically, we looked at faces whose standardized projection scores on the memorability axis were large. We found that the faces contributing strongly and positively to this axis had a relatively small, highly distinctive feature (a distinctive feature taking up a relatively small amount of the face image). For example, the largest contributor to this axis was a male face with a very unusual Elvis-like mouth shape. When the mouth was covered, the face appeared relatively typical. Other examples include a female face with a thick strand of hair hanging down on her forehead and a male face with an unusually shaped, small bushy eyebrow. Although these faces would appear to be very memorable to the observers and hence would generally yield high d's, the model's behavior on these faces is quite understandable. Since these faces were not learned by the model, and since their highly memorable aspects were confined to a relatively small area of the face, the model would still do a reasonably good job of reconstructing the faces (i.e., a high cosine). Thus, the model gives insight into the component of typicality related to memorability. It suggests that this component may well be a function of relatively local, small distinctive features, rather than due to unusual configural properties of the faces.

Additionally, and interestingly, d' appears on the memorability axis at about the same strength and in the same direction as does the cosine, whereas previously it had failed to load. The preceding discussion makes it clear

Table 5A
Simulation 1 with Human Rating and Recognition Performance for Caucasian Faces: Autoassociative Model Trained with a Majority of Caucasian and a Minority of Japanese Faces

	First Four Rotated Factors			
	Memorability	Familiarity	Indictability	Accuracy
d'	**.33**	.18	−.17	**.74**
Criterion	−.08	−.07	**−.87**	.01
Repetition	−.06	−.03	**.85**	−.03
Familiarity	.06	**.78**	.09	.06
Memorability	**.90**	.11	.07	.01
Attractiveness	−.18	**.78**	−.07	−.08
Typicality	**−.88**	**.34**	.05	.00
Cosine	**.30**	.17	−.09	**−.79**
Proportion of variance accounted for by axis	.23	.18	.20	.15
Actual order of axis	1	3	2	4

Table 5B
Simulation 1 with Human Rating and Recognition Performance for Japanese Faces: Autoassociative Model Trained with a Majority of Caucasian and a Minority of Japanese Faces

	First Three Rotated Factors		
	Memorability	Familiarity	Indictability/Accuracy
d'	.16	.02	**.69**
Criterion	−.12	−.28	**.76**
Repetition	−.08	.21	**−.82**
Familiarity	−.03	**.84**	−.22
Memorability	**.86**	−.06	.17
Attractiveness	−.13	**.84**	−.03
Typicality	**−.62**	.11	**−.62**
Cosine	**.50**	−.11	**−.41**
Proportion of variance accounted for by axis	.18	.19	.29
Actual order of axis	3	2	1

Note—Boldface is used for loadings greater than or equal to .30.

that the memorability axis is due, at least in part, to the presence of some faces that are made memorable by the presence of a highly distinctive local feature. These faces are generally well reconstructed by the model, but not in a way that would preserve the highly distinctive feature. These faces create a moderate positive relationship between d' and the cosine, which expresses itself as a moderate loading on the memorability axis. We will see in the second set of simulations a reversal of the positive relationship between d' and the cosine on the memorability axis when the kind of information that the cosine is tapping is shifted.

While the cosine did not load on the general familiarity component at a level that met our criterion, the fact that faces that were judged "memorable" had a small local distinctive feature led us to examine a few analogous faces for the familiarity axis. Faces that were "atypical" for this axis would also be judged generally unattractive and unfamiliar. A few examples of faces fitting this combination of judgments were two female faces with small heads and long skinny necks that lent them a rather ostrich-like appearance and a male face with a very round-shaped hair style. Thus, it would seem that whereas small local distinctive features made faces atypical with respect to the memorability axis, global features such as face and hair shape made faces atypical with respect to the familiarity axis.

Japanese analysis. The qualitative aspects of the Japanese analysis are generally similar to those seen in the Caucasian analysis. The loadings for this analysis appear in Table 5B. Beginning again with the performance-dominated axis, the cosine loads moderately strongly and in an opposing direction to d' and criterion, but in a similar direction to typicality on the accuracy/indictability axis. The relationship between cosine and d' is analogous to that seen with the Caucasian faces—good model reconstructions occurred for new faces that were most similar to the average face. By contrast, good recognition performance for our observers occurred with faces that the model indicated as least similar to the average face. In addition, the presence of typicality on this axis indicates that the cosine is related to the part of typicality that is orthogonal to the memorability component of typicality for these other-race Japanese faces.

For the rating subspace, the cosine appears strongly on the memorability axis, in the same direction as memorability, but opposing typicality—as was seen for the Caucasian faces. We looked at the Japanese faces that contributed strongly to this axis. They too, contained a small distinctive feature. The three largest contributors were faces with unusually protruding ears.

Discussion

In general, the representation quality measure produced by the autoassociative memory was relevant to both human recognition performance accuracy and human facial characteristic ratings. In fact, for Caucasian faces, the autoassociative model's representation quality measure was related more to human recognition performance than were any of the human facial characteristic ratings. We think that the relevance of the autoassociative memory to performance accuracy indicates that human face recognition performance may rely heavily on very subtle variations in the form and configuration of visual information. Since the model makes use of all of the faces learned in creating its representation, these kinds of useful subtle variations are captured well in the model's representation of faces. While facial characteristic ratings attempt to quantify the same kind of information, they fall short of tapping the complicated image-based information that well-practiced human observers may be using for face recognition.

The hypothesis of greater reliance on subtle variations in the form and configuration of visual information for same- versus other-race recognition is consistent with the finding in Experiment 2 that the ratings for other-race faces were more related to performance than were the ratings for same-race faces. Specifically, these kinds of subtle visual features may not be conducive to expression by facial characteristic ratings and hence ratings may be less able to tap recognition processes for same- as opposed to other-race faces.

The model's quality of representation measure was also relevant to the memorability component of typicality. The combination of its counter-intuitive loading direction and an examination of the faces that strongly contributed to this axis suggested that the memorability-related component of typicality is likely to be related to highly distinctive local features. The final set of simulations lends more insight into the kinds of information that may give rise to the different components of typicality.

SIMULATIONS 2A AND 2B

In recent work, we have examined the kinds of information that different eigenvectors provide in reconstructing a face. O'Toole, Abdi, et al. (1991) have shown that the race and sex of a face can be predicted with relatively good accuracy (>80%) by looking at the projection of the face on a single eigenvector. Depending on the racial and sexual homogeneity of the training set, the eigenvector with the greatest predictive value for these categories is usually the eigenvector with the second or third largest eigenvalue.[9] A demonstration of the importance of the second eigenvector in determining some aspects of the opposition between male and female facial characteristics appears in O'Toole et al. (1993). We show that when the second eigenvector is added to the first, the resultant face appears masculine, whereas when the second eigenvector is subtracted from the first, the resultant face appears feminine. Analogously, this type of demonstration can be carried out for the autoassociative memory when it is trained with two races of faces. For example, when an autoassociative memory is trained with half Caucasian and half Japanese faces, good information about the race of the face is found in the second eigenvector. This is illustrated in Figure 1. The first and second eigenvectors of an autoassociative memory trained with 40 Caucasian and 40 Japanese faces appear in the first and second panels, respectively, of Figure 1. The first eigenvector can be interpreted as representing the common face-like aspects of the entire set of images. The second eigenvector appears monster-like, and would, by itself, seem to contain little information about the race of a face. However, when the second eigenvector is added to the first eigenvector (panel 3), the resulting face has a clearly Asian appear-

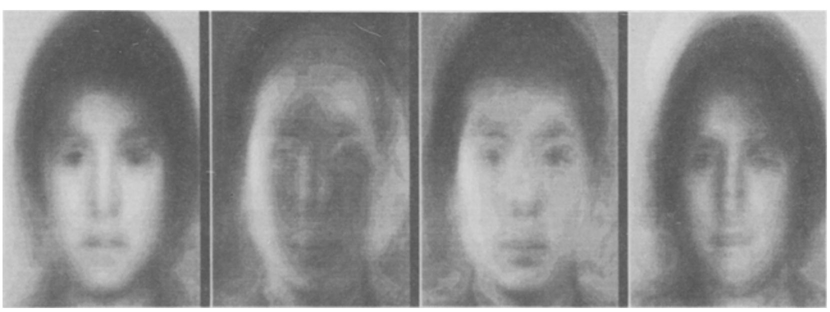

Figure 1. The first and second eigenvectors of an autoassociative memory trained with 40 Caucasian and 40 Japanese faces appear in the first and second panels, respectively. The results of adding the second eigenvector to the first eigenvector appear in panel 3 of Figure 1, whereas panel 4 illustrates the results of subtracting the second eigenvector from the first. The face in panel 3 appears Japanese, whereas the face in panel 4 appears Caucasian.

ance. When the second eigenvector is subtracted from the first (panel 4), the resulting face has a Caucasian appearance. We have argued elsewhere (O'Toole, Abdi, et al., 1991; O'Toole et al., 1993) that the eigenvectors with relatively large eigenvalues contain what Bruce and Young (1986) have referred to as visually derived semantic information about faces. Visually derived semantic information can be extracted from unfamiliar faces and includes visually based categorical information such as the race, sex, and age of a face.

In addition to representing visually derived semantic information, to be able to recognize faces (i.e., discriminate learned from unlearned faces), the autoassociative memory must preserve information that is specific to individual faces. Sirovich and Kirby (1987) and Kirby and Sirovich (1990) have shown previously that only a subset of eigenvectors, those with the largest eigenvalues, is necessary for reconstructing a face to recognizable levels. They proposed that faces could be represented in a low-dimensional space in a way that minimizes the error of reconstruction. Although this minimizes the error of the reconstructions, our recent work (O'Toole et al., 1993) indicates that the model's recognition performance is actually better with ranges of eigenvectors that have relatively small eigenvalues. We have simulated the importance of different ranges of eigenvectors for recognition purposes. The model's d' for the discrimination of learned from unlearned faces was at its optimum with ranges of eigenvectors whose eigenvalues were relatively small. This makes intuitive sense in that the principal components with the largest eigenvalues are those that capture the largest proportion of variance in the stimulus set. As noted, these larger components of variance are likely to be related to large-scale visual categorical dimensions of the faces such as sex and race, which can be ascertained frequently by using primarily the global shape information in faces. Information that is specific to only a single or to a few faces will be captured in eigenvectors with smaller eigenvalues. These eigenvectors should provide better information for discriminating learned from unlearned faces. We will call the information contained in the eigenvectors with smaller eigenvalues *identity-specific* information.

The final set of simulations was carried out to contrast the model's quality of representation measure based on eigenvectors with larger versus smaller eigenvalues. Further, we examined how this measure would relate to human performance when the range of eigenvectors used to reconstruct a face was shifted. It may be expected that the results of Simulation 1 will be replicated in the range of eigenvectors with larger eigenvalues. This is because the eigenvectors with larger eigenvalues will dominate via their larger weights when all of the eigenvectors are used in reconstructing the faces. Recall that these weights are defined as the dot products of the faces with the eigenvectors, which will be larger generally, for eigenvectors that explain larger proportions of the variance in the face set. For the range of eigenvectors with smaller eigenvalues, we would expect that identity-specific information would dominate. If the autoassociative model is getting at aspects of identity-specific information relevant to human rating and recognition performance, we would expect the cosine to load in this principal components analysis, as well.

Method

The method was identical to that described for Simulation 1 with the following exception. For Simulation 2A, Equation 4 contained only the terms for Eigenvectors 1-35, (i.e., the eigenvectors with the 35 largest eigenvalues). For Simulation 2B, Equation 4 contained only the terms for Eigenvectors 20-55. The particular ranges used contain the same number of eigenvectors, but are somewhat arbitrary. The results, however, were robust over several shifts in this range. Sample face reconstructions in these two ranges are illustrated in Figure 2. The leftmost panel shows two original faces. The center panel shows each face reconstructed with the first 35 eigenvectors, whereas the rightmost panel shows the two faces reconstructed with the 20th through 55th eigenvectors. As can be seen, shape information is well preserved in the faces reconstructed with the first 35 eigenvectors. Yet while this representation clearly provides shape information appropriate to these particular faces, much additional detailed information is provided in the faces reconstructed with the 20th through 55th eigenvectors.

Results

Simulation 2A. As in Simulation 1, separate principal components analyses were carried out on the Caucasian and Japanese faces using the human data from Experiments 1 and 2. The cosines from the faces reconstructed in the two eigenvector ranges were incorporated into separate sets of principal components analyses. The loadings for the principal components analyses of Simulation 2A for Caucasian and Japanese faces appear in Tables 6A and 6B. As predicted, the pattern of results is very similar to that seen in Simulation 1 for the Caucasian faces.

Perhaps the only appreciable difference in these results in comparison with those of Simulation 1 occurs for the Japanese faces. On the accuracy axis, the loading of the cosine is reduced considerably. However, the direction of the loading of both the cosine and typicality are the same and are in opposition to d'. We will discuss this result in conjunction with the results of Simulation 2B.

Simulation 2B. The loadings for the principal components analysis of Simulation 2B for Caucasian and Japanese faces appear in Tables 7A and 7B. We will start with the Caucasian faces (see Table 7A). Two interesting differences were observed between this simulation and Simulations 1 and 2A. First, while the cosine again loaded very strongly on the accuracy axis, the direction of the loading with respect to d' in this simulation was reversed in comparison with that seen previously. This is to be expected if the eigenvectors with smaller eigenvalues are tapping identity-specific information. The higher the quality of the representation of identity-specific information, the higher we would expect human recognition performance. This contrasts to the range of eigenvectors with larger eigenvalues. In that case, the higher the quality of representation, the more generally similar the faces were

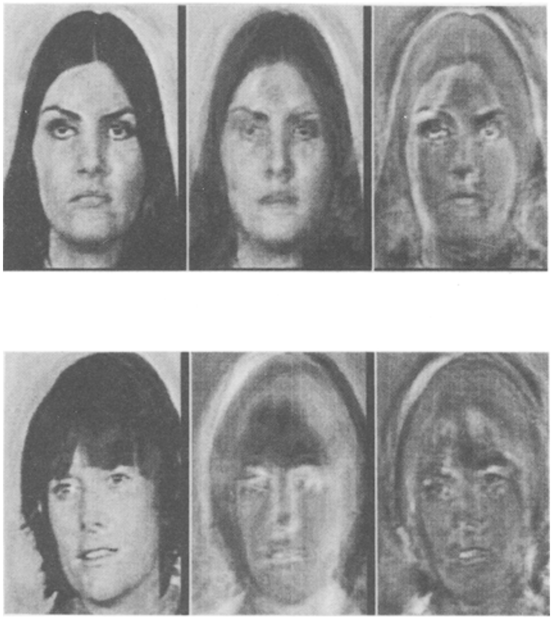

Figure 2. The leftmost panel shows two original faces. The center panel shows each face reconstructed with the first 35 eigenvectors. Global shape information is well preserved in these reconstructions. The rightmost panel shows the two faces reconstructed with the 20th through 55th eigenvectors. Small local details are well preserved in these reconstructions.

to all learned faces, and hence, the worse the human recognition performance.

A second difference was that while the cosine again loaded on the memorability axis, the direction was likewise reversed from the pattern observed in Simulations 1 and 2A. As noted, Simulation 1 indicated that faces contributing strongly to this axis were characterized by the presence of a small distinctive feature. When the eigenvectors used to reconstruct faces contained the information/facial characteristics most common to all faces (eigenvectors with large eigenvalues), these memorable faces were generally well reconstructed—that is, only a small part of the face was badly reconstructed. This is because what was unusual or "memorable" about each of these faces was confined to a very small area of the face. Within a range of eigenvectors with smaller eigenvalues, features specific to a single or small subset of learned faces are well represented, but those more common to the entire set of faces are lost. However, the information necessary to capture the identity-specific distinctive features of the "memorable" faces was not likely to be contained in these particular eigenvectors, since the faces reconstructed in this simulation were not learned by the model. Given that the more categorical information about all faces was also not present in this range of eigenvectors, a highly memorable (unlearned) face that was memorable because of a small distinctive feature would be badly reconstructed relative to the other faces. Hence, faces rated by human ob-

Table 6A
Simulation 2A with Human Rating and Recognition Performance for Caucasian Faces: Eigenvectors 1-35

	First Four Rotated Factors			
	Memorability	Familiarity	Indictability	Accuracy
d'	**.33**	.17	−.17	**.76**
Criterion	−.08	−.07	**−.87**	.00
Repetition	−.05	−.03	**.85**	−.02
Familiarity	.07	**.79**	.08	.04
Memorability	**.90**	.11	.08	.00
Attractiveness	−.19	**.77**	−.06	−.05
Typicality	**−.89**	**.34**	.05	.00
Cosine	**.33**	.19	−.12	**−.77**
Proportion of variance accounted for by axis	.23	.18	.20	.15
Actual order of axis	1	3	2	4

Table 6B
Simulation 2A with Human Rating and Recognition Performance for Japanese Faces: Eigenvectors 1-35

	First Three Rotated Factors		
	Memorability	Familiarity	Indictability/Accuracy
d'	.07	.03	**.71**
Criterion	−.15	−.22	**.77**
Repetition	−.02	.17	**−.85**
Familiarity	−.07	**.79**	−.21
Memorability	**.79**	−.18	.22
Attractiveness	−.01	**.87**	−.03
Typicality	**−.54**	.17	**−.65**
Cosine	**.66**	.09	−.18
Proportion of variance accounted for by axis	.17	.19	.30
Actual order of axis	3	2	1

Note—Boldface is used for loadings greater than or equal to .30.

Table 7A
Simulation 2B with Human Rating and Recognition Performance for Caucasian Faces: Eigenvectors 20–55

	First Four Rotated Factors			
	Memorability	Familiarity	Indictability	Accuracy
d'	**.39**	.20	−.21	**.64**
Criterion	−.06	−.05	**−.89**	−.11
Repetition	−.03	−.01	**.83**	−.13
Familiarity	.07	**.78**	.09	.06
Memorability	**.90**	.10	.08	−.06
Attractiveness	−.19	**.78**	−.06	−.08
Typicality	**−.87**	**.35**	.04	.00
Cosine	**−.32**	−.15	.11	**.81**
Proportion of variance accounted for by axis	.23	.18	.20	.14
Actual order of axis	1	3	2	4

Table 7B
Simulation 2B with Human Rating and Recognition Performance for Japanese Faces: Eigenvectors 20–55

	First Three Rotated Factors		
	Memorability	Familiarity	Indictability/Accuracy
d'	.05	.04	**.73**
Criterion	−.11	−.27	**.78**
Repetition	−.15	.19	**−.81**
Familiarity	−.07	**.82**	−.23
Memorability	**.80**	−.02	.19
Attractiveness	−.13	**.82**	−.03
Typicality	**−.67**	.09	**−.59**
Cosine	**−.60**	.14	.19
Proportion of variance accounted for by axis	.18	.19	.29
Actual order of axis	3	2	1

Note—Boldface is used for loadings greater than or equal to .30.

servers to be highly memorable due to the presence of a small, local, distinctive feature would give rise to a low cosine when the model tried to reconstruct the face with identity-specific information that it had not learned. It is worth noting also that in this range of eigenvectors, the positive relationship between d' and the cosine on the memorability axis that was seen in the first two simulations was also reversed for analogous reasons.

A generally similar pattern of reversals can be seen for the Japanese faces. Whereas the cosine loaded in opposition to d' in Simulations 1 and 2A, it loaded in the same direction in this simulation. Likewise, where the cosine loaded with memorability and against typicality in Simulations 1 and 2A, it opposed memorability and loaded with typicality in this simulation.

GENERAL DISCUSSION

The present study was undertaken to look for qualitative differences in the processing of same- versus other-race faces. Some differences were evident in the human performance data that were clarified through comparison with the performance of the neural network model. The structure of ratings for faces replicated Vokey and Read's (1992) claim that typicality is composed of orthogonal components related to memorability and familiarity for same-race Caucasian faces. This result is impressively robust given that Vokey and Read (1992) used faces of 10th and 12th graders, whereas we used young adult faces. For the other-race Japanese faces, we found a similar rating subspace with the exception that typicality did not load on the familiarity axis, but rather loaded on the accuracy axis.

An examination of faces that were important determiners of the memorability axis, in conjunction with the loading of the cosine in a surprising direction on this axis, gave insight into the differences in typicality on the familiarity and memorability axes. Faces that contributed strongly to the memorability axis were characterized by the presence of a small local distinctive feature. This also appeared to be the case for the Japanese faces.

Since the cosine did not appear on the familiarity axis for any of the present simulations at a level that met our criterion, the model gives less direct insight about its role in typicality and recognition performance for same- versus other-race face processing. However, some speculative and suggestive information can be gleaned from the data as a whole. First, we observed informally that Caucasian faces fitting an atypical profile with respect to this axis (i.e., atypical, unattractive, and unfamiliar) generally deviated from the set of faces in terms of global face and head shapes. Second, the contrast between the appearance of typicality on the familiarity axis for the Caucasian faces and its absence for the Japanese faces, in conjunction with the appearance of typicality on the accuracy/indictability axis for Japanese faces and its absence for the Caucasian faces, suggests the possibility that the kind of information tapped by "typicality" in these two contexts may be related. If this kind of information is related to the global-shape-based properties of faces, as the informal observation of Caucasian faces suggests, it would appear that the effects of the global structural properties of faces may be more related to performance for the Japanese faces than for the Caucasian faces. In other words, it is possible that the global or structurally based typicality status of a face perturbs the performance of Caucasian observers to a greater extent for Japanese faces than for Caucasian faces. This is reasonable, since Caucasian observers have greater expertise with Caucasian faces and hence should be better able to code small variations in facial form for these faces than for Japanese faces. In any case, more data will be required for this hypothesis to be tested explicitly.

Another important finding of this paper was that the cosine measure was more related to human recognition accuracy for Caucasian faces than were any of the observer face ratings. We believe that this result stresses the importance of considering faces as images that provide observers with very rich and elaborate information that is quite difficult to capture in discrete ratings. In the present study, we used face stimuli that were rich in perceptual information, and we have shown that even a relatively simple neural network model with an extremely simple visual code was able to capture more of this information than observer ratings were. An interesting point is that the kind of information that the model captures is relevant to what researchers mean when they ask observers to rate faces for typicality—"How like all other faces is this face?" This is a large part of what the cosine measure captures effectively. Observer ratings, however, appear less able to tap this information, or may simply be highly variable in terms of the strategies that observers use. It may be that the orthogonal components of typicality that we see are due to "typicality's" being a single entity composed of two parts. Alternatively, the separate components could be due to some observers' using a strategy of rating any face with a small distinctive feature "atypical," and other observers' rating any face that deviates in global shape or structure as "atypical." Hence, the components that Vokey and Read (1992) and we have found could be based on strategy differences. More likely, there is a tradeoff between the two strategies within a single observer across faces.

As noted by Valentine and Endo (1992), the simulation approaches used in O'Toole, Deffenbacher, et al. (1991)—and the one we apply here—are broadly similar to their view that the mental representations of faces can be thought of metaphorically as points in a multidimensional space. The present approach is different from their multidimensional model, primarily in that Valentine and Endo do not specify the number and nature of the dimensions of this space and conceptualize these dimensions as abstract parameters derived from faces. In the present study, the dimensions of the autoassociative network emerge spontaneously from the statistical structure of the set of face images learned by the autoassociative model. Hence, the present model, as it captures some aspects of human facial characteristic judgments and recognition performance, may lend insight into the nature of the dimensions that a multidimensional representation implies.

Finally, the existence and form of the accuracy and indictability axes, especially for the Japanese faces, in which typicality, d', and criterion are all interrelated, serve to remind us of the mechanics of a face recognition task and suggest that there are potential problems in interpreting correlations between ratings and single performance measures derived from signal detection theory.

REFERENCES

Bartlett, J. C., Hurry, S., & Thorley, W. (1984). Typicality and familiarity of faces. *Memory & Cognition*, **12**, 219-228.

Bothwell, R. K., Brigham, J. C., & Malpass, R. S. (1989). Cross-racial identification. *Personality & Social Psychology Bulletin*, **15**, 19-25.

Bruce, V. (1988). *Recognizing faces*. Hillsdale, NJ: Erlbaum.

Bruce, V., & Young, A. W. (1986). Understanding face recognition. *British Journal of Psychology*, **77**, 305-327.

Cohen, M. E., & Carr, W. J. (1975). Facial recognition and the von Restorff effect. *Bulletin of the Psychonomic Society*, **6**, 383-384.

Deffenbacher, K. A., & Loftus, E. F. (1982). Do jurors share a common understanding concerning eyewitness behavior? *Law & Human Behavior*, **6**, 15-30.

Feingold, C. A. (1914). The influence of environment on identification of persons and things. *Journal of Criminal Law & Police Science*, **5**, 39-51.

Going, M., & Read, J. D. (1974). The effect of uniqueness, sex of subject and sex of photograph on facial recognition. *Perceptual & Motor Skills*, **39**, 109-110.

Jackson, J. E. (1991). *A user's guide to principal components analysis*. New York: Wiley.

Jordan, M. I. (1986). An introduction to linear algebra in parallel distributed processing. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing: Vol. 1. Foundations* (pp. 365-421). Cambridge, MA: MIT Press.

Kirby, M., & Sirovich, L. (1990). Application of the Karhunen-Loève procedure for the characterization of human faces. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, **12**, 103-108.

Kshirsagar, A. M. (1972). *Multivariate analysis*. New York: Marcel Dekker.

Light, L. L., Kayra-Stuart, F., & Hollander, S. (1979). Recognition memory for typical and unusual faces. *Journal of Experimental Psychology: Human Learning & Memory*, **5**, 212-228.

O'Toole, A. J., & Abdi, H. (1989). Connectionist approaches to visually based feature extraction. In G. Tiberghien (Ed.), *Advances in cognitive psychology* (Vol. 2, pp. 1-13). London: Wiley.

O'Toole, A. J., Abdi, H., Deffenbacher, K. A., & Bartlett, J. C. (1991). Classifying faces by race and sex using an autoassociative memory trained for recognition. In K. J. Hammond & D. Gentner (Eds.), *Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society* (pp. 847-851). Hillsdale, NJ: Erlbaum.

O'Toole, A. J., Abdi, H., Deffenbacher, K. A., & Valentin, D. (1993). Low-dimensional representation of faces in higher dimensions of the face space. *Journal of the Optical Society of America A*, **10**, 405-410.

O'Toole, A. J., Deffenbacher, K. A., Abdi, H., & Bartlett, J. C. (1991). Simulating the "other-race" effect as a problem in perceptual learning. *Connection Science: Journal of Neural Computing, Artificial Intelligence, & Cognitive Research*, **3**, 163-178.

O'Toole, A. J., & Thompson, J. L. (1993). An X Windows tool for synthesizing face images from eigenvectors. *Behavior Research Methods, Instruments, & Computers*, **25**, 41-47.

Shapiro, P. N., & Penrod, S. D. (1986). Meta-analysis of face identification studies. *Psychological Bulletin*, **100**, 139-156.

Shepherd, J. W. (1981). Social factors in face recognition. In G. M. Davies, H. D. Ellis, & J. W. Shepherd (Eds.), *Perceiving and remembering faces* (pp. 55-79). London: Academic Press.

Sirovich, L., & Kirby, M. (1987). Low-dimensional procedure for the characterization of human faces. *Journal of the Optical Society of America A*, **3**, 519-524.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, **117**, 34-50.

Valentine, T., & Bruce, V. (1986). Recognizing familiar faces: The role of distinctiveness and familiarity. *Canadian Journal of Psychology*, **40**, 300-305.

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.

Valentine, T., & Ferrara, A. (1991). Typicality in categorization, recognition and identification: Evidence from face recognition. *British Journal of Psychology*, **82**, 87-102.

Vokey, J. R., & Read, J. D. (1988). Typicality, familiarity and the recognition of male and female faces. *Canadian Journal of Psychology*, **42**, 489-495.

Vokey, J. R., & Read, J. D. (1992). Familiarity, memorability, and the effect of typicality on the recognition of faces. *Memory & Cognition*, **20**, 291-302.

NOTES

1. Although Vokey and Read (1992) label this axis "general familiarity," they note that it is a manifold of attractiveness, typicality, familiarity, and likableness, with attractiveness and likability more strongly weighted than either familiarity or typicality. For consistency, we will use their label of general familiarity for this manifold, regardless of the actual order of loadings.

2. The proportion of variance explained by a particular axis is equal to the eigenvalue associated with the axis divided by the sum of the eigenvalues for all axes.

3. All axes are all those with nonzero eigenvalues.

4. Since all analyses of the human rating data used *Varimax-rotated* principal components analysis, for convenience, henceforth we shall refer to these analyses simply as principal components analysis.

5. We used C, a measure of the displacement in z-score units of the criterion, computed as $-0.5(z_H + z_{FA})$. With this measure, smaller values imply looser criteria (Snodgrass & Corwin, 1988).

6. Our use of the coined word *indictability* is inspired by the eyewitness memory literature and is meant to indicate the tendency of a face to be "recognized" or judged as "familiar" regardless of whether or not it is *actually* familiar. Hence, we try to capture, with our use of this word, the likelihood that a particular face will be falsely recognized—or by analogy to the problem for eyewitness identification—falsely "indicted." Although we realize that this label may not be a perfect description for the axis, the alternative labels, such as "criterion" and "response bias," seemed even less appropriate. "Criterion" seemed problematic because it would be easy to confuse with the criterion variable, while "response bias" seemed problematic in that it is a property of the perceiver and not of the face.

7. Hit and false alarm rates are included as performance measures in this analysis. We have done the principal components analysis both with and without hits and false alarms. We have reported the principal components analyses without them for two reasons: (1) without them, the resulting principal components analysis structures are simpler; and (2) the simpler structures do not change any important conclusions to be drawn from the principal components analysis.

The analysis reported defines repetition as a performance measure, since its status as a performance or rating measure is not entirely clear. When repetition is defined as a rating measure in the canonical correlation analysis, the correlation for Caucasian faces is increased substantially, though, we think, somewhat artificially. Furthermore, with repetition defined as a rating, the analysis still shows a larger correlation for the Japanese faces (.73) than for the Caucasian faces (.64).

8. Both Japanese and Caucasian faces served alternately as the majority and minority race faces for tests yielding Conclusions (1) and (2). Due to a shortage of Japanese faces for the episodic recognition task, Conclusion (3) was tested with only Caucasian faces as the majority race faces.

9. The second eigenvector is related to the sex of the face when the autoassociative matrix is trained with a relatively homogeneous set of faces of a single race. When we have trained the autoassociative model with Caucasian and Japanese faces, the second eigenvector is related to the race of the face and information about the sex of the face moves to the third and fourth eigenvectors (O'Toole, Abdi, et al., 1991).

(Manuscript received August 10, 1992;
revision accepted for publication June 25, 1993.)