Identification of own-race and other-race faces: Implications for the representation of race in face space

GRAHAM BYATT and GILLIAN RHODES
University of Western Australia, Nedlands, Western Australia, Australia

Own-race faces are recognized more easily than faces of a different, unfamiliar race. According to the multidimensional space (MDS) framework, the poor discriminability of other-race faces is due to their being more densely clustered in face space than own-race faces. Multidimensional scaling analyses of similarity ratings (Caucasian participants, $n = 22$) showed that other-race (Chinese) faces are more densely clustered in face space. We applied a formal model to test whether the spatial location of face stimuli could account for identification accuracy of another group of Caucasian participants ($n = 30$). As expected, own-race (Caucasian) faces were identified more accurately (higher hit rate, lower false alarms, and higher A') than other-race faces, which were more densely clustered than own-race faces. A quantitative model successfully predicted identification performance from the spatial locations of the stimuli. The results are discussed in relation to the standard MDS account of race effects and also an alternative "race-feature" hypothesis.

It is well documented that own-race faces are recognized more easily than faces of a different, less familiar race. This "own-race bias" is found across various tasks and is not due to differences in stimulus set or observer characteristics (for reviews and meta-analyses, see Bothwell, Brigham, & Malpass, 1989; Meissner & Brigham, 2001). Furthermore, faces from different racial groups have similar physiognomic variance (Goldstein 1979a, 1979b), so the own-race bias is not due to faces of one racial group being inherently more difficult to recognize than faces from another. Thus, it seems likely that the own-race bias is a result of differences in the mental representation of own- and other-race faces.

The multidimensional space (MDS) framework is currently a popular account of how faces are represented (e.g., Bruce, Burton, & Dench, 1994; Byatt & Rhodes, 1998; Johnston, Kanazawa, Kato, & Oda, 1997; Johnston, Milne, & Williams, 1997; Rhodes, Byatt, Tremewan, & Kennedy, 1997; Valentine & Endo, 1992). The core assumption of the framework is that the spatial relationships between representations in a multidimensional feature space (face space) can explain a variety of face perception phenomena. The concept of exemplar density, which captures the idea that some faces will have more neighbors in face space than will others, plays a major role in these explanations. Differences in relative exemplar density between own- and other-race faces, as shown schematically in Figure 1, are thought to underlie the observed race effects (Valentine, 1991; Valentine & Endo, 1992). The perceptual dimensions that form the basis of face space are "tuned" to capture the subtle difference between faces. As people's experience tends to have occurred with faces of their own racial group, the feature dimensions underlying the space are optimal for the discrimination of own-race faces. The resulting suboptimal encoding of other-race faces leaves them more densely clustered. Thus, the own-race bias results from a lack of experience with other-race faces. Our perceptual expertise with faces of our own race gives rise to perceptual features that do not generalize well to faces of another race (Chiroro & Valentine, 1995; Valentine, 1991; Valentine & Endo, 1992).

An alternative account of the own-race bias holds that race is explicitly encoded for other-race faces but not for own-race faces (Levin, 2000; see also MacLin & Malpass, 2001). This "race-feature" hypothesis holds that the automatic coding of other-race faces emphasizes visual information that specifies race at the expense of individuating information. By contrast, encoding of own-race faces emphasizes individuating information but not race. However, some individuating information must certainly be extracted and stored for other-race faces because they can be individuated, albeit less accurately than own-race faces. This idea of explicitly encoding race differs from the account of race effects normally described within the MDS framework but is still consistent with a face space approach. For example, if features that specify race correspond to a subset of dimensions of face space, the race-
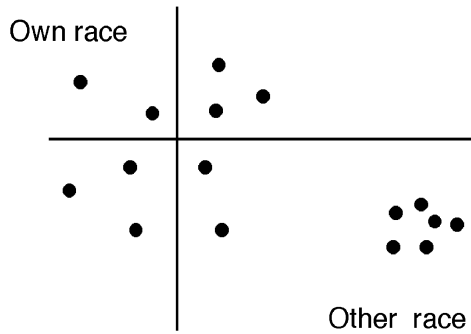
Figure 1. Schematic representation of own- and other-race faces in face space.

feature hypothesis might be formulated in the following way: Dimensions related to race will receive a relatively greater weight for other-race faces than for own-race faces. By this account, there is relatively less individuating information for other-race faces. In contrast, own-race faces are primarily encoded on non–race-related dimensions.

It remains to be seen whether own- and other-race faces form distinct clusters in face space and also whether differences in relative exemplar density between own- and other-race faces are sufficient to explain race effects on performance measures. It is also an open question whether the same dimensions are equally diagnostic for both own- and other-race faces or whether some kind of race-feature coding does occur.

The relative locations of faces within a psychological feature space can be determined by a statistical technique known as multidimensional scaling, which derives a spatial configuration of stimuli from psychological data such as ratings of perceived similarity (Shepard, 1957, 1980). The MDS framework can then be tested by using the derived stimulus coordinates as input to an explicit computational model. An existing model, the generalized context model (GCM; Nosofsky, 1986), is ideally suited for this task, having been shown capable of predicting a wide range of performance measures from stimulus locations derived from multidimensional scaling (e.g., Nosofsky, 1985, 1986). According to the GCM, the probability of correct identification is determined by the similarity relations between stimulus representations (Nosofsky, 1986). This combination of the GCM and multidimensional scaling was used in a previous study (Lee, Byatt, & Rhodes, 2000) to successfully predict identification data for caricatures and anticaricatures of famous male faces. Therefore, the MDS framework can provide a good explanation for the identification of famous faces. In addition, Busey and Tunnicliff (1999) have shown that the framework can also explain old/new recognition performance with unfamiliar faces.

It is worth noting that, although the term "recognition" is often used in a general sense to encompass both episodic recognition (having been seen previously) and also identification (individuating specific exemplars), the majority of

evidence for the own-race recognition advantage comes from episodic recognition studies (Meissner & Brigham, 2001). Given the importance of face *identification*, both ecologically and in the context of cross-race eyewitness memory, more emphasis should be placed on understanding race effects within the identification paradigm.

The main aim of the present study was to explicitly investigate whether the spatial relationships between own- and other-race faces in face space can explain identification performance. Specifically, we sought to determine whether a formal computational model could successfully account for identification performance on own- and other-race faces. Caucasian participants were trained to identify Caucasian and Chinese faces. Subsequent testing used different photographs of the test faces to ensure that the task was based on *face* identification and not *picture* identification. In addition, multidimensional scaling analyses of similarity ratings were able to show whether the poor identification of other-race faces was due to their relatively greater exemplar density in face space. A related question, which could also be answered by examining the multidimensional scaling solution, was whether the Caucasian and Chinese faces would indeed form separable clusters within face space.

As a secondary aim, we wanted to investigate the possibility that race is explicitly coded for other- but not own-race faces (the race-feature hypothesis). If other-race faces are predominantly coded in terms of race-based information, we should find that dimensions differentiating the races play a greater role in identification of other-race faces than of own-race faces. In this case, the best-fitting attention weights should differ, depending on the race of the face.

General procedural details that apply to all aspects of the study are discussed next. Specific discussion of the procedure and results for each section (similarity ratings, identification task, deriving a face space, and modeling the identification data) will then be described in turn, followed by results and discussion.

METHOD

General Method

Participants. All participants ($N = 52$) were students at the University of Western Australia and participated for course credit or $5. All were Caucasian and relatively unfamiliar with Chinese faces.

Stimuli. The stimuli consisted of two sets (smiling and neutral expression sets) of digitized (300×400 pixels, 256 gray levels) photographs of 42 young adult males (21 Caucasian, 21 Chinese). The Caucasian models were students from New Zealand and the United States. The Chinese models were students in New Zealand who came from all over Asia. The smiling versions were used as training stimuli for the identification task. The neutral expression stimuli were used as test faces for the identification task and to obtain similarity ratings. An average face of each race was created from the neutral expression set with Gryphon Morph software. The spatial layout of each face was defined by 656 landmark points corresponding to anatomical landmarks (face outline, ears, nose, mouth, and prominent facial lines). The average Caucasian face was constructed by calculating the mean location of each point and averag-

ing corresponding gray level values across the 21 Caucasian faces. The average Chinese face was made from the 21 Chinese faces, using the same procedure. There were a total of 44 stimuli (22 of each race).

Each face image was placed inside an oval mask in order to cover most of the outer hair while leaving the chin, ears, and inner hairline visible (see Figure 2 for examples of the test stimuli). Each face was scaled to ensure that the visible oval within the mask had the same area (60,000 pixels). The border of the mask was a medium gray (pixel intensity of 128). The region within the oval (i.e., the face itself) was also normalized to a mean pixel intensity of 128 to equate overall intensity for each race.

Similarity Ratings

Participants. Twenty-two Caucasian students (19 female, 3 male) received either $5 or course credit for participating in a 40-min session.

Stimuli. The 42 neutral expression face stimuli and the two averages were used, giving 44 stimuli (22 Chinese, 22 Caucasian).

Procedure. All participants (tested individually) rated the similarity of all 946 unique face-pairs on a 7-point scale (1 = *relatively dissimilar*, 7 = *relatively similar*) with the following instructions:

> You will be shown pictures of pairs of faces. There are Chinese and Caucasian faces in the experiment. Facial characteristics differ across racial groups but some Caucasian faces may still be quite similar to some Chinese faces and vice versa. Faces of the same race may also differ markedly in terms of similarity. Try to consider each pair of faces as two individuals, regardless of race, and answer the following question. How *Similar* are the two faces?

Stimulus presentation and response recording were controlled with SuperLab (Cedrus Corporation) running on a Power Macintosh with a 17-in. monitor (screen resolution, 832 × 624 pixels). The left/right position of each face within pairs and the presentation order were randomized separately for each subject.

Figure 2. Sample of the stimuli used, with Caucasian faces along the top row and Chinese faces along the bottom row.

Identification Task

Participants. Thirty different participants (20 female, 10 male) participated in the identification task.

Stimuli. The test stimuli were identical to those used to obtain the similarity ratings as described previously. The training stimuli consisted of smiling versions of the same faces.

Training procedure. All participants were tested individually in a single 45-min session. They were instructed that they would learn the names for some faces and that a recognition test would follow. The names were randomly assigned to faces for each subject. A HyperCard program, running on a Power Macintosh computer, controlled the presentation of stimuli and data collection.

Training consisted of three phases. First, the 44 smiling faces were divided randomly into 11 sets of 4 (with different random allocations for each subject). The program cycled through each set of 4 faces in different random orders, five times. On the first cycle, each face was displayed with the correct name and remained visible (3-sec minimum) until the subject clicked the mouse. For the remaining four cycles, all four names were shown and each face remained visible until the subject clicked on one of the names. The program provided feedback and the correct name before displaying the next face. This procedure was repeated for all 11 sets of faces. In the second phase, the 44 faces were again randomly split into 11 different sets of 4. The procedure was then similar to that for the first phase except that the correct names were not shown initially, but feedback was still provided. There were only two cycles through each set of 4 faces before moving on to the next set. In the third training phase, there were two cycles through all 44 faces in different random orders. Because this required that the subject choose from among 44 screen buttons, 26 smaller buttons labeled A–Z were available, and clicking on one of these buttons highlighted the names beginning with that initial letter.

Testing procedure. The testing phase was identical to the final phase of training except that there were six cycles through the 44 neutral expression faces (in a different random order on each cycle), and no feedback was provided.

RESULTS

Similarity Ratings

The similarity ratings were highly reliable across participants (Cronbach's alpha = .90). The mean similarity of each Caucasian face to all other Caucasian faces (M = 3.9) was greater than to Chinese faces (M = 2.9) [$t(42)$ = −11.62, p < .0001]. The mean similarity of each Chinese face to all other Chinese faces (M = 4.6) was greater than to Caucasian faces (M = 2.9) [$t(42)$ = 11.64, p < .0001]. These two results suggest that Chinese and Caucasian faces do form distinct clusters. Moreover, as hypothesized, the Chinese faces were rated as more similar to each other than Caucasian faces were to each other [$t(42)$ = 5.47, p < .0001], suggesting that the Chinese faces are more densely clustered than Caucasian faces.

Multidimensional Scaling

The similarity ratings were collapsed across participants, and the resulting mean similarity matrix was subjected to the ALSCAL multidimensional scaling procedure. A six-dimensional solution (Euclidean distance metric) was obtained, which accounted for 99.4% of the variance, with stress = 0.03.[1] The distribution of faces on the six dimensions is shown in Figure 3. It is clear from Figure 3 (top panel) that Dimension 1 separates faces on
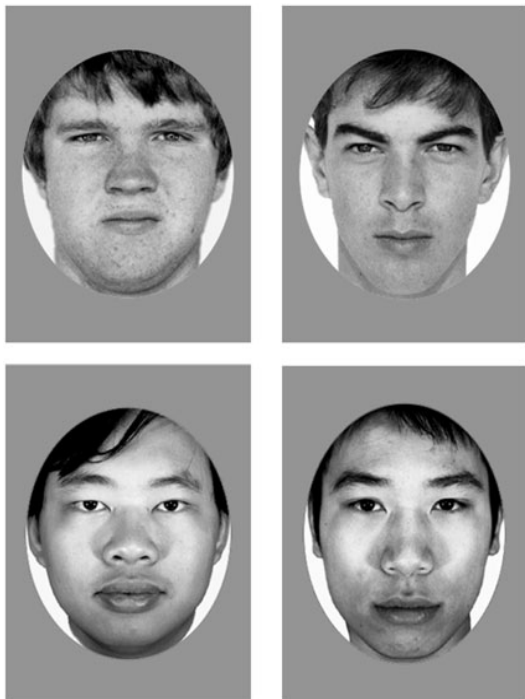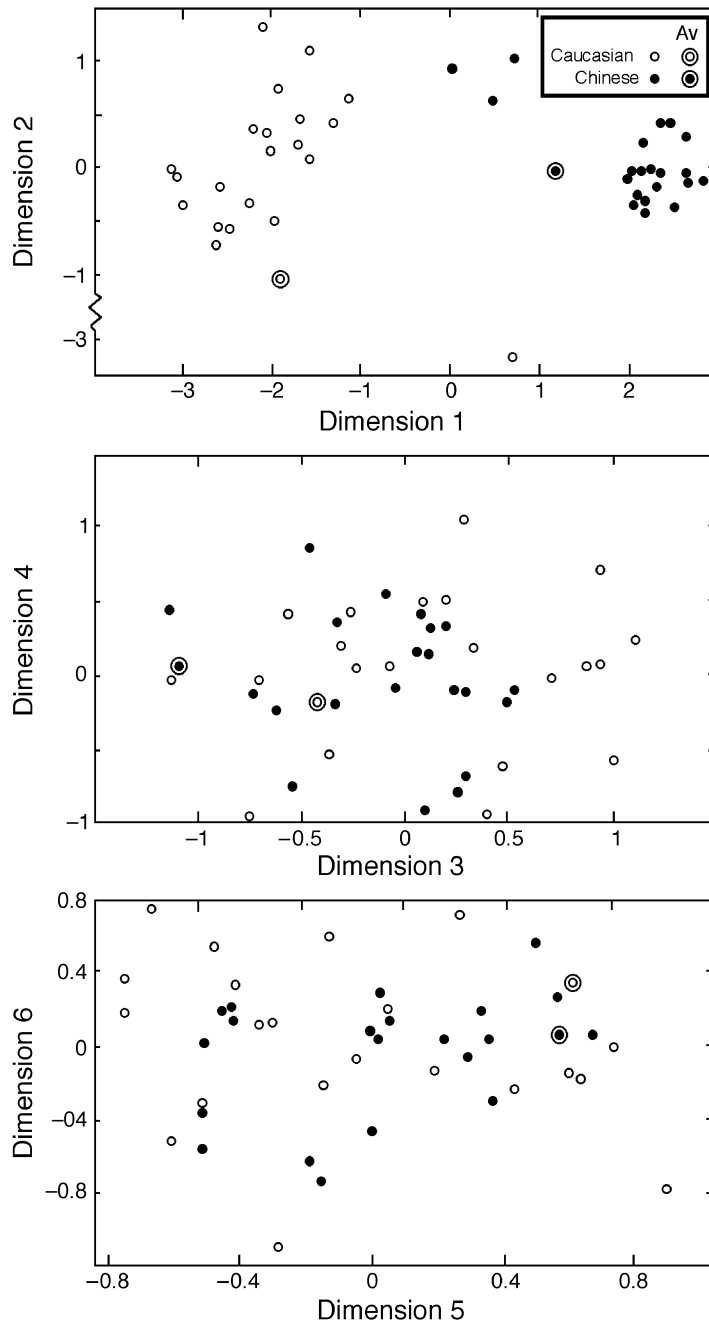
Figure 3. Scatterplots showing the location of each face on all six dimensions.

the basis of race. The other dimensions had no obvious interpretation.[2] Analysis of the overall distances between stimuli showed that as predicted, Chinese and Caucasian faces do constitute separable clusters, with Chinese faces being more densely clustered than Caucasian faces. Specifically, the mean distance of each Caucasian face to all other Caucasian faces ($M = 2.1$) was less than that to all Chinese faces ($M = 4.4$) [$t(42) = 13.29, p < .0001$]. Similarly, the distance of each Chinese face to all other

Chinese faces ($M = 1.5$) was less than to Caucasian faces ($M = 4.4$) [$t(42) = 22.19, p < .0001$]. Furthermore, the Chinese faces were more densely clustered (distance of each Chinese face to all other Chinese faces, $M = 1.5$) than Caucasian faces ($M = 2.1$) [$t(42) = 3.93, p < .0005$]. The morphed average for each face was not located centrally in the relevant race cluster (see Figure 3), perhaps due to its distinct (smoother) skin texture. Lee et al. (2000) have reported a similar result for an MDS for male celebrity faces.

Identification Performance

Separate *t* tests (two-tailed) were carried out on hits, false alarms (FAs)[3], and *A′* scores, with both items (faces) and subjects as random factors (see Table 1). The analysis by faces showed that Caucasian faces enjoyed significantly higher hit rates [$t(42) = 3.43, p < .002$], fewer FAs [$t(42) = 2.49, p < .02$], and higher *A′* scores [$t(42) = 3.51, p < .002$] than did Chinese faces. These results were confirmed with subjects as the random factor (correlated-sample *t* tests, $df = 29$, all ts > 3.5, all ps < .002).

Modeling of Identification Performance

Identification performance was modeled with a quantitative model based on the GCM (Nosofsky, 1985, 1986; see also Lee et al., 2000). The probability that stimulus *i* produces response *j*, $P(R_j | S_i)$, is found by dividing the similarity of stimulus *i* to stimulus *j* by the summed similarity of stimulus *i* to all stimuli:

$$P(R_j | S_i) = \frac{s_{ij}}{\sum_k s_{ik}}.$$

The similarity between stimulus *i* and face *j*, s_{ij}, is defined by a negative exponential decay function of psychological distance (Shepard, 1957):

$$s_{ij} = e^{-d_{ij}}.$$

The psychological distance (d_{ij}) is modeled as the weighted Euclidean distance:

$$d_{ij} = c\left[\sum_M w_m |x_{im} - x_{jm}|^2\right]^{\frac{1}{2}},$$

where *c* is a sensitivity parameter that accounts for the overall discriminability of stimuli in the *M*-dimensional space, x_{im} is the value of face *i* on dimension *m*, and w_m is the weighting of dimension *m* (i.e., the relative importance of each dimension for this task). The attention weights were constrained to sum to one, leaving six free parameters. For simplicity in the formal modeling, we did not include the bias parameters in deriving the fits of the GCM to the data.

The input to the model consisted of the stimulus coordinates from the multidimensional scaling procedure. A computer search was done to find the parameter values that minimized the summed square error (*SSE*) between the predicted performance and the observed iden-

tification data (averaged across participants). The model accounted for 81% of the variance (r^2) in the identification data and accurately predicted the higher identification accuracy for same- than for other-race faces (see Table 2, all faces). From Table 2, it can be seen that relatively little attention was paid to Dimension 1, which is consistent with the interpretation of this dimension as corresponding to race (see Figure 3) and thus being less useful for discriminating among individuals. The remaining five dimensions were approximately equally weighted, suggesting that all five are important for individuation.

Separate *t* tests (two-tailed) carried out on predicted hits, FAs, and *A′* scores mirrored the pattern of results found for the observed data (Table 1). Caucasian faces were predicted to have higher hit rates [$t(42) = 4.97, p < .001$], lower FAs [$t(42) = 4.69, p < .001$], and higher *A′* scores [$t(42) = 5.13, p < .001$] than would Chinese faces. As can be seen from Table 1, predicted hits, FAs, and *A′* scores very closely mirror the observed values.

Model fitting was also carried out separately for each race of face (but still using the same multidimensional scaling solution). This analysis allowed best-fitting attention weights to be determined independently for Chinese and Caucasian faces. As can be seen from Table 2, the dimension that differentiates clearly on race (Dimension 1) is weighted more strongly for Caucasian faces than for Chinese faces, which is the opposite of the weighting that would be expected according to the race-feature hypothesis.

DISCUSSION

According to the MDS framework, the poor recognition of other-race faces is due to their greater exemplar density relative to own-race faces. This dense clustering of other-race faces comes about because they are coded along feature dimensions that are optimally tuned for own-race faces and hence, suboptimal for faces of a different race. The present results support the standard face space account. The similarity ratings showed that other-race faces are perceived to be more similar to each other than are own-race faces. This greater within-race similarity was reflected in the multidimensional scaling solution that showed that own- and other-race faces do indeed constitute distinct clusters, with other-race faces being more densely clustered than own-race faces. Furthermore, the modeling results showed that, for both

Table 1
Observed and Predicted Identification Performance for
Chinese and Caucasian Faces

	Hits (%)		FA (%)		*A′*	
	Chinese	Caucasian	Chinese	Caucasian	Chinese	Caucasian
Observed	36.6	56.8	1.4	1.1	.824	.885
Predicted	37.0	55.3	1.6	0.9	.828	.882

Note—FA, false alarm.

Table 2
Results of Model Fitting Showing the Fit Sum Square Error
(SSE), r^2, and the Relative Weighting Given to Each Dimension

Data Fitted	SSE	r^2	c	Attention Weights					
				1	2	3	4	5	6
All faces	2.16	.813	12.6	.05	.22	.19	.14	.27	.13
Chinese	1.36	.632	9.7	.08	.14	.22	.21	.34	.00
Caucasian	0.69	.910	11.5	.12	.33	.14	.12	.07	.22

Note—r^2, proportion of variance accounted for; c, overall scaling parameter.

races, the spatial location of faces was an accurate predictor of the own-race identification advantage.

An alternative account, the race-feature hypothesis, suggests that the encoding of other-race faces emphasizes information specifying race at the expense of individuating information. There was no evidence that other-race faces are explicitly coded primarily in terms of race-specifying information. It is clear from the scaling solution that only Dimension 1 corresponds to race (see Figure 3), but attention allocated to this dimension was no higher for other-race faces than it was for own-race faces. Subjects weighted Dimensions 2–5 highly for other-race faces and did not rely solely on Dimension 1. The fact that participants were clearly able to discriminate among the other-race faces (Chinese faces, $A' = .824$) shows that more than just race information was being encoded. It should be noted that participants were *trained* to name the faces, so they were motivated to focus on individuating information. However, the race-feature hypothesis (Levin, 2000) holds that race is automatically emphasized *at the expense of individuating information* for other-race faces, and this is not subject to task demands. Furthermore, the predicted identification performance was based on similarity ratings from different participants who were merely asked to judge the similarity of face pairs, and the accuracy of the predictions suggests that even under these relatively passive viewing conditions, people extract individuating information from both their own race faces and faces from a less familiar race. Thus, it does not seem to be the case that the poor identification of other-race faces results from selective encoding of race-specific information.

It should be noted that although we have been speaking in general terms about own- and other-race faces, the participants in this study were all Caucasian. The important point for this study is that we observed an identification bias in favor of own-race faces that could be accounted for by the spatial locations of stimuli in psychological face space. We expect that this result would generalize to other races of observers (and faces) because, as discussed earlier, the "own-race bias" is extremely robust, and there is considerable evidence that it is not driven by either stimulus or observer characteristics (Bothwell, et al., 1989; Goldstein, 1979a, 1979b; Meissner & Brigham, 2001) and because there is no reason to think that the Chinese faces were more physically similar than the Caucasian ones (see the Method section).

Our aim here was to take a straightforward MDS approach and to demonstrate that the application of formal methods enables the framework to serve as a plausible psychological theory providing a basis for explanation and precise quantitative prediction. That the approach proved generally successful suggests that this aim was met. We do not, however, suggest that this particular combination of face space and formal methodology is necessarily optimal. Alternative formulations might provide a better description of the perceptual representations and cognitive processes underlying face perception. For example, it may be too simplistic to treat faces as single points in space (as noted by Valentine, 2001), and perhaps it would be better to treat individual representations as multivariate probability distributions rather than discrete points, given a certain amount of "noise" in perceptual or memorial systems (Ashby, 1992; Nosofsky, 1992). It may also be possible that face space consists of many separate multivariate normal distributions, similar to the "manifold model" discussed by Craw (1995). A comparison of different methods is beyond the scope of this article and remains for future study.

It is also worth noting that the multidimensional scaling solution is rotation invariant, and perhaps some optimal rotation would provide a better fit and allow "meaningful" dimensions to emerge. Some researchers do indeed provide interpretations of all the dimensions of face spaces derived from multidimensional scaling. The experimenter decides a priori what the "features" of the space should be, rotates the solution to fit these variables, and then interprets the dimensions of face space. Our view is that there is no reason to assume that the dimensions of face space correspond to any easily articulated feature (see, for example, Hancock, Burton, & Bruce, 1996; O'Toole, Abdi, Deffenbacher, & Valentin, 1993).

To conclude, the results suggest that people encode individuating information from all faces and that own- and other-race faces are encoded in terms of the same set of features. It seems likely that, as Valentine proposed, these features are tuned through a process of perceptual learning to capture the natural variation in own-race faces and are suboptimal for discriminating other-race faces. We have shown that the resultant increased exemplar density of other-race faces within face space can explain why they are more difficult to identify.

REFERENCES

Ashby, F. G. (1992). Multidimensional models of categorization. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 449-484). Hillsdale, NJ: Erlbaum.

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

Bruce, V., Burton, M. A., & Dench, N. (1994). What's distinctive about a distinctive face? *Quarterly Journal of Experimental Psychology*, **47A**, 119-141.

Busey, T. A., & Tunnicliff, J. L. (1999). Accounts of blending, distinctiveness, and typicality in the false recognition of faces. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **25**, 1210-1235.

Byatt, G., & Rhodes, G. (1998). Recognition of own-race and other-race caricatures: Implications for models of face recognition. *Vision Research*, **38**, 2455-2468.

Chiroro, P., & Valentine, T. (1995). An investigation of the contact

hypothesis of the own-race bias in face recognition. *Quarterly Journal of Experimental Psychology*, **48A**, 879-894.

CRAW, I. (1995). A manifold model of face and object recognition. In T. Valentine (Ed.), *Cognitive and computational aspects of face recognition: Explorations in face space* (pp. 183-203). London: Routledge.

GOLDSTEIN, A. G. (1979a). Facial feature variation: Anthropometric data II. *Bulletin of the Psychonomic Society*, **13**, 191-193.

GOLDSTEIN, A. G. (1979b). Race-related variation of facial features: Anthropometric data I. *Bulletin of the Psychonomic Society*, **13**, 187-190.

HANCOCK, P. J. B., BURTON, A. M., & BRUCE, V. (1996). Face processing: Human perception and principal components analysis. *Memory & Cognition*, **24**, 26-40.

JOHNSTON, R. A., KANAZAWA, M., KATO, T., & ODA, M. (1997). Exploring the structure of multidimensional face-space: The effects of age and gender. *Visual Cognition*, **4**, 39-57.

JOHNSTON, R. A., MILNE, A. B., & WILLIAMS, C. (1997). Do distinctive faces come from outer space? An investigation of the status of a multidimensional face-space. *Visual Cognition*, **4**, 59-67.

KRUSKAL, J. B., & WISH, M. (1978). *Multidimensional scaling*. Beverly Hills, CA: Sage.

LEE, K. J., BYATT, G., & RHODES, G. (2000). Caricature effects, distinctiveness, and identification: Testing the face-space framework. *Psychological Science*, **11**, 379-385.

LEVIN, D. T. (2000). Race as a visual feature: Using visual search and perceptual discrimination tasks to understand face categories and the cross-race recognition deficit. *Journal of Experimental Psychology: General*, **129**, 559-574.

MACLIN, O. H., & MALPASS, R. S. (2001). Racial categorization of faces: The ambiguous race face effect. *Psychology, Public Policy, & Law*, **7**, 98-118.

MEISSNER, C. A., & BRIGHAM, C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, & Law*, **7**, 3-35.

NOSOFSKY, R. [M.] (1985). Overall similarity and the identification of separable-dimension stimuli: A choice model analysis. *Perception & Psychophysics*, **38**, 415-432.

NOSOFSKY, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General*, **115**, 39-57.

NOSOFSKY, R. M. (1992). Similarity scaling and cognitive process models. *Annual Review of Psychology*, **43**, 25-53.

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.

RHODES, G., BYATT, G., TREMEWAN, T., & KENNEDY, A. (1997). Facial distinctiveness and the power of caricatures. *Perception*, **26**, 207-223.

SHEPARD, R. N. (1957). Stimulus and response generalization: A stochastic model relating generalization to distance in psychological space. *Psychometrika*, **22**, 325-345.

SHEPARD, R. N. (1980). Multidimensional scaling, tree-fitting, and clustering. *Science*, **210**, 390-398.

VALENTINE, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *Quarterly Journal of Experimental Psychology*, **43A**, 161-204.

VALENTINE, T. (2001). Face-space models of face recognition. In M. J. Wenger & J. T. Townsend (Eds.), *Computational, geometric, and process perspectives on facial cognition: Contexts and challenges* (pp. 83-113). Mahwah, NJ: Erlbaum.

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.

NOTES

1. Kruskal and Wish (1978) suggest that a stress value of 0.05 indicates a good fit, whereas values above 0.20 represent a bad fit.

2. We prepared scatterplots as in Figure 3 but with the actual faces in place of symbols and asked several colleagues and students to identify characteristics of the faces that might correspond to each dimension—with no success.

3. A false alarm is recorded against a face when that face's name is given in response to some other face. A hit is recorded against a face when it is named correctly.