

Do "foreign" faces really look alike?

ALVIN G. GOLDSTEIN and JUNE CHANCE
University of Missouri, Columbia, Missouri 65211

White subjects judged the degree of apparent resemblance of pairs of Japanese and pairs of white American faces using a 10-step psychometric scale. Judgments were made either in a "perceptual" condition (both faces seen simultaneously) or in a "memory" condition (sequential presentation, zero time between faces). Mean resemblance judgments were not affected by the method of presenting the pairs. Contrary to the results obtained in earlier investigations, Japanese and all women's faces were judged to be significantly more alike than white and all men's faces. No interactions were significant. These data, considered alone, support the hypothesis that Japanese faces are perceived to be more alike than white faces. However, these results, when considered in the context of the results of the earlier studies, are not compelling evidence for the homogeneity hypothesis. It is unlikely that the other-race memory effect can be accounted for by the differences in judged similarity of Japanese compared to white faces.

For reasons that still remain unclear, portraits of briefly seen white faces, in contrast to either Japanese or black faces, are more accurately recognized by white subjects (Chance, Goldstein, & McBride, 1975; Malpass & Kravitz, 1969). This finding, which might be labeled the "other-race"¹ phenomenon, has attracted the attention of investigators with theoretical and practical interests. The social and interpersonal implications of this admittedly mild memory deficit are perhaps clearer than its theoretical implications; nonetheless, a solution to the other-race puzzle might shed light on what kind of information is coded and stored in long-term memory for visual stimuli. Thus, in an attempt to account for the increased error rate in response to other-race faces, several hypotheses have been proposed.

The present experiment is the most recent in a series of investigations designed to test the hypothesis that psychological similarity (i.e., degree of within-group stimulus homogeneity) among Japanese faces is greater than psychological similarity among white faces. Briefly, the hypothesis asserts that Japanese faces in a recognition memory study are more confusable than white faces because the psychological distances between Japanese faces on a similarity dimension are smaller than the distances between white faces on the same dimension. In five earlier investigations, measures of psychological similarity were collected in a variety of laboratory tasks, with all procedures utilizing simultaneous viewing of at least one pair of faces (Goldstein & Chance, 1976, 1978). With one exception, discussed below, evidence for differential degrees of within-groups psychological similarity was

entirely lacking. Those studies suggested that white college-age subjects, when judging resemblance or likeness, do not judge groups of Japanese faces to be more (or less) homogeneous than groups of white American faces, although some subjects still remarked that "they all look alike." We will return to this issue later in this article.

In the last experiment reported (Goldstein & Chance, 1978, Experiment 4), a match-to-sample search task was used to measure the time taken by subjects to find a duplicate portrait embedded among numerous same-race and same-sex distractor portraits. This method assumes that if Japanese faces on the average are more alike than white faces, then locating a particular Japanese face in an array of Japanese faces should take longer than locating a white face because of the higher target-distractor similarity among Japanese faces. Although the main effect of race was not significant, the interaction of race and gender of the face obtained was significant, a finding not compatible with the results of the four earlier investigations. White faces of both sexes were located after equal search times (9.5 and 9.3 sec), but Japanese women's faces were located more rapidly (8.2 sec) than all other faces, and Japanese men were the most difficult to find (11.7 sec). Because these data resisted coherent interpretation, they stimulated further research.

In this experiment, within-races similarity judgments of Japanese and white faces were collected on a 10-point scale potentially more sensitive to differences in subjective evaluation than the more or less resemblance judgments employed in the earlier research. Also, it was not yet clear whether the basis of the other-race phenomenon could be a memory rather than a perceptual deficit. To illuminate this problem, half the subjects in this study made their judgments with both members of a pair of faces in view, and the remaining subjects

This research was assisted by grants from the Graduate School and Office of Research awarded by the University Research Council of the University of Missouri.

saw first one face of the pair, followed immediately by the second member of the pair.

METHOD

Stimuli

Japanese and white American portraits were used as stimuli. The source of the white faces was original photographs used in the production of a university yearbook; the Japanese portraits were photographs that had accompanied applications for admission to a Japanese university.² All photographs were achromatic, showing a full-face pose of the head only. Japanese photos were slightly smaller than white photos, but the quality of both was about equal. Each portrait was mounted on a cardboard sheet for ease of handling.

Subjects and Procedure

The experiment involved two conditions of viewing (simultaneous or sequential), two kinds of faces (Japanese and white), and male and female faces. Forty white subjects, volunteers from introductory psychology classes, were randomly assigned to one of the two viewing conditions. Similarity judgments were made in the simultaneous condition while both portraits were in view for 10 sec; in the sequential condition, each portrait of a pair was shown separately for 5 sec, with a fraction of a second interval between the first and second portrait. Timing was approximate; the experimenter used the second hand of a standard timer to estimate the 5- and 10-sec intervals. Stimuli were displayed by hand; each subject was tested individually, and responses were recorded by the experimenter. Subjects in both conditions heard the same instructions (below), except the words in parentheses were inserted for either the successive or sequential groups. Full instructions are reproduced below.

"It is well known that two people who are related, such as brothers or cousins, may resemble each other. We often say that there is a family resemblance. However, relatives are not the only people who may look alike. I'm sure you have known two people unrelated to each other who resembled each other. Sometimes the resemblance is strong and sometimes it is weak, but still noticeable. In this experiment we are trying to find out whether people can judge the *amount* of resemblance between two people. I am going to show you two portraits (one at a time, for 5 sec each) (simultaneously for 10 sec). Please look at each portrait, and after you have looked at both, I would like you to assign a number from the scale in front of you that you think will express the degree or amount of similarity you see between the two people in the pictures. In other words, if you think that there was no resemblance between the two people in the photos, then tell me this by saying zero or one. If, on the other hand, the portraits resemble each other rather strongly, then tell me this by saying eight or nine, or any number in between. Please remember that you have only a few seconds to look at each portrait."

On a table directly in front of the subjects at all times was a small calibrated cardboard ruler with the numbers 1-10 inscribed along its length, and the words, "no resemblance" and "strong resemblance" above the 0 and 10 points, respectively. Each subject judged 12 pairs of Japanese faces and 12 pairs of white faces randomly intermixed in the same testing session. Pictures of men and women were equally represented within both groups of faces. Face pairs were randomly composed, except both faces had to be of the same sex and race. Two subjects, one each from the perceptual and memory conditions, saw a unique set of 24 pairs in a unique random order. Thus, with 20 subjects in each condition, all the faces in the experiment were randomly paired on 20 different occasions, and two subjects were tested with each of the random sets.

RESULTS

Preliminary analysis of the individual responses to each pair of faces revealed that subjects utilized values at all points along the scale, suggesting that they found the discrimination requested meaningful. For example, 36 of the 40 subjects used 0 or 1 and 8, 9, or 10 in response to at least two pairs of faces. The grand mean, 3.88, was below the midpoint of the scale. This value is reasonable, since most pairs of randomly selected faces should appear to be unrelated.

Analysis of variance computed on the data, summarized in Table 1, revealed that judgments of resemblance in the simultaneous and sequential conditions were not reliably different [$F(1,38) = 1.74$, $p > .05$], nor were any of the interactions involving this factor. Japanese face pairs were judged to resemble each other to a significantly greater degree than white face pairs [$F(1,38) = 16.02$, $p < .01$], and women's faces were judged to be more alike than men's faces [$F(1,38) = 4.53$, $p < .05$]. Neither sex nor race of portrait interacted with any other factor (all $F_s < 1.00$).

DISCUSSION

What the present study shows is that subjects given a 10-point scale to express their psychological judgments about resemblances of faces find greater similarity among Japanese than among white faces. Simultaneous or sequential viewing of the pair of faces being judged had almost no effect on the similarity judgment, suggesting that perceived similarity is not increased by memory processes, or at least not by processes brought into play by short delays. The observed effect—Japanese faces judged to be more alike than white faces—is less than 1 unit (.8) on the 10-point scale and represents slightly more than a 20% increase in apparent similarity for the Japanese faces. This result, along with the additional finding that women's faces are judged to be more alike than men's faces, is at odds in one way or another with performance data obtained from subjects who were asked to rapidly detect perceptual similarities and differences in both kinds of faces (Goldstein & Chance, 1976) and with the results of Experiments 1, 2, and 3 of an earlier investigation where similarity was judged using a simple more-less scale (Goldstein & Chance, 1978).

One of the more important discrepancies in outcome may be due to the relatively coarse scale of judgment used in the earlier investigations. This interpretation implies that subjects may perceive Japanese faces as being more alike, but only under special conditions do they judge them to be more alike. This disparity in findings suggests that the effect being measured is not robust.

Table 1
Mean Resemblance Judgments in Response to White and Japanese Face Pairs as a Function of Display Condition

	Simultaneous		Sequential	
	Men	Women	Men	Women
White	3.4	3.9	3.3	3.5
Japanese	4.6	4.5	3.6	4.2
Mean	4.1		3.7	

Note—Scale: 0 = no resemblance, 10 = strong resemblance.

It is important to keep in mind when considering all the studies together that different methods were employed in measuring similarity. The fact that performance measures (e.g., latency for locating a matching face) do not agree with judgment measures of similarity may indicate that the two tasks do not measure the same behavior. It could also indicate that Japanese faces as stimulus configurations are really less similar in internal construction than they appear to the observers (Goldstein, Note 1). Thus, groups of Japanese faces may "look alike" on a naive, phenomenological level, but when the observer is pressed to make psychometric judgments, another level of behavior is involved, one that emphasizes perceptual analysis.

It is clear that at least two problems are involved in the other-race effect. First, white subjects do not recognize Japanese faces as well as they recognize white faces, and second, white people often report that Japanese faces all look alike. The former could be caused by the latter, but the two events could just as easily be unrelated. In fact, if the two questions are treated separately, some insight into the other-race response might be gained. For instance, at least for Japanese and black faces, the other-race effect appears to be well documented. But, as our research has demonstrated, it is difficult to show that recognition memory performance errors are related to the perceived similarity among Japanese faces. At this time, we conclude that the other-race memory deficit is probably not accounted for by the subjective homogeneity hypothesis. Although this conclusion does not point to an explanation of the other-race effect, it still serves a purpose, since it suggests where not to look for an answer to the problem.

The subjective state described by the phrase, "they all look alike to me," apparently is common to individuals from diverse racial and cultural groups, although the existing evidence seems to be mostly anecdotal. Nevertheless, anthropologists attest to personal experiences with peoples who react this way to "outsiders" or those different from them in skin color, place of origin, and so on (Robbins, Note 2), so it appears reasonable to suppose that "something" is happening when we look at foreign faces. But this event has proved to be elusive and difficult to measure, and therefore may not contribute much to our understanding of the other-race effect. In speculating on the reasons an observer reports that foreign faces look similar, several hypotheses come to mind. Perhaps the observer is reporting that his immediate reaction to the other-race person is a feeling of unfamiliarity, strangeness, or at least "difference" and, since all members of the other-race group elicit this reaction, they all look alike. In the same vein, an initial shock or preoccupation with differences when looking at a foreign face may interfere with the more customary encoding process. At least, in a real-life situation, this slight encoding disruption could lead to a less adequate memory trace and a subsequent feeling that the poorly encoded faces are similar. In addition, an immediate "labeling" reaction, "Japanese face," may focus attention upon a restricted set of stimuli (e.g., eye and cheek region in Japanese) and, again, the overall impression would be one of sameness or likeness. Faces are important social stimuli, and face recognition is a critical skill for maintaining adequate interpersonal relations, but we are rarely involved with other-race people in important or long-term interpersonal relationships, since often they are performing low-status services for us. Under these circumstances, other-race faces are less influential in their role as social stimuli than own-race faces. Stimuli lacking potency receive less attention; discrimination is impaired as a consequence of reduced attention. Repeated experiences with difficult discriminations might lead to the belief that the members of the other-race groups are more alike.

Finally, it is reasonable to suppose that other-race faces may be more difficult to "read" and therefore might elicit fewer abstract responses such as friendly, intelligent, generous, kind, and so on.³ It is also often said by whites that distinguishing age in black and oriental faces is difficult. These

perceptual difficulties should enhance the impression of facial similarity because many foreign faces would be seen as lacking "personality." That is, other-race faces tend to look alike because some of the inferences that we normally make when we see new faces are missing, and it may be these inferences that help us distinguish among faces of our own group, or at least assist us in reducing the homogeneity among own-race faces. It is worth noting that several of these speculations suggest that, in terms of the Craik and Lockhart (1972) depth-of-processing memory model, other-race faces are likely to be processed more superficially than own-race faces.

REFERENCE NOTES

1. Goldstein, A. G. *Race-related variation of facial features: Anthropometric data. I*. Submitted for publication, 1979.
2. Robbins, M. Personal communication, 1978.

REFERENCES

- CHANCE, J., GOLDSTEIN, A. G., & MCBRIDE, L. Differential experience and recognition memory for faces. *Journal of Social Psychology*, 1975, **97**, 243-253.
- CRAIK, F. I. M., & LOCKHART, R. S. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 1972, **11**, 671-684.
- GOLDSTEIN, A. G., & CHANCE, J. Measuring psychological similarity of faces. *Bulletin of the Psychonomic Society*, 1976, **7**, 407-408.
- GOLDSTEIN, A. G., & CHANCE, J. Judging face similarity in own and other races. *Journal of Social Psychology*, 1978, **98**, 185-193.
- MALPASS, R. S., & KRAVITZ, J. Recognition for faces of own and other race. *Journal of Personality and Social Psychology*, 1969, **13**, 330-334.
- MUELLER, J. H., CARLOMUSTO, M., & GOLDSTEIN, A. G. Orienting task and study time in facial recognition. *Bulletin of the Psychonomic Society*, 1978, **11**, 313-316.
- SHOEMAKER, D. J., SOUTH, D. R., & LOWE, J. Facial stereotypes of deviants and judgments of guilt or innocence. *Social Forces*, 1973, **51**, 427-433.

NOTES

1. Race is used only to refer to people who share in common facial features that perceptually distinguish them from members of other groups.
2. The authors are grateful to Professor T. Kaneko of the University of Education, Tokyo, Japan, for supplying the Japanese portraits.
3. Inferring personality traits from faces appears to be a common response. Several sources of evidence strongly suggest that many faces are labeled with attributes like "intelligent" and "trustworthy" with no more evidence than a photograph. In fact, one study has demonstrated that subjects will attribute specific deviant behaviors (murderer, rapist, etc.) to portraits of randomly selected (presumed innocent) middle-aged men, and these attributions are attached to individual portraits with better-than-chance consensual agreement (Shoemaker, South, & Lowe, 1973). In addition, it is common for experimenters to ask subjects to rate a series of portraits on a variety of personality traits (honesty, kindness, etc.; e.g., see Mueller, Carlomusto, & Goldstein, 1978), a task that is patently impossible, but to our knowledge, subjects never refuse to comply with these instructions. Implied in these examples is the behavioral truth, people perceive abstract, personality-related qualities in human faces. Whether, and to what degree, white people make these kinds of attributions in response to nonwhite faces is an empirical question we are investigating in our laboratory.

(Received for publication November 20, 1978.)