Bulletin of the Psychonomic Society
1986, 24 (4), 257-260

Recognition memory for infant faces: An analog of the other-race effect

JUNE E. CHANCE, ALVIN G. GOLDSTEIN, and BLAKE ANDERSEN
University of Missouri, Columbia, Missouri

College adults' recognition memory for and perception of resemblance among pictures of infants were tested in three experiments. Subjects had more difficulty recognizing briefly seen pictures of infants than pictures of adults, although those reporting some experience with infants did better than inexperienced subjects. In a second experiment, subjects rated adult male pictures as resembling each other more than infant pictures, but in a third experiment, when pairs of infant pictures and pairs of adult pictures were directly compared, infant pairs were more often chosen as being more alike than adult pairs.

In three experiments reported here, we tested whether adults have greater difficulty discriminating and identifying pictures of newborn infants than they do pictures of adults. Many adults believe that newborn infants look more alike than older children, just as they believe that people of ethnic groups other than their own look more alike than members of their own group (the other-race effect). That newborn infants may be more difficult for adults to distinguish from one another is also suggested by the measures employed in hospital nurseries to ensure that children are accurately identified. Nonetheless, in spite of identification bracelets, foot- and handprints, and photographs, cases still occur that suggest that babies in hospital nurseries have been mixed up (e.g., Kenney, 1979; "Maternity Ward Nightmare," 1978). If infants are not really more alike in facial structure than other humans, and adults do have more trouble telling them apart, then we may have an effect similar to the other-race effect based on similar reasons.

Various aspects of the other-race effect have been examined by a number of different investigators and the effect has been found to be quite robust (Malpass, 1981; Shepherd, 1981). An early study (Chance, Goldstein, & McBride, 1975) found that white college students recognized pictures of blacks less well than they recognized pictures of whites, and that, reciprocally, black college students recognized pictures of whites less well than pictures of blacks. Both black and white subjects recognized a sample of Japanese photographs less well than they recognized either white or black portraits. In order to interpret these results, we examined a variety of other possible differences between white and Japanese faces. Physical anthropometric data indicated no significant differences in heterogeneity of physical facial structure among many different ethnic groups of faces, including white and Japanese (Goldstein, 1979a, 1979b). In other studies, using several different techniques, the perceptual difficulty subjects experience in making discriminations among faces of individuals within groups of portraits of whites and Japanese was assessed (Goldstein & Chance, 1976, 1978, 1979, 1980). No reliable differences were observed in the accuracy or speed with which subjects discriminated among individuals within different groups of faces.

Although prejudiced ethnic attitudes might account for differential recognition, our Midwestern subjects had no systematic reason to be either pro- or anti-Japanese in their attitudes. Moreover, earlier literature examining relationships of ethnic attitudes and memory for faces failed to show that negative attitudes are systematically associated with either better or worse memory for faces (Malpass, 1981; Shepherd, 1981). Chance et al.'s (1975) findings also would not support the argument that differences found in studies using white and black faces are artifacts of poorer photographic resolution of black faces owing to darker skin, because the very poorly remembered Japanese are relatively lighter skinned than most American blacks.

Chance et al.'s (1975) findings do support different degrees of experience with persons of various ethnic groups as an important factor in the other-race effect. In that study, black and white subjects each had some—although perhaps limited—experience with the other group. In contrast, neither white nor black subjects, presumably, had more than very casual experience with Orientals, and it was on those faces that they did most poorly. Other investigators have also shown that differential memory for faces of different groups is associated with geographic distribution (Shepherd, Deregowski, & Ellis, 1974) and racial integration of living conditions (Cross, Cross, & Daly, 1971; Feinman & Entwhistle, 1976).

We tested the influence of experience more directly in a cross-sectional developmental study of face recognition among children ranging in age from 6 to 14 years and among college-aged subjects (Chance, Turner, & Gold-

Requests for reprints should be addressed to J. E. Chance, Department of Psychology, University of Missouri, Columbia, MO 65211.
The authors are grateful to Walter Goddard and Lisa Goldstein for providing the infant pictures used in the study.

stein, 1982). If different amounts of experience with various groups of faces were associated with differential recognition of faces of other groups, then older children should show greater disparity between their skills in recognizing white and Oriental faces than should younger children. This prediction was confirmed. Older subjects recognized all faces better than did younger subjects, but older subjects showed much larger differences in their performances between white and Japanese faces than younger subjects did. In another study, college women selected because of their initially very poor recognition of Japanese faces were trained to equal their excellent recognition of white faces by means of repeated exposures to new sets of Japanese faces in a paired associated learning task (Goldstein & Chance, 1985). Differences in experience definitely appear to be implicated in the other-race effect.

The use of infants' faces as another means to assess the role of experience in face recognition was suggested by two observations: first, infant faces, as a population of stimuli, are relatively rarely seen by most adults; and second, beliefs about the greater perceptual homogeneity of infant faces and of the faces of ethnic groups besides our own are similar. Moreover, the assumption made earlier that prejudiced attitudes were not an issue could be tested by using pictures of infants, toward whom most of our subjects should be positively or, at worst, neutrally disposed.

We undertook three experiments using infant faces as stimuli. In the first we asked whether infant faces were more difficult for subjects to recognize than were faces of adults, and whether subjects' experiences with real infants were associated with their performance. In the other two experiments, we used two different methods to assess subjects' perceptions of the relative similarity of infant and adult faces.

EXPERIMENT 1

In Experiment 1, we tested whether adult subjects would more accurately recognize briefly seen photographs of the faces of other adults than photographs of the faces of newborn infants. A brief questionnaire was administered to assess the degree of subjects' experiences with infants, to test whether experience with real infant faces was associated with performance on photographs of infant faces.

Method

Thirty-four women and 34 men looked at color slides of faces, and, after a brief interval, were asked to pick out the faces they had just seen from a larger sample of similar faces. Each subject saw two independent sets of target faces and was tested immediately afterward on two larger sets similar to the target faces. Small groups of 1 to 3 persons were tested. Before any pictures were shown, all subjects were administered a short questionnaire which asked about their experience with infants: "Have you ever taken care of newborn infants? How many? For how long? Have you recently held or played with an infant? Describe the situation." They were also asked whether they thought adults were especially attracted to infants and to what degree they agreed with the idea that most infants look alike. Half the subjects saw and were tested on sets of infant pictures first; the other half saw and were tested on pictures of adult females first. Approximately equal numbers of men and women participated in each of the two orders. Also, one half of the subjects in each order saw one set of 10 randomly selected target faces of each type, while the other half saw another set of 10. Each recognition test included 10 target faces mixed with 22 distractor faces. Study slides were shown one at a time for 3 sec each; recognition test slides were shown for 5 sec each. As each slide was shown, subjects marked on their answer sheets whether or not they thought they had seen that face before.

In the brief interval between looking at target pictures and performing the recognition tests, all subjects answered a brief personality questionnaire. When the first recognition test was completed, subjects were shown the alternative set of target pictures, and after answering another brief questionnaire, were again tested for recognition.

Infant pictures were duplicates of those taken for newborn-nursery hospital records. The adult photos were of college women. Adult faces were chosen to include no idiosyncratic features such as scars, jewelry, or glasses. The adults wore a smock-like drape in order to conceal their clothing. The adults' faces were presented in slightly different poses at study and at test. The infants were photographed in their cribs in the newborn nursery. Unfortunately, many of these pictures also included nonfacial details, such as the print on a gown, ostensibly making these pictures easier to recognize later; however, the prints and other details were repeated, allowing us to select pictures so that several infants in any set of photos would appear with the same details (e.g., a polka-dot blanket). Pictures of infants whose faces had distinctive aspects, such as many red blotches, were not used. Since the infants appeared sexless, we used photos of college women only, rather than both sexes.

Results

Very few of our subjects reported more than casual experiences with infants; therefore, the answers to the questions cited above about experiences with infants were used to make a very crude classification of subjects. We called subjects "experienced" if they reported any amount of experience beyond baby-sitting an infant once or twice. Two raters independently rated 25 subjects as "experienced" and 43 as "inexperienced," with 97% agreement.

Although we had assumed that many adults believe that infants look more alike than other people, 98% of our college sample strongly disagreed that infants do look alike. Our subjects also said, in about the same majority, that they thought adults were especially attracted to infants.

A three-way analysis of variance (subject sex × subject experience × picture type, with the latter as a repeated measure) was employed to test the recognition memory data of Experiment 1 (d′ scores). Although we found that sex of subjects made no difference, either alone or in interaction with the other variables, both subject experience [$F(1,64) = 7.698, p < .001, MSe = 1.853$] and picture type [$F(1,64) = 63.353, p < .0001, MSe = 0.839$] were strongly associated with recognition performance. Experience and picture type also interacted significantly [$F(1,64) = 5.714, p < .02, MSe = 0.839$]. All subjects recognized pictures of adults better than those of infants. Mean d′ scores were 2.79 for adult photos and 1.86 for infant photos. Experienced subjects performed better than inexperienced subjects on all pictures, regardless of type; means were 2.57 and 2.09, respectively. However, as the significant interaction suggested, the difference between

experienced and inexperienced subjects was larger for infant pictures than for adult pictures. Means were 2.90 and 2.66, respectively, when experienced and inexperienced subjects viewed adult faces, but were 2.24 and 1.48, respectively, when experienced and inexperienced subjects viewed infant faces.

EXPERIMENT 2

Although the subjects in Experiment 1 had indicated on a questionnaire that they strongly disagreed with the idea that infants looked more alike than older individuals, we wanted to ask about relative degrees of perceived resemblance in a more systematic and structured way. Would subjects making judgments about resemblance using actual faces of infants and adults evidence perceptual difficulties that could be associated with poorer memory for infant faces?

Method

In Experiment 2, pairs of adult male and infant faces were presented alternately to 13 women and 7 men. As each pair was presented, the subject was asked to rate the degree to which the members of the pair resembled each other. A rating scale running from 0 to 10 was laid on the table beneath the pair of face photographs to be rated. Subjects were told that a rating of 0 signified a total lack of resemblance between the faces and that 10 indicated a resemblance "like identical twins."

Subjects were tested individually, on two occasions separated by 2 weeks. On the first test, the experimenter gathered the set of infant photos in one deck and the adult male photos in another. Each deck was shuffled independently and the experimenter randomly selected pairs to be rated from the two decks alternately. The experimenter recorded both the subject's ratings for each pair and the identifying numbers of the pictures in each pair. Twenty-four pairs of pictures were rated.

Subjects rated these same pairs of photos a second time, 2 weeks later. We wanted to know whether subjects would rate the infant faces as resembling each other more than do adult faces, but we also wanted to know whether their ratings would be reliable, which would suggest that the ratings had meaningful relationships to the stimulus pictures.

Results

We summed each subject's ratings for infant and for adult picture pairs. A three-way ANOVA, with sex as the between-subjects factor and picture type and times of testing as the within-subjects factors, was performed. Sex of the subjects was not related to resemblance ratings on either testing; however, both men and women rated adult pictures as resembling each other more than infant pictures [$F(1,18) = 7.80$, $p < .01$, $MSe = 116.19$].

Although all ratings on the second test tended to move toward less perceptual resemblance, this change was not significant, with ratings averaging 44.89 on the first testing and 41.65 on the second. However, the picture type×times of testing interaction was significant [$F(1,18) = 4.26$, $p < .05$, $MSe = 157.60$]. Ratings of adult pictures changed more than ratings of infant pictures. Subjects' average ratings of resemblance between adult photos was 51.98 on the first testing and 45.70 on the second; the corresponding ratings for infant photos were 37.80 and 37.59. Pearson product-moment correlations between subjects' two ratings of faces were +.59 for adult faces and +.91 for infant faces.

EXPERIMENT 3

Experiment 3 was like Experiment 2, except that the procedure to assess judged resemblance was changed.

Method

Twelve women and 7 men were again asked to judge resemblance, but in this procedure two pairs of pictures—one of adult males and one of infants—were presented to the subject simultaneously, and the subject was asked to select which of the pairs resembled each other more. Pairs of pictures were selected randomly by the experimenter as described in Experiment 2, and the subjects were asked to make the same choice from the same sets of pairs 2 weeks later.

Results

On both occasions, the majority of subjects selected a larger number of infant pairs as being more alike. Wilcoxon's matched-pairs signed-ranks test was used to evaluate the differences [$T(19) = 16$ on the first occasion and 9 on the second occasion, both $ps < .005$]. A subject's consistency of choice was scored on the percentage of times in 12 possibilities that the subject chose the same pair as more alike on the second test as on the first. Subject percentages of agreement ranged from 92% to 58%, with a mean for all subjects of 75%.

DISCUSSION

In Experiment 1, college-aged subjects of both sexes had greater difficulty in recognizing briefly seen photos of newborn infants than in recognizing photos of college women. Subjects reporting even minimal experience with infants recognized more infant pictures than those reporting no experience. However, the majority of these college-aged subjects strongly disagreed with the statement that newborn infants look very much alike, and also said they thought adults were especially attracted to infants. Both opinions suggest positive attitudes toward infants.

Subjects in two other experiments rated and judged perceptual similarity of pairs of infant and adult male portraits. When a rating scale was used and adult pairs and infant pairs were presented alternately, subjects rated adult males as resembling each other more closely than infants resembled each other. This finding concerning judged resemblance corroborates the finding in Experiment 1 whose subjects said on the questionnaire that infants do not look more alike to them than other people do. It also suggests that judgments may reflect attitudes as well as perceptual discrimination. However, when adult and infant pairs were presented simultaneously and subjects were forced to choose which of the pairs resembled each other more, infant pairs were chosen over adult pairs in significant numbers. This finding is more in line with the investigators' original assumption, as well as more consistent with the memory data. Ratings under both conditions were repeated by subjects 2 weeks later with a very high degree of consistency.

The finding that infant pictures are harder for adult subjects to recognize than adult pictures supports the conclusion that some difficulties in face recognition may be due to the subject's lack of relevant past experience that would enable him/her to develop efficient face-processing schemata (Goldstein & Chance, 1980). This conclusion is further supported by the finding that particular subjects' reports of personal experience with infants are associated with better infant face recognition among those subjects. That experience with infant faces as a class should influence effectiveness in recognizing new infant faces is congruent with the findings about memory for faces reported in earlier studies as well as with other data in the memory literature. For instance, memory for chess moves (Chase & Simon, 1973; Chi, 1978), bird pictures (Peeck & Zwarts, 1983), bridge hands (Charness, 1979), and digit matrices (Ericsson & Chase, 1982) have all been shown to be better among subjects experienced with the class of stimuli than among inexperienced subjects.

That prejudicial attitudes might provide an alternative explanation (as can be argued about studies showing the other-race effect) does not seem to be the case. Subjects' self-reports suggest that a vast majority have quite positive attitudes regarding infants. Another alternative interpretation of the memory differential suggests that physical facial structure is less variable among infants than among adults, and that in consequence, infants are really more difficult to discriminate and to remember. Although we lack facial measurements of samples of newborn infants, analysis of facial structural measurements of 1-, 2-, and 3-year-olds reported by Herskovitz and by Miklashevskaya (both cited in Goldstein, 1979a) do not suggest that structural variability at early ages differs markedly from that of adolescents or adults. Structural differences also cannot explain the effects of experience on memory demonstrated in Experiment 1.

This investigation has provided an equivocal answer to the question of whether infant faces are perceived by adults to resemble each other more than do adult faces, or less. The process of perception comprises factors of structural differences between stimuli, factors of prior experience, and possibly factors of attitude. The outcome of the second experiment, which permitted a rating of each pair, could have been more heavily influenced by attitudes than was the outcome of the third experiment, which used forced choices.

We conclude that infant faces are more difficult for adults to recognize than are faces of older individuals. Of the possible factors examined that might account for this difference, experience with infant faces receives the most support from our results.

REFERENCES

CHANCE, J. E., GOLDSTEIN, A. G., & MCBRIDE, L. (1975). Differential experience and recognition memory for faces. *Journal of Social Psychology*, **97**, 243-253.

CHANCE, J. E., TURNER, A. L., & GOLDSTEIN, A. G. (1982). Development of differential recognition of own- and other-race faces. *Journal of Psychology*, **112**, 29-37.

CHARNESS, N. (1979). Components of skill in bridge. *Canadian Journal of Psychology*, **33**, 1-50.

CHASE, W. G., & SIMON, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing* (pp. 215-281). New York: Academic Press.

CHI, M. T. H. (1978). Knowledge structures and memory development. In R. S. Seigler (Ed.), *Children's memory* (pp. 73-96). Hillsdale, NJ: Erlbaum.

CROSS, J. F., CROSS, J., & DALY, J. (1971). Sex, race, age, and beauty as factors in recognition of faces. *Perception & Psychophysics*, **10**, 393-396.

ERICSSON, K. A., & CHASE, W. G. (1982). Exceptional memory. *American Scientist*, **70**, 607-615.

FEINMAN, S., & ENTWHISTLE, D. R. (1976). Children's ability to recognize other children's faces. *Child Development*, **47**, 506-510.

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.

GOLDSTEIN, A. G., & CHANCE, J. E. (1976). Measuring psychological similarity of faces. *Bulletin of the Psychonomic Society*, **7**, 407-408.

GOLDSTEIN, A. G., & CHANCE, J. E. (1978). Judging face similarity in own and other races. *Journal of Psychology*, **98**, 185-193.

GOLDSTEIN, A. G., & CHANCE, J. E. (1979). Do foreign faces really look alike? *Bulletin of the Psychonomic Society*, **13**, 111-113.

GOLDSTEIN, A. G., & CHANCE, J. E. (1980). Memory for faces and schema theory. *Journal of Psychology*, **105**, 47-59

GOLDSTEIN, A. G., & CHANCE, J. E. (1985). Effects of training on Japanese face recognition: Reduction of the other-race effect. *Bulletin of the Psychonomic Society*, **23**, 211-214.

KENNEY, M. (1979, July). Right name, wrong baby. *Parents*, pp. 62-65.

MALPASS, R. S. (1981). Training in face recognition. In G. Davies, H. Ellis, & J. Shepherd (Eds.), *Perceiving and remembering faces* (pp. 271-285). London: Academic Press.

Maternity ward nightmare. (1978, August 28). *Time*, p. 42.

PEECK, J., & ZWARTS, J. (1983). Recognition memory for pictures of birds in relation to bird-watching skill. *American Journal of Psychology*, **96**, 553-556.

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

SHEPHERD, J., DEREGOWSKI, K., & ELLIS, H. (1974). A cross-cultural study of recognition memory for faces. *International Journal of Psychology*, **9**, 205-212.

(Manuscript received for publication March 3, 1986.)