

Discriminability of different parts of faces

MARIANNE S. LACROCE, LEONARD BROSGOLE, and REX G. STANFORD
St. John's University, Jamaica, New York

Two face-recognition studies were run, in which different groups of subjects were presented with photographs of fully exposed faces, faces with their lower halves masked, and faces with their upper halves masked. In the first experiment, subjects were shown a series of 50 pictures of 50 different people and had to keep a mental tally of how many different people were shown. The presentation of such a series produced a confusion among faces, resulting in errors' being made in the final tallies given. The greatest number of errors (lowest tallies) was associated with masking the upper halves of the faces, and the fewest (highest tallies) occurred with no masking. Similar results were obtained in the second experiment, when subjects had to sort a set of 25 different facial stimuli into piles, with each pile supposedly containing pictures of the same person. The smallest number of piles was created when the upper halves of the faces were masked, and the largest number, with no masking. The finding that intrusion errors were associated primarily with the lower region of the face supports the proposition that faces are most distinctive in the upper half and most similar in the lower portion.

Howells (1938) conducted a face-recognition study in which he employed a masking procedure. He found that face recognition was somewhat more difficult when the lower halves of facial photographs were covered. Of course, recognition was best when the faces were fully exposed. Howells's results suggested that the features in the lower half of the face contribute more to face recognition than those in the upper half. However, his procedure was seriously flawed in several respects. A principal problem was that the test stimuli (compared to the inspection pictures) were transformed in size, clarity of focus, clothing, angle of view, and facial expression. The use of such a design produced a low level of recognition accuracy, even with fully exposed faces (62%). This fact alone would lead us to dismiss Howells's findings.

The question of what portion of the face contributes most to recognition accuracy was investigated again by Goldstein and Mackenberg (1966). They presented kindergarten, first-grade, and fifth-grade children with pictures of their classmates. The pictures were covered in 13 different ways, and the children had to name the person under each mask. Full faces were never shown. Fifth graders were correct on 100% of the trials when the left half of the face was exposed. There was over 95% accuracy when the upper half was shown and approximately 40% accuracy when the lower half was seen. Although accuracy declined overall as a function of decreasing age, the relative contribution of the different parts of the face to correct identification still occurred among the younger children.

The findings of Goldstein and Mackenberg (1966) were supported by McKelvie (1976), who studied the effect of

masking either the eyes or the mouths of inspection and test facial photographs on the recognition of unfamiliar faces by adults. McKelvie found that masking the eyes had a large detrimental effect on the level of recognition. Covering the mouths, however, produced essentially the same high level of performance as did full face exposure.

McKelvie's (1976) findings were upheld by pilot research in our own laboratory, but only with the use of a long inspection series of 24 faces.¹ Effects were not obtained with the use of only 8 inspection faces. (It is interesting that McKelvie used 27 inspection faces.) In comparison with the short inspection series, a long series generated substantially and significantly more errors. This was particularly the case when the upper halves of the faces were covered and the lower halves were exposed. Perhaps the numerosity, complexity, and contrast of feature cues in the upper region of the face contributed to the distinctiveness of the face, making it highly discriminable from other faces. Conversely, faces should be similar to one another below the midpoint of the nose, because of the fewer feature cues that are available for recognition. If that were the case, one would have expected a long inspection series to produce recognition errors primarily with the exposure of lower halves of faces, because of the intraserial interference generated by similarity. Such thinking was used to explain the data that were obtained.

In accord with the logic just specified, the present study was aimed at determining whether it is easier to differentiate between faces by using upper region feature cues or by using lower region cues.

EXPERIMENT 1

Method

Subjects. The subjects consisted of 90 unpaid volunteers (78 females, 12 males), ranging in age from 18 to 65 years. They were all university students who were recruited from graduate and undergraduate classes. Each subject was randomly assigned to one of the nine test conditions,

This research is based on a dissertation that was submitted in partial fulfillment of the doctoral requirements at St. John's University. Requests for reprints should be sent to L. Brosgole, Psychology Laboratory, St. John's University, Jamaica, NY 11439.

resulting in 10 subjects per condition. All participants signed an informed consent form.

Materials. The stimuli were full face, color photographic prints ($3\frac{1}{2} \times 5$ in. glossies) of 26 college-age females and 24 college-age males, taken against a neutral gray ground. Flash pictures of the models were taken with a Cannon AE1 camera, using a 50-mm, f 1.8 lens in combination with a $2 \times$ telextender from a distance of 54 in. Vertically oriented photographs were obtained for each model with neutral, smiling, and pouting expressions, in that sequence. For the pouting expression, the models were shown a picture of a person exhibiting a pout and asked to imitate that person's expression. The pouting picture was obtained from an article by Field, Woodson, Greenberg, and Cohen (1982). All details of clothing were eliminated by draping a neutral gray sheet over the models' shoulders. The female models' hair was pinned up, to control for hair length. The models were asked to remove all jewelry and eyeglasses, and no males with mustaches and/or beards were included. The stimuli were restricted to pictures of Caucasian students only, although there were some variations in skin tone. (As a result, all subjects were Caucasian, to avoid the potential problem of a failure to recognize faces because of racial differences.)

Three copies of each of the 50 faces for inspection were obtained for each of the three expressions (neutral, smiling, pouting) to form nine experimental albums. For all nine albums, the 50 pictures for inspection were randomly arranged, 1 on a page. The same picture arrangement was repeated for all albums. In Albums 1, 2, and 3, all the photographs were of the neutral expression. For Album 1, gray duct tape was used to mask the entire lower regions of the faces, beginning at the middle of the nose. For Album 2, the upper regions of the faces were masked with tape, beginning at the midpoint of the nose. Album 3 contained full faces with no masking. Albums 4-6 were the same as 1-3, except that they contained smiling faces. Similarly, Albums 7-9 contained pouting faces.

Procedure. All subjects were tested individually in a quiet room in a laboratory setting. The same procedure was followed for every subject. Each subject was told that he/she would be shown a series of different pictures of people and that some pictures might be shown more than once in the series. Actually, no duplicate pictures were shown. The subjects had to keep a mental count of how many different people had been shown in the series. The subjects were advised that we would stop the procedure on a couple of occasions and ask for a tally of how many different individuals they had seen up to that point. The subjects were encouraged to pay close attention to the task.

Each picture was shown for 4 sec. The first tally was requested after the first 25 pictures had been shown, and a final count was taken after all 50 pictures had been presented. All subjects indicated that they were not familiar with any of the people appearing in the pictures.

Design. Ten different subjects were tested with each album, meaning that facial expression (neutral, smiling, and pouting) and part of face (full, upper, and lower) were between-subject factors. Each subject gave two tallies. One occurred half-way through the procedure, and the second occurred upon its completion. Therefore, part of series (first half and second half) was a within-subject factor.

Results and Discussion

Every subject was shown 50 different facial photographs. Perfect performance would have required a final count of 50 by each subject. A lower count would constitute intrusion errors—that is, mistakes produced by the confusion of pictures with ones seen previously. So, for example, a final tally of 45 would indicate that a subject thought that 5 pictures had already been seen, and that subject would have made five errors. The error data were analyzed by means of a $2 \times 3 \times 3$ split plot analysis of variance (ANOVA). The analysis yielded significant main effects of part of face [$F(2,81) = 34.79, p < .01$] and of part of series [$F(2,81) = 59.35, p < .01$]. In addition, there was a significant interaction between part of face and part of series [$F(2,81) = 5.93, p < .01$]. No other significant main effects or interactions were found.

The significant main effect for part of series indicated directly that there were reliably more errors made during the second half of the series than during the first half ($p < .01$), with a mean of 35.27 mistakes in the second half and 24.36 in the first. With regard to part of face, 12.4 errors were made with full faces, 18.93 with the upper halves, and 28.3 with the lower regions. The significant main effect for part of face was further assessed with the Newman-Keuls test, because that factor contained three levels. It was found that significantly fewer errors were made when full faces were exposed than when either the upper or the lower regions were shown ($p < .01$ for each comparison). In addition, exposure of the upper halves produced reliably fewer errors than did exposure of the lower halves ($p < .01$).

The series \times part-of-face interaction was analyzed by means of simple effects, and every comparison was found to be significant. There was an average of 5.33 mistakes made with full faces in the first half of the series, and there were 7.07 in the second [$F(1,81) = 4.50, p < .05$]. The exposure of the upper regions resulted in 7.73 and 11.2 errors, respectively [$F(1,81) = 18.01, p < .01$]. With the exposure of the lower halves, 11.3 mistakes were made in the first half of the series, and 17.0 in the second [$F(1,81) = 48.69, p < .01$]. (See Table 1.) The simple effects across parts of face were followed by Newman-Keuls tests, and the part-of-face main effect was obtained for both halves of the series. The significant interaction reflected the fact that the increase in errors from the first to the second half of the series with the lower regions (50% increase) was greater than that obtained with exposure of the upper halves (45%) and full faces (33%). A further analysis, utilizing the difference scores between the two halves of the series for each part of face, indicated that the increase in errors with exposure of the lower regions was significantly greater than the increment with either the upper half or full face presentation ($p < .01$ for each comparison). In addition, upper half presentation resulted in a significantly greater increase in mistakes than did full face exposure ($p < .05$).

These data have shown that the presentation of a series of faces resulted in intrusion errors; the longer the series, the greater the number of intrusions. In addition, the greatest number of confusions was produced when only the lower halves of the faces were shown, and fewer confusions were associated with exposure of the upper halves. This supports the speculation that the upper half of the face is more distinctive, more discriminable, and hence

Table 1
Mean Number of Errors by Part of Face and Part of Series

Part of Face	Part of Series			SD
	First Half	Second Half	Total	
Full face	5.33	7.07	12.40	4.26
Upper region	7.73	11.20	18.93	4.57
Lower region	11.30	17.00	28.30	3.81
Total	24.36	35.27		

more informative than the lower half. The results also showed that information from both halves of the faces was useful in combination, in that exposure of full faces produced the smallest number of mistakes. This suggests that the upper region of the face does not contain all of the cues that are useful for face recognition.

EXPERIMENT 2

The results of Experiment 1 depended on the use of memory. Experiment 2 was an assessment of whether comparable effects would appear perceptually. Therefore, subjects were handed 25 pictures of faces and were required to sort them into piles. Each pile was to contain pictures of the same person. Perfect performance would have required that the subject create 25 piles, with one picture per pile. A smaller number of piles would indicate that there had been a failure of perceptual discrimination. Some subjects sorted a pack of fully exposed faces. Others sorted by upper region alone, and still others, by lower region. This allowed us to ascertain whether the upper region of the face was more distinctive and informative than the lower region on a perceptual level.

Method

Subjects. A total of 90 unpaid volunteers (55 females and 35 males) participated. They were all university students, who ranged in age from 18 to 65 years. They all signed consent forms.

Materials. The facial stimuli were the same as those used in the first half of the picture series in Experiment 1. As before, there were smiling, pouting, and neutral faces, with upper, lower, and full face presentation. In this experiment, the pictures were removed from the albums to create nine separate packs of photographs. All of the pictures in a given pack (or deck) contained the same expression and were masked in the same fashion. Faces of the same models appeared in each set of 25 pictures. Thirteen models were females and 12 were males.

Procedure. All subjects were tested individually in a quiet location. The same procedure was followed for every subject. The subjects were told to sort the pictures into piles for as many different individuals as possible. They were allowed to handle the sorting task in any way that they liked (e.g., by initially spreading all pictures face up on the table and then comparing them, or by selecting pictures one at a time from the deck as they made comparisons). The subjects were allowed to change their decisions at any time during the task, if they thought that they had made errors regarding the placement of pictures into particular piles. No time limit was imposed for completing the sorting task. The experimenter then counted and recorded the number of piles into which the subjects sorted the pictures. The subjects were not familiar with any of the models in the pictures.

Design. Different subjects were tested on each pack of pictures. Therefore, this was a between-subject design with only two factors. The first was part of face (upper, lower, full) and the second was expression (neutral, smiling, pouting). The sorting errors made by the subjects were analyzed statistically.

Results and Discussion

The subjects were required to sort the pictures into piles of different faces. Perfect performance would have re-

quired sorting into 25 piles, with one face in each. A smaller number of sorts would indicate confusion errors, presumably based on the similarity between pictures. These errors were analyzed in 3×3 completely randomized block factorial ANOVA. The analysis yielded a significant main effect of part of face [$F(2,81) = 7.62, p < .01$]. No other significant main effects or interactions were found.

With regard to the main effect of part of face, the subjects committed an average of 1.8 errors (7%) with full face exposure, 3.5 errors (14%) with the upper regions, and 6.33 errors (25%) with the lower halves of the faces. Newman-Keuls tests showed that the subjects were significantly more accurate in sorting the full face exposures than they were in sorting the upper halves ($p < .05$) and lower halves ($p < .01$). In addition, significantly fewer mistakes were made in sorting the upper as opposed to the lower regions ($p < .01$).

The procedure used in this experiment was very liberal. The subjects were allowed to approach the task by using any strategy that they desired and were then permitted to correct themselves on a continuing basis. Even under such circumstances, a part-of-face difference emerged. It is obvious, then, that across a population of faces, there is the potential for greater distinctiveness in the upper region and more redundancy in the lower. This is the case upon direct perceptual input, as well as memory.

One surprising outcome was consistent across studies: the factor of expression failed to emerge as a significant main effect and did not enter into any significant interactions. It was felt that adding expression to the faces might enhance the distinctiveness of the lower regions in particular, and thereby diminish a part-of-face difference. This thinking was not supported by the data.

REFERENCES

- FIELD, T. M., WOODSON, R., GREENBERG, R., & COHEN, D. (1982). Discrimination and imitation of facial expressions by neonates. *Science*, **218**, 179-181.
- GOLDSTEIN, A. G., & MACKENBERG, E. J. (1966). Recognition of human faces from isolated facial features: A developmental study. *Psychonomic Science*, **6**, 149-150.
- HOWELLS, T. H. (1938). A study of the ability to recognize faces. *Journal of Abnormal & Social Psychology*, **33**, 124-127.
- McKELVIE, S. J. (1976). The role of eyes and mouth in the memory of a face. *American Journal of Psychology*, **89**, 311-323.

NOTES

1. These data were pirated by a former student of the second author and published in a foreign journal.