

The recognition of faces and expressions

DEBRA COHEN-PAGER and LEONARD BROSGOLE
St. John's University, Jamaica, New York

The accuracy of face recognition was compared with that for expressional recognition. A series of 16 pictures of faces was presented for inspection. The series comprised 8 pictures of smiling faces and 8 pictures of neutral faces. The expressions were presented in random order in Condition 1. Condition 2 involved the block presentation of expression. In a recognition test that followed 5 days later, the levels of both facial and expressional recognition were assessed. Face recognition was significantly better than expressional recognition under both conditions. In addition, Condition 1 produced significantly better expressional recognition than did Condition 2. The theoretical implications of these findings are discussed.

Various studies have shown that features contained in the upper portion of the face are used more reliably for purposes of facial recognition and identification than lower region features are. Such studies cut across the infant, child, and adult literature.

With regard to infants, Bushnell (1982) used a habituation-recovery design to show that 5-week-old babies were capable of discriminating between photographs of their mothers and of strangers. These babies were able to discriminate between mothers and strangers by processing certain contour information, specifically, the boundary separating the hair from the face. By the 19th week of age, infants came to rely on information regarding the eyes in making such discriminations. These findings were fully consistent with the face scanning data of Hainline (1978) and Maurer and Salapatek (1976). Thus, it would appear that the features located in the upper half of the face contained all of the information that infants required for face recognition. Similar findings were obtained with young children.

Goldstein and Mackenberg (1966) tested kindergarten, first-, and fifth-grade children for recognition of familiar faces. The children were shown pictures of 13 of their classmates, whom they had to identify. However, each of the 13 pictures was masked in a different fashion. The highest recognition levels were obtained for pictures in which the eyes and hair were exposed in some combination, and low recognition levels were gotten with the exposure of lower region feature combinations. There have been similar findings for adults.

McKelvie (1976) used a masking procedure with adults to assess their ability to recognize pictures of previously exposed unfamiliar faces. He found that whereas masking the eyes seriously impaired face recognition, masking the mouth had virtually no detrimental effect. Although it is clear in the literature that subjects perform accurately on

face recognition tasks by using features located in the upper half of the face, a question emerges with respect to the extent to which the features in the lower half of the face are registered in memory during the inspection phase of a recognition design. Even if those features (e.g., the mouth) are registered in memory, a further question may be raised with respect to whether the facial expressions created by them are encoded in memory as well. The latter question was addressed directly in an often cited experiment by Galper and Hochberg (1971), who compared expressional and facial memory.

Galper and Hochberg (1971) presented subjects with an inspection series of 24 slides. The slides consisted of 16 faces in photographic positive and 8 faces in photographic negative. The positive slides contained facial expressions ranging from a resting state (neutral) to a full smile. The slides were exposed one at a time, in a semi-randomized order. The inspection series was followed by a recognition test 5 days later.

The recognition test consisted of presenting subjects with pairs of pictures. One member of each pair was taken from the inspection series, and one was an unfamiliar picture. The subject had to indicate which picture had been shown previously. The recognition errors made to the negatives were used to establish an empirical chance level of recognition. The positive slides were used to test for facial and expressional memory.

The level of facial recognition was assessed over trials on which familiar faces (i.e., ones from the inspection series) were paired with unfamiliar faces. The degree of expressional recognition was determined over trials; each trial consisted of the presentation of a familiar picture paired with a new picture of the same person with the opposite expression. Galper and Hochberg (1971) found that expressional memory was significantly better than chance, but significantly worse than facial memory. From this, they concluded that "the characteristics which differ in resting state and smile play some role in face recognition" (p. 354).

The present study was essentially a replication of the Galper and Hochberg (1971) study, in that 16 pictures

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.

were presented in photographic positive and later followed by a recognition test. Just as in Galper and Hochberg, recognition was tested with the use of two types of tasks for each subject. In one task, the subjects were presented with a familiar and an unfamiliar face, and in the other task, the subjects were presented with a familiar and an unfamiliar expression on a familiar face.

This experiment was aimed at determining whether the results obtained by Galper and Hochberg (1971) could be attributed to a potentially serious methodological flaw—namely, the randomization of expressions in the inspection phase. Such randomization may have had the net effect of transmitting the irrelevance of expression, resulting in a decreased level of expressional recognition as opposed to facial recognition. In order to assess whether their results were due to the randomization of expressions in the inspection series, two conditions were run in which the distribution of the expressions on the faces was manipulated systematically.

METHOD

Subjects

Twelve male and 28 female undergraduates were tested. They were naive with respect to the purpose of the study. The subjects ranged in age from 17 to 50, their average age being 21. Each subject was assigned randomly to one of two groups, resulting in 20 subjects per group. All participants gave informed consent by signing an agreement and were treated in accordance with the ethical principles of the American Psychological Association.

Materials

The stimuli consisted of 35-mm color prints (3.5×5 in.) of 96 college-age males and females. Equal numbers of photographs of males and females were used in each condition. The photographs were taken with a 50-mm lens in conjunction with a telexender from a distance of 4 ft. Thus, each print contained only the head and shoulders of the model being photographed. These stimuli were used previously by LaCroce (1985) and Drago (1985).

Each model was photographed twice, once with a neutral and once with a smiling expression. All details of clothing were eliminated by draping a neutral gray cloth over the models' shoulders. All of the models removed their jewelry and/or eye glasses, and no males with moustaches and/or beards were included. In order to control for the women's hair length, their hair was pinned up. Background details were eliminated by photographing the models against a neutral gray projection screen.

Procedure

The subjects viewed a series of 16 photographs in the inspection phase. The pictures were exposed for 7 sec with a 3-sec interval between exposures. The subjects were told to "look at each picture with the aim of later being able to remember it." The expressions on the faces viewed in the inspection phase were varied systematically by condition. Five days later, the subjects viewed 16 pairs of pictures and were asked to "point to the picture that most resembles the one you saw the other day." The sequence of the presentations of the test series corresponded to the sequence of the pictures viewed in the inspection phase.

In the test series, the subjects viewed a series of 16 pairs of photographs, with one in each pair corresponding to the 16 photographs viewed in the inspection phase. Two types of pairs were used in the test phase. These types were randomly varied, without either type appearing for more than three consecutive test trials. One pair consisted of two photographs of the same face as the original target face, with one smiling and one neutral. The other pair consisted of the original target face and a new face, both of whom had the same expression as did the original target face.

In Condition 1, the subjects viewed 16 inspection faces with the two expressions randomized over trials. This condition paralleled Galper and

Hochberg's (1971) design. In Condition 2, half of the subjects viewed 8 neutral faces followed by a block of 8 smiling faces. The remaining subjects viewed a block of 8 smiling faces followed by 8 neutral faces in the inspection phase. The nonrandomization of expression in the inspection phase of Condition 2 was expected to promote a sense that expression should be taken into account for the recognition test that was to follow. Therefore, it was expected that Condition 2 would produce a significantly higher level of expressional recognition than would Condition 1. Additionally, it was entirely possible that there would be no difference between face and expressional recognition in the latter condition.

RESULTS

The number of facial and expressional errors were recorded for each subject under each condition. The data were subjected to a 2×2×2×2 split-plot analysis of variance (ANOVA). There were two between- and two within-subject factors. The between-subject factors were conditions (random vs. block presentation of expression), and type of first test trial (expressional vs. facial). The first within-subject factor was the type of recognition test (expressional vs. facial). The second factor was the type of inspection face (smiling vs. neutral). The ANOVA revealed three significant main effects and one significant two-way interaction. No other significant differences were found. Table 1 gives the cell means.

The first significant main effect related to conditions [$F(1,36) = 4.43, p < .05$]. Forty-nine errors were made in Condition 1, and 71 errors in Condition 2. Thus, significantly fewer errors were made with random presentation of expression in the inspection series than with block presentation.

The second main effect related to type of test [$F(1,36) = 23.28, p < .01$]. As can be seen in Table 1, 84 expressional errors were made, in comparison with 36 facial errors. Therefore, there were significantly more expressional errors committed than facial errors.

The third main effect related to type of inspection face [$F(1,36) = 4.39, p < .05$]. Overall, more errors were committed in response to smiling faces as opposed to neutral faces (71 vs. 49).

The significant two-way interaction occurred between conditions and type of recognition test [$F(1,36) = 4.40, p < .01$]. To evaluate this significant interaction further, simple effects were run first across type of recognition test for each condition separately. The simple effects showed that there were significantly more expressional than facial errors in both conditions ($p < .01$ in each instance). Then, simple effects were run across conditions for each type

Table 1
Mean Number of Expressional and Facial Errors Committed in Response to the Random and Block Presentation of Expression in the Inspection Series

Presentation	Type of Error		Total
	Expressional	Facial	
Random (Condition 1)	31	18	49
Block (Condition 2)	53	18	71
Total	84	36	120

of recognition test separately. They indicated that the conditions differed from one another reliably with regard to expressional errors only [$F(1,36) = 23.28, p < .01$]. Significantly more expressional errors were committed in Condition 2 (block) than in Condition 1 (random). This difference between conditions can be observed in Table 1. Interestingly enough, the two conditions did not differ from one another at all in the number of facial errors made.

DISCUSSION

This study demonstrated clearly that expressional recognition was substantially inferior to facial recognition. These results both confirmed the findings of Galper and Hochberg (1971) and demonstrated that those prior findings were quite robust. Contrary to original expectation, the difference between expressional and facial recognition was enhanced by the block presentation of expression in the inspection phase of the study (Condition 2). This finding was quite surprising, but it may have been due to a decrease in arousal level produced by grouping the expressions together. Of course, an argument against this thinking is that a lowering of arousal should have negatively affected the level of face recognition as well. As indicated previously, this was not the case. However, it could be reasoned that perhaps manipulation of arousal level would influence expressional recognition alone, since expressional recognition is so much more difficult than facial recognition. Unfortunately, such reasoning brings us full circle, back to the intent of this study—namely, that of explaining the difference found between the levels of facial and expressional recognition. Part of the robustness of the Galper and Hochberg study is that their findings survived not only the main manipulation used in the present experiment, but also a variation in the type of first test trial given.

It was assumed that the very first recognition trial would set the tone for the trials to follow. Accordingly, it was reasoned that a set to recognize expressions would be established if the first test trial was an expressional one, and that expressional recognition might be better than facial recognition under such circumstances. Of course, the opposite would hold if the first test trial were a facial one. Galper and Hochberg (1971) were not clear about whether or not the type of first trial given was varied over subjects. It is entirely possible that their results may have been an artifact of their subjects' encountering a facial trial as the very first one in the test session. Such a possibility was ruled out in the present study, for facial recognition was superior to expressional recognition, regardless of the type of first test trial.

Galper and Hochberg (1971) did not evaluate their data to determine whether their findings were due to a failure of subjects to recognize one expression only. The results of the present study indicated that neutral faces were recognized significantly better than smiling faces (perhaps because subjects have adapted to seeing smiling faces in photographs). Regardless of the overall difference in recognition level, however, face recognition was better than expressional recognition for both types of faces. This finding added even further strength to the findings obtained previously by Galper and Hochberg.

Galper and Hochberg's (1971) results are solid, undiminished by a host of experimental manipulations. Yet a problem remains—namely, that no hint was given about the underlying process. That is, no substantive reasons have been advanced to explain the relative inferiority of expressional recognition. Galper and Hochberg's conclusion that expression plays some role in face recognition fails to foster an understanding of the difference between facial and expressional recognition. In fact, that conclusion did not even derive from their data. They simply compared facial and expressional recognition. The role that expression plays in face recognition can be better addressed through a transformation study, such as the type run by Bruce (1982).

Different reasons can be offered to explain the difference between expressional and facial recognition. First, it could be contended that

parts (e.g., the mouth) are not remembered as well as wholes (e.g., the entire face), and that expression is communicated principally by variations in just one feature (the mouth) of the entire face. From this point of view, a test for expressional recognition would tap the memory for an isolated part of an overall configuration. This line of reasoning does not make much sense, however, from a configurational standpoint. If the face has certain emergent properties that are contingent on an interaction between its parts, variations in expression should give rise to alterations in the overall facial configuration. This means that expressional and facial recognition should be equivalent. In fact, Asch, Ceraso, and Heimer (1960) demonstrated that both parts and wholes are remembered equally. The fact that expressional recognition is poorer than facial recognition would argue against a configurational interpretation of face perception. It would seem to favor a piecemeal understanding of how faces are seen (Carey & Diamond, 1977; Diamond & Carey, 1977).

The piecemeal explanation of face perception essentially comes out of structuralism and can be applied generally to the perception of a geometric form. From this point of view, the sensing of a face is built up from a limited number of invariant isolated features. These invariant features are contained in the upper region of the face and include the eyes, eyebrows, hairline, and hair style. Because they are unchanging, these features lead most reliably to face identification. The features that are less reliable (e.g., the mouth) need not be incorporated into the facial percept. From this point of view, then, facial recognition should be outstanding and expressional recognition dismally poor. The fact that expressional recognition is above chance may be purely incidental. Such reasoning seems to be in full accord with the manner in which Galper and Hochberg (1971) explained their data.

REFERENCES

- ASCH, S. E., CERASO, J., & HEIMER, W. (1960). Perceptual conditions of association. *Psychological Monographs: General & Applied*, *74* (3, Whole No. 490).
- BRUCE, V. (1982). Changing faces: Visual and non-visual coding processes in face recognition. *British Journal of Psychology*, *73*, 105-116.
- BUSHNELL, I. W. R. (1982). Discrimination of faces by young infants. *Journal of Experimental Child Psychology*, *33*, 298-308.
- CAREY, R., & DIAMOND, S. (1977). From piecemeal to configurational representation of faces. *Science*, *195*, 312-314.
- DIAMOND, R., & CAREY, S. (1977). Developmental changes in the representation of faces. *Journal of Experimental Child Psychology*, *23*, 1-22.
- DRAGO, R. J. (1985). *The effect of transformations of expression on the recognition of full and part faces*. Unpublished doctoral dissertation, St. John's University, Queens, NY.
- GALPER, R. E., & HOCHBERG, J. (1971). Recognition memory for photographs of faces. *American Journal of Psychology*, *84*, 351-354.
- GOLDSTEIN, A. G., & MACKENBERG, E. J. (1966). Recognition of human faces from isolated facial features: A developmental study. *Psychonomic Science*, *6*, 149-150.
- HAINLINE, L. (1978). Developmental changes in visual scanning of a face and nonface patterns by infants. *Journal of Experimental Child Psychology*, *25*, 90-115.
- LACROCE, M. (1985). *An assessment of the discriminability of the different parts of face and facial expressions by the use of two sorting techniques*. Unpublished doctoral dissertation, St. John's University, Queens, NY.
- MAURER, D., & SALAPATEK, P. (1976). Developmental changes in the scanning of faces. *Child Development*, *47*, 523-527.
- McKELVIE, S. J. (1976). The role of eyes and mouth in the memory of a face. *American Journal of Psychology*, *89*, 311-323.