

Facial inversion effects: Parts and whole relationship

SAM S. RAKOVER and BROSH TEUCHER
Haifa University, Haifa, Israel

“Facial inversion effects” refers to the findings that recognition of inverted faces is less accurate than recognition of upright faces. We now report inversion effects for isolated facial features: forehead, eyes, nose, mouth, and chin. This shows that configurational information extracted from a whole face (i.e., from spatial relationships among the facial features) is not necessary for obtaining the inversion effects. Other factors, such as “upright-orientation,” mental rotation, and feature saliency, account for the inversion effects both in a whole face and in its isolated features. We propose a simple formula that satisfactorily predicts the recognition of a whole face and the inversion effects for that face on the basis of its individual features.

“Facial inversion effects” refers to the findings that recognition of inverted faces is less accurate than recognition of upright faces. Moreover, the inversion of faces impairs recognition more than does inversion of such other objects as houses and landscapes (for a review, see Valentine, 1988). However, since the purpose of the present study was to test whether facial inversion effects were caused by the processing of individual features versus a whole face, we will not consider, in this paper, the effects of inversion on objects other than faces.

There are several hypotheses that attempt to explain these effects (e.g., see Valentine, 1988). One important hypothesis, which we shall call the “configurational” hypothesis, is based on the distinction between featural and configurational information that can be extracted from a face. The first is about the individual features of a face (e.g., saliency of the eyes), and the second is about the spatial relations among these individual features. The inversion of a face disrupts the second kind of information (the configurational) and forces one to use the first type of information (the featural) in attempting to recognize a face (see, e.g., Carey & Diamond, 1977; Diamond & Carey, 1986; Rhodes, 1995; Rhodes, Brake, & Atkinson, 1993; Rock, 1973; Sergent, 1984). In accordance with Bruce (1988), we “will use the term ‘feature’ in its every day sense to mean a discrete component part of a face such as a nose or a chin, whereas the term ‘configuration’ to mean the spatial interrelationship of facial features” (p. 38).

If facial inversion effects are caused by the destruction of configurational information and limiting one to the

use of featural information only, then the presentation of isolated inverted features, such as forehead, eyes, and nose, should not result in a reduction of recognition, because there is no configurational information to be extracted from an isolated feature either in the upright or inverted position. Configurational information is extracted from a whole face and cannot be extracted from an isolated feature.

Similarly, Rhodes et al. (1993) tested the configurational hypothesis by presenting eyes or mouth as isolated features without being displayed in a face. They argued that if the inversion effects were caused by spatial relations among facial features, then these effects should be reduced considerably when these features were presented alone (see also Rhodes, 1995).

There are two technical differences between the research of Rhodes et al. (1993) and the present study. First, Rhodes et al. tested the configurational hypothesis by presenting two facial features alone (i.e., the eyes and the mouth), whereas we presented all five major features of a face—hair and forehead, eyes, nose, mouth, and chin—alone. These features are naturally well discriminated, and are central for face reconstruction in the Identikit and the Photofit system. Furthermore, there is experimental evidence suggesting that cells in the monkey brain are particularly sensitive to regions of the hair, the eyes, and the mouth (e.g., Perrett et al., 1988). Second, Rhodes et al. used digitized versions of full-face photos that looked like drawings, whereas we used photographs of full faces and their features generated from the Photofit system (e.g., Penry, 1971a, 1971b). It seems that the latter technique approximates reality a bit better than the former one.

The configurational hypothesis contrasts with the supposition that the inversion effects are caused by other factors: “upright orientation,” mental rotation, and feature saliency. Given a yes/no recognition task consisting of study and test stages, according to the upright-orientation hypothesis, a general schema for processing faces and their features in the upright orientation develops as a con-

This research was supported by the United States–Israel Binational Science Foundation (Grant 89-00254). The present experiments consisted of Brosh Teucher’s Master’s thesis, supervised by the first author. We would like to express our gratitude to Micah Leshem, Ilan Fisher, and Ramzi Suleiman, who read an earlier draft of the present paper and made helpful suggestions. Requests for reprints should be addressed to S. S. Rakover, Department of Psychology, Haifa University, Haifa, Israel 31905 (e-mail: rps742@haifauvm.bitnet).

sequence of exposure from birth to a huge number of faces and their features viewed upright (see, e.g., Bruce, 1988; Goldstein & Chance, 1980; Rock, 1973, 1974; Valentine, 1991).

The ability to recognize faces improves with age and is related to the increase in the efficiency of a schema for processing facial information. There are some indications that the development of face recognition is from featural strategy, which dominates childhood, to configurational strategy, which prevails in adulthood. In line with this, there is evidence that children show weaker inversion effects than adults (for reviews, see Chung & Thomson, 1995; Johnston & Ellis, 1995).

Hence, a straightforward prediction of this hypothesis is that an isolated facial feature presented to adult subjects will be better recognized in the upright orientation in the study and test stages of a recognition task (i.e., upright–upright or UU) than they will in the inverted–inverted (II) condition.

The mental-rotation hypothesis deals with the upright–inverted (UI) and the inverted–upright (IU) conditions. Since, in these conditions, features are presented in the study and in the test stages in opposite orientations, subjects have to mentally rotate one feature in order to match it to the orientation of the other feature. Mental rotation describes a cognitive operation in which a visual figure is rotated in the mind to a desirable direction. For example, Valentine and Bruce (1988) have presented target faces in an upright orientation and test faces in different degrees of orientation. They found that the subjects' response time to judge whether or not a test face was the same as a target face increased linearly with degree of mental rotation required to align the faces. This mental operation causes errors in recognition (see, e.g., Rock, 1973; Shepard & Metzler, 1971; Valentine & Bruce, 1988; Yin, 1969) and in predictions that recognition will be higher in the UU and II conditions than in the UI and IU conditions.

According to the feature-saliency hypothesis, since facial configuration information cannot be extracted from an isolated feature, the visual information of an individual feature has to be considered. One such important factor is the saliency of facial features. Different features of a face are perceived and remembered to different degrees, where the general order of saliency is forehead > eyes > nose > mouth > chin (see, e.g., Davies, Ellis, & Shepherd, 1977; Haig, 1984, 1986; Rakover & Cahlon, 1989; Shepherd, Davies, & Ellis, 1981). While there have been many experiments showing that face recognition improves with age, there seems to be no convincing evidence to support the hypothesis that this improvement is due to the use of different facial features by children and adults (see, e.g., Chung & Thomson, 1995). We propose that inversion of isolated features interacts with their saliency. However, the nature of this interaction is hard to predict. Thus, for example, it has been found that, in general, inversion does not change the order of feature saliency presented in the upright orientation (e.g., Bruyer & Coget, 1987; Endo,

1986). However, Phillips (1979) has found that recognition of inverted internal features (i.e., eyes, nose, mouth) is more impaired than are inverted external features (i.e., forehead, ears and chin) (see also Ellis, Shepherd, & Davies, 1979).

In view of the above, to measure the effects of inversion on face recognition, one can use three approaches: one can compare UU with UI, UU with II, and UU with IU. Although no experiment (other than Yin's, 1969) has used all three approaches, a review of the literature reveals that researchers have used either the first approach (UU vs. UI; e.g., Rhodes, Brake, & Taylor, 1989; Scapinello & Yarmey, 1970; Valentine & Bruce, 1988; Yarmey, 1971) or the second (UU vs. II; e.g., Diamond & Carey, 1986; Rhodes et al., 1993). Apparently, no one has used the third (UU vs. IU).

Which of these approaches is most appropriate for testing of the configurational hypothesis? The answer depends on the interaction between the cognitive processes involved in recognizing an inverted face, such as encoding and retrieval processes, and the specific demands of the task to be performed. If one proposes that inversion is associated with the encoding process, then one would use the UU versus II or the UU versus IU comparisons. However, both of these comparisons are also based on other processes. The former comparison deals with a presentation of the same inverted face in both the study and test stages. Hence, the inversion effects can be attributed to encoding and/or retrieval processes. The comparison of UU versus IU, while discriminating between the study and test stages, requires mental rotation. If one proposes that inversion is associated with the retrieval process, then one would use the UU versus UI comparison. However, this comparison involves mental rotation.

It seems that there is no recognition task that does not involve several complex cognitive processes. Consider, for example, a simple straightforward procedure for testing the inversion effects, where one has to identify by name faces presented upright or inverted of known persons. Even this procedure involves several cognitive processes, such as the comparison between the presented faces and the upright remembered faces—a process that demands mental rotation when inverted faces are presented.

Given the above, we decided to analyze our results by utilizing all three approaches, with special attention being paid to the UU versus II and the UU versus UI comparisons.

In sum, we proposed that several factors might be involved in the facial inversion effects: featural information (e.g., feature saliency), configurational information, upright orientation, and mental rotation. In Experiment 1, we attempted to test whether configurational information was crucial by finding out if there were inversion effects of isolated features. Assuming such effects, the next question was whether the inversion effects of a whole face could be accounted for by the inversion effects of its isolated features. In Experiment 2, we intended to answer this question by testing the recognition of a whole face

composed of the features used in Experiment 1, and in the same experimental conditions.

EXPERIMENT 1

Method

Subjects. Eighty students in the Department of Psychology at Haifa University volunteered to participate in the experiment. Their average age was 22.2 years. Seventy-two were female and 8 were male. The subjects were assigned randomly to the experimental groups and were tested in groups of 4.

Stimuli and Apparatus. A face is composed of values (e.g., small *mouth*, blue *eyes*), each of which belongs to a different facial dimension or feature (e.g., *mouth*, *eyes*). In the examples in parentheses, the italicized words stand for facial dimensions or features, which, in conjunction with their adjectives, represent their values. We used Penry's Photofit kit, where a face is divided into five major, well-discriminated, and natural facial dimensions and each dimension consists of many different values (Penry, 1971a, 1971b). The Photofit is a popular technique used by the police and in many laboratory experiments. In comparison with the other popular technique, called the Identikit, which uses line-drawn features, the Photofit uses photographs of human faces (see, e.g., Bruce, 1988; Davies, 1981, 1986). Twenty different values of each of the facial dimensions: hair and forehead, eyes and eyebrows, nose, mouth, and chin were randomly selected from Penry's Photofit kit. All were black and white pictures of a male Caucasian face. Ten values of each facial dimension were chosen to be study values, and 10 were chosen as distractors.¹

Twenty video cassettes were prepared, each presenting 30 values in two successive stages. In the study stage, 10 study values of a facial dimension were presented in random order. In the test stage, 20 test values were presented in random order. Ten test values had been presented in the previous stage, and 10 were new distractor values. Thus, each cassette presented 30 values of only one facial dimension. There were five facial dimensions, two orientations of values presented in the study stage (upright or inverted), and two orientations in the test stage (upright or inverted), resulting in 20 different cassettes (i.e., $5 \times 2 \times 2 = 20$). Each study value was presented for 15 sec, and each test value was presented for 10 sec. (These timed intervals, which had been determined on the basis of a pilot study, allowed the subjects enough time to perceive the study values and respond to the test values as natural parts of a human face. Also note that, in comparison with 6 and 24 sec, a 15-sec exposure of a face results in medium recognition—see Read, Vokey, & Hammersly, 1990.) The values were presented on a black-and-white 24-in. TV screen.

Design and Procedure. Three independent variables were manipulated in a $2 \times 2 \times 5$ factorial design with repeated measurements on the last variable. The first variable, study orientation, presents values either upright or inverted. The second variable, test orientation, also presents values upright or inverted. The third variable, facial dimension, consists of forehead, eyes, nose, mouth, and chin. Since subjects viewed all facial dimensions in succession, the groups are differentiated by the first two variables. Subjects in the upright-upright (UU) group ($n = 20$) were asked to view the study values and to decide for each test value whether it was old or new, that is, whether or not the test value had been exposed in the study stage. They viewed all of the five facial dimensions in succession. They were informed that the values would be presented in the upright orientation in the study stage and in the test stage, that is, the UU sequence. The order of presentation of the facial dimensions was counterbalanced between subjects. Subjects in the upright-inverted (UI) group ($n = 20$) were given the same instructions as the UU group but were informed that the values would be presented in UI sequence. Subjects in the inverted-upright (IU) group ($n = 20$) were similarly instructed but were informed that the values would

be presented in IU sequence. Subjects in the inverted-inverted (II) group ($n = 20$) were informed that the values would be presented in II sequence.²

Results

A 2 (study orientation) $\times 2$ (test orientation) $\times 5$ (facial dimension) analysis of variance (ANOVA) with repeated measurements on the last variable was applied to each of the following measures: correct choices (V), hits, A' (a sensitivity measure), and B'' (a response-bias measure; see, e.g., Macmillan & Creelman, 1991). However, since no significant differences were found with B'' , and since the results of hits were similar to V and the results of A' were very close to V , we shall report only the results of V . These measures do not add significant information to the analysis of V . (Furthermore, note that similar results and relations among these measures were obtained in Experiment 2.)

Figure 1 depicts V and P (change) as a function of study orientation, test orientation, and facial dimension. (Note that facial dimensions are ordered from top to bottom of a face.) P (change) = $(UU - X)/UU$, where UU represents V of the UU group, X represents V of the UI, II, or IU group. As can be seen from Figure 1, V is higher when the values presented in the study stage were upright rather than inverted [$F(1,76) = 5.94 p < .05$]. The interaction between study and test orientation shows that V is higher when values are presented in the same orientation in the study and test stages (i.e., for the UU and II groups) than when they are presented in different orientations (i.e., UI and IU groups) [$F(1,76) = 64.01 p < .01$]. Subsequent Newman-Keuls tests at $\alpha = .05$ revealed that V in the UU group (averaged over the five facial dimensions) was greater than V s in the II, UI, and IU groups, and that V in the II group was greater than V s in the UI and IU groups. V shows the following order for the facial dimensions; forehead > eyes > nose > chin > mouth [$F(4,304) = 52.71 p < .01$]. The interaction between study orientation, test orientation, and facial dimension shows that, in comparison with that of the UU group, the V of the internal values (eyes, nose, and mouth) decreased in the UI and the IU groups more than did V of the external values (forehead and chin), whereas V for high salient values decreased in the II group more than did V for low salient values [$F(4,304) = 4.23 p < .01$]. (These relations are presented in the inset panel of Figure 1.)

Subsequent planned comparisons with the Newman-Keuls test at $\alpha = .05$ revealed the following significant differences for each facial dimension (see Table 1).

These results are substantiated by the comparisons made between the UU and II and the UU and UI groups. A 2 (UU vs. II) $\times 5$ (facial dimension) ANOVA with repeated measurements on the last variable reveals that V is higher in the UU group than in the II group [$F(1,38) = 10.29, p < .1$]. V shows the following order for the facial dimensions: forehead > eyes > nose > mouth > chin [$F(4,152) = 23.76, p < .01$]. Subsequent planned comparisons with the Newman-Keuls test at $\alpha = .05$ reveals that V s in the UU group are higher than V s in the II group

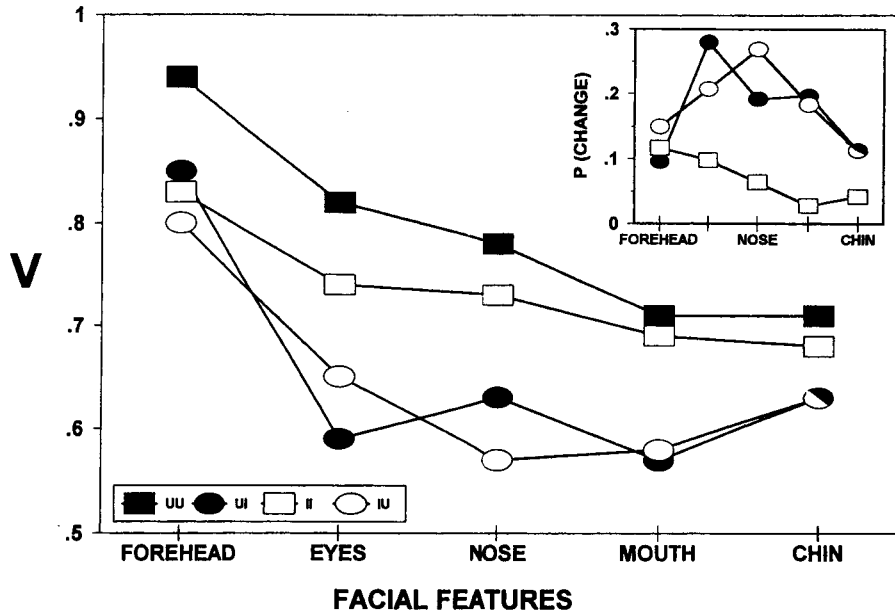


Figure 1. Probability of correct choices (*V*) and *P* (change) as a function of the study orientation, test orientation, and facial dimension.

for the forehead and eyes. (Note that the difference in *V* of the eyes between the UU and the II groups in Table 1 is significant at $\alpha = .0526$, because the MS_e of this analysis is increased due to inclusion of the UI and the IU groups.)

A 2 (group: UU vs. UI) \times 5 (facial dimension) ANOVA reveals that *V* in the UU group is higher than the *V* in the UI group [$F(1,38) = 67.62, p < .01$]. *V* shows the following order for the facial dimensions: forehead > (eyes = nose) > chin > mouth [$F(4,152) = 37.63 p < .01$]. The interaction between groups and facial dimension shows that in comparison with the UU group, *V* of the forehead in the UI group was much higher than it was for the other four facial features [$F(4,152) = 3.34 p < .05$]. Planned

comparisons with the Newman-Keuls test at $\alpha = .05$ reveals that *V*s in the UU group were higher than *V*s in the UI group for each facial dimension.

A 3 (group: II, UI, IU) \times 4 (facial dimension) ANOVA was applied to *P* (change) after it had been calculated for each subject. *P* (change) is higher in the UI and IU groups than in the II group [$F(2,57) = 11.91 p < .01$]. *P* (change) displays an inverted U-shaped function of facial dimension [$F(4,228) = 5.30 p < .01$]. *P* (change) of the internal features in the UI and the IU groups is higher than *P* (change) of the external features, whereas *P* (change) in the II group tends to decrease as a function of the decrease in feature saliency [$F(8,228) = 2.61 p < .05$]. Planned comparisons with the Newman-Keuls test at $\alpha = .05$ reveals that only *P* (change) of the internal features is higher in the UI and IU groups than in the II group.

Table 1
Correct Choices (*V*) as a Function of Group and Facial Dimension

Dimension	Group			
	UU	II	UI	IU
Forehead	.94	.83	.85	.80
		UU	UU	UU
Eyes	.82	.74	.59	.65
		UU*	UU	UU
Nose	.78	.73	.63	.57
			UU	UU
Mouth	.71	.69	.57	.58
			UU	UU
Chin	.71	.68	.63	.63
			UU	UU

Note—UU and/or II in the columns indicate that their *V*s are significantly higher than the *V* below which they appear, Neuman-Keuls test at $\alpha = .05$. * $p < .0526$.

Discussion

The main findings of Experiment 1 are as follows. First, *V* (correct choices) in the UU group (averaged over the five facial dimensions) is higher than *V* in the II, UI, and IU groups. Second, *V* in the II group (averaged over the five dimensions) is higher than *V* in the UI and IU groups. Third, *V* in the UU group is higher than *V* in the II group for the forehead and the eyes. Fourth, *V* in the UU group is higher than *V* in the UI group or in the IU group for each facial dimension. Fifth, *P* (change) of the internal features in the UI and IU groups is higher than *P* (change) of the external features, whereas *P* (change) in the II group tends to decrease as a function of the decrease in feature saliency.

If we accept the proposition that the spatial relations among the five facial dimensions generate the configu-

rational information of a face (i.e., second-order relational properties), then Carey and Diamond's (1977) and Diamond and Carey's (1986) configurational hypothesis is disconfirmed. Whatever the comparison between the UU group and each of the other three groups (II, UI, IU), the results present a severe problem for this hypothesis.

The present results can be accounted for by upright orientation, mental rotation, and feature saliency. The upright-orientation hypothesis predicts that V will be higher in the UU group than in the II group. The mental-rotation hypothesis predicts that V will be higher in the UU and the II groups than in the UI and the IU groups.

According to the feature-saliency hypothesis, the effects of inversion depend on the saliency of facial features. As can be seen from P (change) in Figure 1, and in accordance with Phillips (1979), inversion of the internal features (i.e., eyes, nose, and mouth) produced lower V in the UI and the IU groups than did inversion of the external features (i.e., forehead and chin). If, however, one compares V in the UU and II groups, one finds that the effects of inversion decrease as a function of the decrease in feature saliency. One plausible explanation for this difference in V or P (change) between the UI and IU groups, on the one hand, and the II group, on the other, is that in the former groups the subjects performed mental rotation. Note that there is no significant difference in V between the UI and IU groups. This finding can be interpreted as indicating that the same mental process is operating in both cases.

Although unfamiliar faces seem to be processed differently from familiar faces, inversion effects have been found in both (e.g., Bruce, 1988; Stevenage, 1995; Valentine, 1988). However, the present experiment raises special questions regarding the relation between unfamiliarity and inversion, since isolated facial features have been used here. For example, how do subjects handle these individual features? Do they think of them as parts of a face or as peculiar visual forms? Do subjects discriminate among these isolated features in the same way as they do among different faces?

We believe, for the following reasons, that isolated features are perceived as parts of a face: (1) We used five features that were naturally well discriminated. (2) We exposed the features for fixed timed intervals to allow enough time for the processing of these features as natural parts of a human face. (3) The order of recognition obtained in the UU group for isolated facial features (see Figure 1) is very close to the order of saliency obtained in the literature for a whole face. (4) The findings that there is no response bias in the data and that the different measures employed (i.e., hits, A' and V) are consistent with each other, indicate that the subjects did not use cognitive processes in addition to those proposed here. Had the subjects used many different processes, one would expect divergent decision criteria and high variability, which would have caused inconsistency in these measures and response bias.

Finally, it should be mentioned that recognition of facial features cannot be accounted for in terms of their size,

since their order of size (forehead > chin > eyes > nose > mouth) does not correspond to their recognition order.

EXPERIMENT 2

In Experiment 1, facial inversion effects were obtained for isolated facial dimensions. This raised the question of the relation between the inversion of isolated features of a face and the inversion of a whole face. We therefore replicated Experiment 1, except that a whole face was presented instead of its isolated features. A whole face was composed of the facial features used in Experiment 1.

A major aim of Experiment 2 was to show that the recognition of a whole face in each of the conditions—UU, UI, II, and IU—was based on the recognition of the five isolated facial features. If we could show this, then the inversion effects of the whole face would have resulted from recognition of the isolated features. Suppose that we used a function, F , which predicted the recognition of a whole face on the basis of recognition of the five isolated facial features, so that $W_p = F(V_f, V_e, V_n, V_m, V_c)$, where W_p stands for the predicted recognition of a whole face, f for the forehead, e for the eyes, n for the nose, m for the mouth, and c for the chin. If, for example, W_p was .91 in the UU group, .76 in the II group, and .66 in the UI group, then the predicted inversion effects would be $.91 - .76 = .15$ and $.91 - .66 = .25$. If we were to find that these figures equaled the obtained figures, so that the obtained recognition of a whole face (W_o) in the UU group was about .91, and so forth, then the inversion effects of the whole face would be completely predicted by recognition of its isolated features.

Rakover (1994) proposed the following general constraints for developing a mathematical formula that relates the recognition of a whole face to the recognition of its isolated features:

1. A face is a mono-oriented visual form composed of well-discriminated features such as forehead, eyes, nose, mouth, and chin. The recognition of a whole face is based on the recognition of its features.

2. The recognition of the whole face and its isolated features has to be measured by the same experimental method.

3. $V_L \leq W_p \leq V_H$, where V_L denotes recognition of the lowest V of the facial features and V_H denotes recognition of the highest V .

4. The formula $W_p = F(V)$ will not include free parameters.

These constraints are based on the following rationale:

1. We view a face as a mono-oriented visual form composed of well-discriminated features, such as eyes and nose. If the features of the face are not well discriminated, we believe that the proposed W_p formula (see below) would fail to predict recognition of the whole face. This is indirectly supported by the finding that a well-discriminated part in a visual stimulus is better recognized as belonging to that stimulus than is a poorly discriminated part (e.g., Bower & Glass, 1976; Palmer, 1977). Biederman (1987) proposed that if part of the components

of an object were identified, identification of the object would be accurate.

2. Since different methods of facial recognition may result in different measurements, we require that a face and its isolable features be measured by the same recognition task. Thus, an experiment that tests recognition of a whole face following the presentation of its isolated feature(s) poses questions (and will yield results) different from those of the present experiment.

3. Let us assume that one invests cognitive effort in recognizing a whole face composed of five isolated features. We propose that the cognitive effort in processing a whole face might be distributed among the five features presented simultaneously in a face, so that: $V_L \leq W_p \leq V_H$. Thus, compared with V_H , W_p is decreased by V_s lower than V_H , and compared with V_L , W_p is increased by V_s higher than V_L .

Assuming that W_p predicts W_o , one can test empirically the inequality that W_o falls between V_H and V_L .

4. To avoid the strong association between parameter estimations and a specific experimental situation, we developed a formula that connects W_p with V without using free parameters.

Given the above constraints, Rakover (1994) proposes a variety of formulas for $W_p = F(V_i)$. One of these formulas is based on Osgood, Suci, and Tannenbaum's (1957) formula for computing the meaning (e.g., attitude) of a compound stimulus and its extension as proposed by Triandis and Fishbein (1963). Accordingly,

$$W_p = \frac{\sum_{i=1}^5 V_i^2}{\sum_{i=1}^5 V_i}, \tag{1}$$

where V_i denotes the recognition scores of isolated facial features and W_p denotes the predicted recognition score of a whole face. Note that the W_p formula mathematically fulfills the requirement that $V_L \leq W_p \leq V_H$. The formula is based on a normalization rule, where V_i is multiplied by $V_i / \sum V_i$. The latter expression can be viewed as a weight multiplying V_i , so that $W_p = w_i V_i$, where $w_i = V_i / \sum V_i$, and interpreted as a weight value relative to all the values comprising a whole face.

The calculation of W_p used the data obtained in Experiment 1. Each subject in Experiment 1 had five values of correct choice: V_f , V_e , V_n , V_m , and V_c . Let's say a subject has the following values: $V_f = 90\%$, $V_e = 85\%$, $V_n = 80\%$, $V_m = 70\%$, and $V_c = 68\%$. Given these values, the subject's $W_p = 79.5\%$. Similarly, individual W_p s were computed for each subject in Experiment 1. Given these W_p s, mean W_p s were calculated for the four groups of Experiment 1—the UU, UI, IU, and II groups. These four W_p means were compared with the four W_o means obtained for Experiment 2 (UU, UI, IU, and II). If recognition of a whole face is based on recognition of its individual features, then there should be no significant difference between W_p and W_o . Since each of these four groups in-

volves different cognitive processes, we propose that the predictive ability of Formula 1 is maximal, when no significant differences between W_p and W_o are found for any of the four groups (UU, UI, IU, and II). (Note that in testing Formula 1, nonrejection of the null hypothesis is regarded as supporting the W_p formula. This procedure, however, raises a debate about statistical problems that go beyond the scope of the present paper. The interested reader is referred to Badia, Haber, & Runyon, 1970.) The predictive ability of Formula 1 was also compared with five alternative formulas proposed by Rakover (1994): the arithmetic mean, the harmonic mean, the geometric mean, the logarithmic mean, and the stimulus uncertainty. If the predictive strength of Formula 1 is higher than each of the alternative formulas, then Formula 1 gains additional empirical support.

Method

Subjects. Eighty students in the Department of Psychology at Haifa University volunteered to participate in the experiment. Their average age was 22.4 years. Seventy-three were female and 7 were male. The subjects were assigned randomly to the experimental groups and were tested in groups of 4.

Stimuli and Apparatus. There were 10 study faces and 10 distractor faces. The study faces were composed of the study values used in Experiment 1; the distractor faces were composed of the distractor values. Study Face 1 was composed by randomly selecting 1 out of 10 study foreheads, 1 out of 10 study pairs of eyes, 1 out of 10 study noses, 1 out of 10 study mouths, and 1 out of 10 study chins. Using Penry's Photofit kit, these features generated Study Face 1, where the distance between forehead and eyes was at the D level and the distance between mouth and chin was at the Y level in Penry's Photofit kit.

Similarly, Study Face 2 was generated by randomly selecting a value from the nine remaining study values. This method of random sampling without replacement generated the rest of the study and distractor faces.

Each of the four video cassettes presented 30 faces in two stages. In the study stage, 10 faces were presented in random order. In the test stage, 10 study faces mixed randomly with 10 distractor faces were presented. Each study face was presented on a black-and-white TV for 15 sec; each test face was presented for 10 sec.

Design and Procedure. Two independent variables were manipulated in a 2 x 2 factorial design. For the first variable, study orientation, faces were presented either upright or inverted, as was done with the second variable, test orientation. There were four groups—UU, UI, II, and IU—each with 20 subjects. The subjects were asked to view the study faces and to decide for each test face whether or not it had been previously presented. The subjects were informed of the orientation in which the faces would be presented.

Results

Table 2 presents the means of W_o , W_p , V_H , and V_L as a function of each of the four experimental groups. Mean W_p was calculated after W_o was computed for each subject in Experiment 1. The means of V_H and V_L are taken from Table 1.

A 2 (study orientation) x 2 (test orientation) ANOVA was applied to W_o . Table 2 shows that, in the study stage, W_o is higher for faces presented in the upright orientation than it is for faces presented in the inverted orientation [$F(1,76) = 8.32 p < .01$]. W_o is also higher when faces pre-

Table 2
 W_o, W_p, V_H, V_L as a Function of Group

Group	Correct Choices			
	W_o	W_p^*	V_H	V_L
UU	.89	.81	.94	.71
		†	†	†
UI	.69	.68	.85	.57
		n.s.	†	†
II	.74	.75	.83	.68
		n.s.	†	n.s.
IU	.71	.67	.80	.57
		n.s.	†	†

Note— V_H is the recognition score of the forehead and V_L is the recognition score of the nose, mouth, or chin. The values of V_H and V_L are taken from Table 1. * W_p is a mean of W_p s calculated for each subject. † $p < .01$ for $t(38)$ test comparisons between W_o and W_p or V_H or V_L .

sented in the test stage are in the upright orientation than when they are presented in the inverted orientation [$F(1,76) = 15.07, p < .01$]. W_o is also higher when faces are presented in the same orientation in the study and test stages (i.e., the UU and II groups) than when they are presented in differing orientations (i.e., UI and IU groups), that is, the interaction between the study and test orientations is significant [$F(1,76) = 23.82, p < .01$].

Subsequent planned comparisons using the Newman-Keuls test at $\alpha = .05$ revealed that W_o was higher in the UU group than in the other three groups.

Table 2 shows that $V_L \leq W_o \leq V_H$, as proposed by constraint No. 3.

The table also shows that the W_p formula generated values that were very close to W_o values, except for the UU group, in which $W_o > W_p$. (To recapitulate, the comparisons between W_p and W_o were made after an individual W_p was calculated for each subject in Experiment 1; that is, the W_p reported in Table 2 is the mean of these individual W_p s.) While Formula 1 does not show maximal predictive ability, it predicts W_o with greater accuracy than the five alternative formulas: the arithmetic mean, the harmonic mean, the geometric mean, the logarithmic mean, and the stimulus uncertainty.

Another way to test the accuracy of prediction of the W_p formula is as follows: if $W_o = W_p$, then the regression line for these variables should have the form of a linear function $y = ax + b$ with $a = 1$ and $b = 0$. However, the test of this prediction requires that values of x and y be measured on the same subject, a requirement that is not fulfilled in the present study. To approximate this demand, we matched values of W_o and W_p in the following way: The subjects in the UU group in Experiment 1 were ranked from high to low values of the forehead only, and the subjects in the UU group in Experiment 2 were ranked from high to low values of W_o . We then matched the subjects of these two groups on both ranks. Similarly, we matched the subjects in both experiments for the rest of the three groups. We then calculated the regression line of W_o and W_p : $W_o = 1.06 W_p - .01$, where $R(80) = .62, p < .01$. We found, as expected, that 1.06 does not differ sig-

nificantly from unity and that .01 does not differ significantly from zero. These results give additional empirical support to the proposition that the W_p formula predicts W_o .

This conclusion should be reached with caution, since there are several differences between the present matching procedure and the standard one (e.g., matched random assignment). While the purpose of the standard procedure is to control for variables, such as age, sex, and intelligence, which are expected to affect the dependent variable, the aim of the present procedure is to generate matched pairs of subjects, so that their two responses (W_p in Experiment 1 and W_o in Experiment 2) could be viewed as two responses made by the same subject. The variables used for matching subjects in the present procedure are not independent of the dependent variable as required by the standard procedure, but are part of the dependent variable in Experiment 1 (i.e., V_f is a component of W_p) and the dependent variable in Experiment 2 (i.e., W_o). The reason for using V_f for the matching procedure is that V_f is different from W_o (see V_H in Table 2). It is the most salient feature in a face, and is the first feature a subject processes when viewing a photograph of a whole face (e.g., Hines & Braun, 1990). Finally, in contrast to the standard procedure, in the present procedure random assignment of the subjects is made before matching.

Discussion

The main results of Experiment 2 are as follows. First, W_o in the UU group is higher than the W_o in the II, UI, and IU groups. Second, $V_L \leq W_o \leq V_H$. Third, the W_p formula predicts W_o of each group except the UU group.

The obtained inversion effects of the whole face are: $.89(\text{UU}) - .74(\text{II}) = .15$ and $.89(\text{UU}) - .69(\text{UI}) = .20$. In comparison, the predicted inversion effects are: $.06$ and $.13$, respectively. These differences stem from the fact that in the UU group W_p is significantly smaller than W_o ($W_o - W_p = .08$). This may be related to the finding that while there is no significant effect of test orientation on recognition of isolated features, it has a significant effect on recognition of a whole face. Assuming that study orientation is related to memory encoding and that test orientation is related to retrieval, retrieval plays an important role in the recognition of a whole upright face, in comparison with the recognition of an isolated feature, by utilizing configural information. Thus, recognition of a whole face in the UU group is not entirely explained by recognition of the isolated facial features as proposed by the W_p formula.

Nevertheless, the similarity of the results of Experiment 1 and Experiment 2 and the W_p formula suggests that (1) the inversion effects of a whole face can be accounted for by upright orientation, mental rotation, and feature saliency, as proposed above, and (2) the recognition of the isolated features provides a substantial contribution to recognition of a whole face.

The fact that $V_L \leq W_o \leq V_H$ suggests that recognition of a whole face cannot be explained by the most salient isolated feature or by assuming simple additivity of isolated

features. Furthermore, this finding disconfirms the hypothesis that W_o is based on the addition of featural and configural information, which predicts that $W_o > V_H$, because the recognition of V_H is based on featural information only, whereas W_o is based on both featural and configural information.

GENERAL DISCUSSION

While the results of Experiment 1 present serious problems to the configural hypothesis, since inversion effects have been found with isolated facial features, the results of Experiment 2 propose that recognition of isolated features contributes considerably to recognition of the whole face. Hence, configural information extracted from the whole face is not a necessary condition for obtaining the inversion effects. There are other factors that account for our findings: upright orientation, mental rotation, and feature saliency. These findings contrast with other hypotheses that explain the facial inversion effects on the basis of whole-face configural information, such as facial expression and prototype or schema hypotheses (for these hypotheses, see Goldstein & Chance, 1980; Valentine, 1988, 1991).

The configural hypothesis might be salvaged if we assume the "form" hypothesis, according to which an isolated facial feature is a visual form made up of basic components (for similar ideas, see Biederman, 1987; Palmer, 1977). For example, an eye is a visual form composed of such natural components as eyebrow, eyelid, eyelash, eyeball, pupil, and iris. (Note that each of these components can itself be conceived of as a visual form composed of more basic constituents.) Hence, it may be proposed that the inversion effects of an isolated facial feature is caused by impairment of spatial relations among its basic components, that is, by the impairment of its configural information.

There are two problems with the "form" hypothesis. First, the fundamental atomic facial unit or constituent is hard to define theoretically and operationally. Second, it follows from the "form" hypothesis that a major factor in the recognition of a face is the configural information extracted from an individual feature. If this is correct, then the inversion effects of a whole face can be predicted on the basis of the inversion effects of individual facial features. The results of Experiment 2 lend support to this proposal. Hence, the impairment of configural information extracted from spatial relations among the features of a face is not necessary for explaining the inversion effects of a whole face.

In view of the above, it is possible that the inversion effects of a whole face are derived from expertise for spatial relations of the components of an individual facial feature, and not from expertise for configural information of a whole face, that is, of the spatial relations among facial features (on expertise and face recognition, see Diamond & Carey, 1986; Rhodes, 1995; Stevenage, 1995).

As mentioned above, we suggest that our findings can be accounted for by upright orientation, mental rotation,

and feature saliency. These factors operate interactively rather than additively. In situations in which mental rotation is not required, the inversion effects are explained by appeal to upright orientation and feature saliency. In both experiments, recognition is higher in the UU group than in the II group, and recognition in the II group of Experiment 1 decreases as a function of the reduction in feature saliency. When mental rotation is required, recognition is higher in the UU group than in the UI or IU group. Note that the results show that the effect of mental rotation on recognition is higher than the effect of upright orientation: In both experiments, the difference in recognition between the UU group and the UI or IU group is higher than the difference between the UU and II groups.

In Experiment 1, where mental rotation is required, the inversion effects are greater in internal features than in the external features. One possible explanation for this difference is that inverted eyes and mouth (and probably nose, especially one without nostrils) are still seen as unusual upright features (see Rock, 1973, 1974). In contrast, an inverted forehead or chin is not seen as a bizarre upright feature, but as an inverted feature. This may increase errors of recognition of internal features, and may be involved in the Margaret Thatcher illusion, where inverted eyes and mouth generate a grotesque expression (e.g., Bartlett & Searcy, 1993; Thompson, 1980).

While the W_p formula presents a normalization rule by which one computes W_p as a prediction of W_o , it does not describe the cognitive processes for generating recognition of a whole face from its isolated features. One possible cognitive interpretation of the W_p formula is the justification of Constraint 3 and its basis in a normalization rule. Thus, the cognitive processing involved in recognition of a whole face is distributed among its features in such a way that (1) the recognition of a whole face falls between V_L and V_H , and (2) salient features demand more processing than do nonsalient features. Similarly, Palmer (1977) proposes that the processing of well-discriminated parts of a figure is more efficient and accurate than are poorly discriminated parts.

The finding that the W_p formula predicts W_o of each group (except the UU group) provides additional support for the hypothesis that isolated facial features were processed as a whole face. Had isolated features been processed differently from a whole face, one would expect the predictive ability of the W_p formula to be less than that obtained. (Note that this hypothesis does not negate the "prototype theory," which proposes that faces are encoded as deviations from a norm or a prototype face (see, e.g., Valentine, 1991), since faces and prototypes of faces consist of facial features that can be processed individually.)

The attempt to predict recognition of a whole face on the basis of recognition of its isolated features is part of the attempt to understand higher order face perception and recognition by the combinations of underlying components. For example, Ekman and his colleagues (e.g., Ekman, 1979; Ekman & Friesen, 1982) have proposed a system called the "Facial Active Coding System," which generates various facial movements and expressions on

the basis of primary facial "action units." Benson (1995) has discussed how high-order facial qualities, such as attractiveness, honesty, and friendliness, can be understood on the basis of three basic facial properties: gender, age, and ethnicity. He stresses the importance of the development of an appropriate facial algebra which will take into consideration the interaction among elementary properties. The interactive nature of facial features is demonstrated by such phenomena as the following: Subjects are sensitive to inward rather than outward eye movements, and sometimes they perceive different faces generated by slightly different combinations of the same facial features (e.g., Haig, 1984, 1986).

In view of the above and of the fact that significant interactions are obtained in the independent variables of both experiments, we do not propose that the perception of a whole face is the sum of the perceptions of its isolated features. Furthermore, we do not propose that facial configurational information does not play an important role in face perception and remembering (e.g., Bruce, Doyle, Dench, & Burton, 1991; Tanaka & Farah, 1993). However, we do suggest (1) that configurational information is not necessary for obtaining inversion effects in a whole face, and (2) that in the recognition of a whole face, configurational information extracted from an upright face is less important than featural information. According to the W_p formula, one can account for 91% ($.81/.89 = .91$) of the recognition of an upright whole face by its isolated features.

REFERENCES

- BADIA, P., HABER, A., & RUNYON, R. P. (1970). *Research problems in psychology*. MA: Addison-Wesley.
- BARTLETT, J. C., & SEARCY, J. (1993). Inversion and configuration of faces. *Cognitive Psychology*, **25**, 281-316.
- BENSON, P. J. (1995). Perspectives on face perception: Directing research by exploiting emergent prototypes. In T. Valentine (Ed.), *Cognitive and computational aspects of face recognition: Explorations in face space* (pp. 204-224). London: Routledge.
- BIEDERMAN, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, **94**, 115-147.
- BOWER, G. H., & GLASS, A. L. (1976). Structural units and the redintegrative power of picture fragments. *Journal of Experimental Psychology: Human Learning & Memory*, **2**, 456-466.
- BRUCE, V. (1988). *Recognizing faces*. Hillsdale, NJ: Erlbaum.
- BRUCE, V., DOYLE, T., DENCH, N., & BURTON, M. (1991). Remembering facial configurations. *Cognition*, **38**, 109-144.
- BRUYER, R., & COGET, M. C. (1987). Features of laterally displaced faces: Saliency or top-down processing? *Acta Psychologica*, **66**, 103-114.
- CAREY, S., & DIAMOND, R. (1977). From piecemeal to configurational representation of faces. *Science*, **195**, 312-314.
- CHUNG, M.-S., & THOMSON, D. M. (1995). Development of face recognition. *British Journal of Psychology*, **86**, 55-87.
- DAVIES, G. M. (1981). Face recall systems. In G. M. Davies, H. D. Ellis, & J. W. Shepherd (Eds.), *Perceiving and remembering faces* (pp. 227-250). London: Academic Press.
- DAVIES, G. M. (1986). The recall and reconstruction of faces: Implications for theory and practice. In H. D. Ellis, M. A. Jeeves, F. Newcombe, & A. Young (Eds.), *Aspects of face processing* (pp. 388-397). Dordrecht: Martinus Nijhoff.
- DAVIES, G. [M.], ELLIS, H. [D.], & SHEPHERD, J. [W.] (1977). Cue saliency in faces as assessed by the Photofit technique. *Perception*, **6**, 263-269.
- DIAMOND, R., & CAREY, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, **115**, 107-117.
- EKMAN, P. (1979). About brows: Emotional and conversational signals. In J. Aschoff, M. von Carnach, K. Foppa, W. Lepenies, & D. Ploog (Eds.), *Human ethology* (pp. 169-202). Cambridge: Cambridge University Press.
- EKMAN, P., & FRIESEN, W. V. (1982). Measuring facial movement with the Facial Action Coding System. In P. Ekman (Ed.), *Emotion in the human face* (2nd ed., pp. 178-211). Cambridge: Cambridge University Press.
- ELLIS, H. D., SHEPHERD, J. W., & DAVIES, G. M. (1979). Identification of familiar and unfamiliar faces from internal and external features: Some implications for theories of face recognition. *Perception*, **8**, 431-439.
- ENDO, M. (1986). Perception of upside-down faces: An analysis from the viewpoint of cue saliency. In H. D. Ellis, M. A. Jeeves, F. Newcombe, & A. Young (Eds.), *Aspects of face processing* (pp. 53-58). Dordrecht: Martinus Nijhoff.
- GOLDSTEIN, A. G., & CHANCE, J. E. (1980). Memory for faces and schema theory. *Journal of Psychology*, **105**, 47-59.
- HAIG, N. D. (1984). The effect of feature displacement on face recognition. *Perception*, **13**, 505-512.
- HAIG, N. D. (1986). Exploring recognition with interchanged facial features. *Perception*, **15**, 235-247.
- HINES, D., & BRAUN, J. A. (1990). Order of feature recognition in familiar and unfamiliar faces. *Brain & Cognition*, **14**, 165-184.
- JOHNSTON, R. A., & ELLIS, J. D. (1995). The development of face recognition. In T. Valentine (Ed.), *Cognitive and computational aspects of face recognition: Explorations in face space* (pp. 1-23). London: Routledge.
- MACMILLAN, N. A., & CREELMAN, C. D. (1991). *Detection theory: A user's guide*. New York: Cambridge University Press.
- OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. (1957). *The measurement of meaning*. Urbana: University of Illinois Press.
- PALMER, S. E. (1977). Hierarchical structure in perceptual representation. *Cognitive Psychology*, **9**, 441-474.
- PENRY, J. (1971a). *Looking at faces and remembering them: A guide to facial identification*. London: Blek Books.
- PENRY, J. (1971b). *Photofit kit*. Leeds: John Waddington of Kirkstall.
- PERRETT, D. I., MISTLIN, A. J., CHITTY, A. J., HARRIS, M. H., NEWCOMBE, F., & DEHAAN, E. (1988). Neuronal mechanisms of face perception and their pathology. In C. Kennard & F. C. Rose (Eds.), *Physiological aspects of clinical neuro-ophthalmology* (pp. 137-154). London: Chapman & Hall.
- PHILLIPS, R. J. (1979). Some explanatory experiments on memory for photographs of faces. *Acta Psychologica*, **43**, 39-56.
- RAKOVER, S. S. (1994). *On the relationship between a face and its parts*. Unpublished manuscript.
- RAKOVER, S. S., & CAHLON, B. (1989). To catch a thief with a recognition test: The model and some empirical results. *Cognitive Psychology*, **21**, 423-468.
- READ, J. D., VOKEY, J. R., & HAMMERSLY, R. (1990). Changing photos of faces: Effects of exposure duration and photo similarity on recognition and the accuracy-confidence relationship. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 870-882.
- RHODES, G. (1995). Face recognition and configural coding. In T. Valentine (Ed.), *Cognitive and computational aspects of face recognition: Explorations in face space* (pp. 647-680). London: Routledge.
- RHODES, G., BRAKE, S., & ATKINSON, A. P. (1993). What's lost in inverted faces? *Cognition*, **47**, 25-57.
- RHODES, G., BRAKE, S., & TAYLOR, K. (1989). Expertise and configurational coding in face recognition. *British Journal of Psychology*, **80**, 313-331.
- ROCK, I. (1973). *Orientation and form*. New York: Academic Press.
- ROCK, I. (1974). The perception of disoriented figures. *Scientific American*, **230**, 78-85.
- SCAPINELLO, K. F., & YARMEY, A. D. (1970). The role of familiarity and orientation in immediate and delayed recognition of pictorial stimuli. *Psychonomic Science*, **21**, 329-330.
- SERGEANT, J. (1984). An investigation into component and configural pro-

- cesses underlying face perception. *British Journal of Psychology*, **75**, 221-241.
- SHEPARD, R. N., & METZLER, J. (1971). Mental rotation of three dimensional objects. *Science*, **171**, 701-703.
- SHEPHERD, J. W., DAVIES, G. H., & ELLIS, H. D. (1981). Studies of cue saliency. In G. M. Davies, H. D. Ellis, & J. W. Shepherd (Eds.), *Perceiving and remembering faces* (pp. 105-131). London: Academic Press.
- STEVENAGE, S. V. (1995). Expertise and caricature advantage. In T. Valentine (Ed.), *Cognitive & computational aspects of face recognition: Explorations in face space* (pp. 24-46). London: Routledge.
- TANAKA, J. W., & FARAH, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology*, **46A**, 225-245.
- THOMPSON, P. (1980). Margaret Thatcher—A new illusion. *Perception*, **9**, 483-484.
- TRIANDIS, H. C., & FISHBEIN, M. (1963). Cognitive interaction in person perception. *Journal of Abnormal & Social Psychology*, **67**, 446-453.
- VALENTINE, T. (1988). Upside-down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology*, **79**, 471-491.
- VALENTINE, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *Quarterly Journal Experimental Psychology*, **43A**, 161-204.
- VALENTINE, T., & BRUCE, V. (1988). Mental rotation of faces. *Memory & Cognition*, **16**, 556-566.
- YARMEY, A. D. (1971). Recognition memory for familiar "public" faces: Effects of orientation and delay. *Psychonomic Science*, **24**, 286-288.
- YIN, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, **81**, 141-145.
- the whole range of values of each facial dimension. The following numbers represent different values in Penry's technique (the numbers in parentheses represent the number of values in each dimension). *Foreheads* (204): Study values—7, 27, 41, 69, 95, 112, 128, 149, 152, 167; distractors—5, 22, 38, 67, 93, 114, 126, 146, 148, 153. *Eyes* (103): Study values—6, 35, 36, 38, 43, 59, 70, 80, 82, 93; distractors—4, 27, 32, 34, 44, 55, 67, 73, 81, 87. *Noses* (100): Study values—7, 18, 20, 33, 36, 38, 40, 62, 88, 95; distractors—4, 9, 17, 27, 30, 34, 50, 57, 73, 92. *Mouths* (112): Study values—3, 13, 23, 42, 54, 74, 78, 81, 87, 102; distractors—7, 10, 16, 30, 56, 61, 69, 71, 80, 103. *Chins* (93): Study values—2, 5, 24, 29, 41, 48, 62, 70, 81, 88; distractors—3, 9, 18, 32, 45, 50, 56, 66, 79, 83.

2. Most of the experiments investigating recognition memory (including face recognition) have used either the yes-no (YN) procedure or the two-alternative forced-choice (2AFC) procedure. Although it seems reasonable to assume that different tasks are associated with different measurements and different cognitive processes, signal detection theory proposes that there is a relatively simple theoretical relation between the YN procedure and the 2AFC procedure: that is, $d'_{2AFC} = 1/\sqrt{2} d'_{YN}$, where d' is a sensitivity measure. For the subject, the 2AFC procedure is easier than the YN procedure since it is based on two observations, whereas the YN procedure is based on only one observation (see Macmillan & Creelman, 1991). In view of this, one reason for employing the YN procedure in the present research is that if our hypotheses are confirmed with the YN procedure, there is a good chance that they will be confirmed with the 2AFC procedure. Another reason is to avoid the possibility that a choice in the 2AFC procedure could be made because the subject is sure that the distractor face is not the target face and not because the subject identifies the target face.

NOTES

1. Values were randomly sampled from each facial dimension with one major constraint: The study values and the distractors would cover

(Manuscript received August 21, 1995;
revision accepted for publication August 15, 1996.)