Features and their configuration in face recognition

JAMES W. TANAKA and JOSEPH A. SENGCO
Oberlin College, Oberlin, Ohio

Tanaka and Farah (1993) have proposed a holistic approach to face recognition in which information about the features of a face and their configuration are combined together in the face representation. An implication of the holistic hypothesis is that alterations in facial configuration should interfere with retrieval of features. In four experiments, the effect of configuration on feature recognition was investigated by creating two configurations of a face, one with eyes close together and one with eyes far apart. After subjects studied faces presented in one of the two configurations (eyes-close or eyes-far), they were tested for their recognition of features shown in isolation, in a new face configuration, and in the old face configuration. It was found that subjects recognized features best when presented in the old configuration, next best in the new configuration, and poorest in isolation. Moreover, subjects were not sensitive to configural information in inverted faces (Experiment 2) or nonface stimuli (i.e., houses; Experiments 3 and 4). Importantly, for normal faces, altering the spatial location of the eyes not only impaired subjects' recognition of the eye features but also impaired their recognition of the nose and mouth features—features whose spatial locations were not directly altered. These findings emphasize the interdependency of featural and configural information in a holistic face representation.

What is the relation between the features of a face and their configuration in face recognition? As shown in Figure 1, our recognition of a well-known face is disrupted when the spatial distances between the features are slightly altered. Since Galton's time, researchers in face recognition have suggested that the spatial relations of the facial features might be as important to recognition as the features themselves. What is less clear is how information about features of a face and their configuration are encoded in the underlying representation. According to the holistic hypothesis of face recognition, information about the features of a face and their configuration interact such that changes in one source of information (e.g., configuration) should disrupt the processing of the other source of information (e.g., features). In this paper, we will begin by providing an operational definition of holistic processing. Next, we will describe previous research that examined the effects of configural information on face perception and recognition. We will then report four experiments that examined the holistic theory by testing the effects of configural change on the recognition of features.

Holistic Face Recognition

It is generally agreed that faces are recognized not on the basis of their individual features[1] but in terms of the *whole* that emerges from the features. While this claim is not controversial, operational definitions of holistic and featural processing have been lacking. Tanaka and Farah (1993) defined the concept of holistic recognition in the following task: In a learning phase, subjects memorized a set of normal faces and a set contrast stimuli of scrambled faces, inverted faces, or houses. After learning, in a two-choice test of recognition, subjects identified the face parts (eyes, nose, and mouth) and the house parts (small window, large window, and door) presented in isolation and in the whole object. For whole-object test items, the targets and foils were constructed such that they differed only with respect to the critical feature under test. For example, if the nose feature was being tested, the other features of the target and foil faces (i.e., face outline, hair, eyes, and mouth features) were identical. Therefore, the difference in subjects' performance between isolated-part and the whole-object test conditions should reflect the extent to which the object is represented as a whole. The central finding was that, for normal faces, subjects recognized parts better in the whole face than in isolation. In contrast, subjects were no better at recognizing parts of scrambled faces, parts of inverted faces, and parts of houses when shown in the whole-object condition than when shown in isolation. Thus, whereas recognition of a part from a normal face was facilitated by presence of the whole face, recognition of parts from the contrast stimuli was not facilitated by whole-object information. From these data, Tanaka and Farah (1993) suggested that faces are represented holistically.

Configural Information in Face Recognition

What role does configuration (i.e., the spatial distances between features) play in face and object recognition?

Figure 1. On the left is an unmodified photograph of Bill Clinton; on the right is a modified version of the same photograph in which the spatial positions of the eyes have been altered.

Diamond and Carey (1986) proposed that object configurations can be classified along a continuum according to their *first-order* and *second-order* relational properties. First-order relational properties refer to the categorical relations of object features. For example, a first-order description of a face would include such statements as "The eyes are *above* the nose" and "The mouth is *below* the nose." In contrast, second-order relational properties specify the distances that describe first-order relational properties in either relative or absolute terms. For example, a second-order relational description might describe the distance that the mouth is below the nose as "wide" or as "22 mm." It has been speculated that because all faces share the same first-order relational properties—that is, everybody's eyes are above their noses and mouths below their noses—recognition of individual faces depends on the encoding of second-order relational properties.

Consistent with this claim, a number of studies have shown the importance of second-order relational information in face processing. For example, in a multidimensional scaling study, Rhodes (1988) found that the degree to which two faces were perceived to be similar was dependent on the similarity of the individual facial features and the spatial relations of the features. Hosie, Ellis, and Haig (1988) found that changing the spacing of in-

ternal facial features (e.g., eyes, nose, and mouth) disrupted recognition more than did altering the external features (e.g., hairline). In a related study, Haig (1984) demonstrated the extent to which face recognition processes are sensitive to configural information. After briefly showing subjects an unfamiliar face, Haig reported that they were able to detect very slight spatial changes between the eyes, nose, and mouth features—in some instances, approximating one minute visual angle, close to the limits of visual acuity (Haig, 1984). Bruce, Doyle, Dench, and Burton (1991) compared configural effects in face and house recognition. They presented subjects with sets of computer-generated faces with identical features but with slightly different spatial configurations. In an incidental memory task, they found that subjects abstracted the prototypical configuration for each set, and that this tendency to identify the prototype as most familiar was greater for faces than for houses. Collectively, these studies demonstrate that, in tasks involving the perception and recognition of faces, humans are highly sensitive to the configural information contained in the face stimulus.

While the above studies addressed the role of configural information in face processing, relatively few studies have investigated the interaction between configural

and part information in face processing. An exception was Sergent's (1984) perceptual matching experiment in which subjects judged two simultaneously presented faces as *same* or *different*. The *different* faces varied with respect to their features, the spatial relations of their features, or both the features and spatial relations. She reasoned that, for *different* judgments, the combined condition should be no faster than either the featural or the spatial relations condition if subjects were using only one source of information. Instead, Sergent found that subjects were faster to respond "different" when the faces varied in featural (e.g., chin contour) and configural (e.g., distance between mouth and nose features) information than when they differed in featural information or configural information alone. Because this effect was found for upright faces, but not for inverted faces, Sergent argued that, in normal face perception, the sources of featural and configural information are combined into a single representation.

Although the above studies stress the importance of configural information in face processing, they do not bear directly on the question of part and configural information in face recognition for several reasons. First, the multidimensional scaling studies (Hosie et al., 1988; Rhodes, 1988) and matching study (Sergent, 1984) involved judgments of direct perception, not of recognition. It is possible that subjects use different strategies for a perceptual discrimination task, in which the face stimuli are physically present than for a recognition task in which a face stimulus is matched to a representation in memory. Second, with the exception of the Bruce et al. (1991) study, the results obtained with faces in these paradigms were not compared with results obtained with stimuli other than faces. This is important if we wish to assess the extent to which the holistic or featural representation is special to faces. Finally, none of the cited studies manipulated spatial information and examined its effect on feature identification. In this regard, the holistic account makes a straightforward prediction: If subjects encode featural and configural information together in a unitary face representation, changes in a configural information should impair recognition of its features.[2]

This prediction of the holistic hypothesis can be directly tested in the parts and wholes paradigm. After subjects learn a series of faces, their memory for the parts can be assessed under conditions where there is no configural information (i.e., in isolation), new configural information, and old configural information. The holistic hypothesis predicts that subjects should be more impaired at recognizing facial features when presented in a new configuration than when presented in an old configuration. Moreover, given the interdependency of parts and their configuration in a holistic representation, altering the spatial relation of one part should affect the recognition of *all* parts in the configuration.

EXPERIMENT 1

Previous results have shown that holistic encoding was eliminated when the features of a face were completely scrambled (Tanaka & Farah, 1993, Experiment 1). However, it could be argued that scrambling a face violates its first-order relational properties, which, in turn, disrupts holistic encoding. Therefore, it is possible that if first-order relational properties were preserved, subjects might demonstrate holistic processing. In Experiment 1, subjects studied faces in which the eyes were spaced either close together or wide apart. After the study phase, subjects identified parts of the stimulus faces presented in isolation, in a new face configuration, and in the old face configuration. As shown in Figure 2, the only difference between the new face and old face configuration conditions was the eye spacing. Moreover, note that the target and foil faces in the new and old face conditions differed by only the part being tested. For example, in Figure 2, the target and foil faces varied only with respect to the critical nose feature; the other eye and mouth features were the same in the target and foil faces.

According to the holistic theory, subjects should recognize face features better in the old configuration condition, where second-order relational information is preserved, than in the new configuration condition, where second-order relational information is changed. The holis-
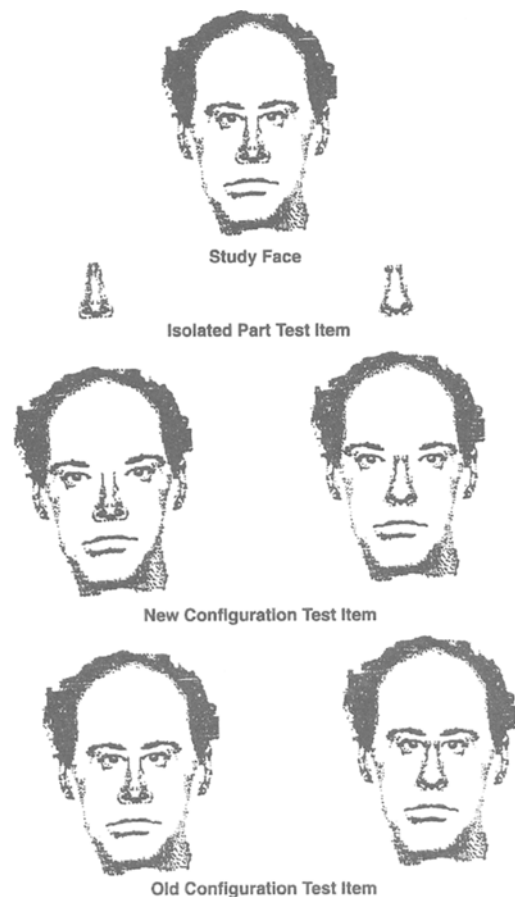

Figure 2. Example of study face with close-together eye spacing. In this example, memory for nose feature is tested when shown in isolation, in a new configuration (i.e., far-apart eye spacing), and in the old configuration (i.e., close-together eye spacing).

tic hypothesis further predicts that changes in configuration should also disrupt subjects' memory for the other features (e.g., nose and mouth)—features whose spatial locations were not directly manipulated.

Alternatively, it is possible that people encode only the first-order relational properties about a face and are not sensitive to its second-order relational properties. According to a first-order relational hypothesis, subjects should recognize the face parts as well in the new configuration condition as in the old configuration condition and better in the new and old configuration conditions than in the isolated condition. In Experiment 1, the predictions of the first-order relational hypothesis were tested against the predictions of the holistic hypothesis.

Method

Subjects. Twenty-four students from Oberlin College participated in the experiment. The subjects were tested individually and received course credit for their participation.

Materials. Six male target faces were generated on a Macintosh computer using the Mac-a-Mug software program. The six target faces were composed of a set of eyes, nose, and mouth exemplars placed in a common face outline. Each exemplar appeared only once in the target face. Two versions of each target and foil face were generated: one version where the distance separating the eyes was 4 pixels (close-together eye spacing) and one version where the distance separating the eyes was 14 pixels (far-apart eye spacing). To create face foils, the six target faces were paired, and the foil exemplars for a target face were taken from its paired counterpart. For example, if the faces of Joe and Bob were paired together, when memory for Joe's nose was tested, Bob's nose served as the foil feature, and visa versa. Thus, each feature exemplar appeared an equal number of times as a target and foil. Importantly, the foil face differed from the target face only with respect to the critical feature being tested; all other facial features, such as the face outline and the untested features, were held constant. Face images subtended a visual angle of approximately 6° in the horizontal direction and 7° in the vertical direction.

Procedure. The subjects were shown a series of six male faces paired with male names. Their task was to learn the name–face pairings. A single learning trial consisted of a name presented on a computer screen for 2 sec followed by a face. The face remained on the screen until the subject pressed the space bar, at which time the next name–face learning trial was presented following a 1-sec intertrial interval. A study block consisted of the six faces, and presentation of the faces was randomized within a study block. For three of the six faces in the study block, the eye spacing was close together; for the other three faces, the eye spacing was far apart. Assignment of the close-together and far-apart eye spacing to the study faces was counterbalanced across subjects. There were five study blocks in the experiment for a total of 30 learning trials.

Immediately following the study phase, memory for the faces was assessed in a forced-choice recognition test. The subjects saw a prompt describing the face part of one of the target faces (e.g., Bob's nose) on the computer screen for 2 sec followed by two pictures positioned on the left and right side of the screen. If the picture on the left matched the part in question, the subjects indicated their response by pressing the key on the keyboard labeled "Left." If the picture on the right matched the part description, the subjects pressed the key labeled "Right."

The subjects' memories for a face feature (i.e., eyes, nose, or mouth) was tested when shown in isolation, in a new configuration, and in an old configuration. In the *isolated-part* condition, the target and foil features were presented without the surrounding face context. In the *new-configuration* condition, target and foil features

were presented in whole faces whose eye spacings were different from the study face. For example, if a feature was studied in a face with eyes close together, then the target and foil features were presented in a face with eyes far apart. In the *old-configuration* condition, the target and foil features were presented in faces whose eye spacing was the same as that of the study face. In the isolated-part, new-configuration, and old-configuration test conditions, the target and foil items varied only by the critical feature being tested; all other feature information was held constant. The subjects' recognition for the eyes, nose, and mouth parts from each of the six study faces was tested in isolation, in the new configuration, and in the old configuration. Thus, there were a total of 48 test trials in Experiment 1. Trials were randomly presented to the subject.

Results and Discussion

As shown in Table 1, the subjects correctly identified parts shown in isolation on 65% of the trials. When the same parts were tested in a new face configuration where first-order relational information was preserved, performance increased to 72%. When part memory was tested in the old face configuration where second-order (and by default, first-order) information was preserved, recognition increased to 77%. The difference in recognition across the three test configurations was confirmed by a nonparametric Friedman test [$\chi^2(2) = 18.812, p < .001$], with a within-subjects confidence interval (CI; Loftus & Masson, 1994) of ±3%. Comparisons between the three conditions revealed that they were reliably different from each other ($p < .05$). Consistent with the holistic hypothesis, part recognition was shown to be better in the old-configuration condition, than in the new-configuration condition, suggesting that people encode facial configuration according to its second-order relational properties.

A second prediction of the holistic hypothesis was that configural changes of one feature should also disrupt subjects' memory for the other features. Comparing recognition of the spatially unaltered nose and mouth features in the two configural conditions showed that the subjects' recognition was marginally worse when tested in the new configuration than in the old configuration [Wilcoxon test, $T(17) = 39.00, p < .08$]. Thus, altering the spatial location of the eye feature also adversely affected the recognition of the spatially unaltered nose and mouth features.

In summary, when second-order relational properties of a face were changed, the subjects were impaired in their recognition of the features. Thus, evidence for the interdependency of part and configural information, as predicted by the holistic hypothesis, was found in Experiment 1. In Experiment 2, we further explored the relation between featural and configural information in face rec-

Table 1
Percent Correct Recognition for Eyes, Nose, and Mouth Features Shown in Isolation, in a New Configuration, and in the Old Configuration

| | Features | | | |
	Eyes	Nose	Mouth	M
Isolated parts	61	63	70	65
New configuration	73	70	74	72
Old configuration	79	72	81	77

ognition by comparing recognition of features from normal faces with recognition of features from a set of inverted faces.

EXPERIMENT 2

A number of studies have shown that people have more difficulty recognizing upside-down faces than other types of inverted objects, such as houses, stick figures, costumes, and dog faces (Yin, 1969). The *face-inversion effect* has been shown for well-known faces and less familiar faces (Scapinello & Yarmey, 1970; Yarmey, 1971), photographic faces (Carey & Diamond, 1977; Diamond & Carey, 1986), and simple line-drawn faces (Yin, 1969). The face-inversion effect has also been demonstrated in different experimental paradigms, including forced-choice recognition (Yin, 1969) and "old" versus "new" judgments (Valentine & Bruce, 1986).

Although the face-inversion effect is a robust experimental phenomenon, less is known about the changes in face processing brought about by inverting a face. It has been speculated that inversion disrupts holistic processes that are used for upright faces (Farah, Drain, & Tanaka, 1995; Tanaka & Farah, 1993; Young, Hellawell, & Hay, 1987). In a study by Young et al. (1987), subjects were asked to identify the top half of photographs depicting famous people. Half of the stimuli were composite faces where the top and bottom halves of the photographs were precisely aligned so as to give the impression of forming a new face. For other faces, the top and bottom were offset so the two halves appeared not to form a composite face . When subjects identified the famous person in just the top half of the photograph, their response times were reliably slower for composite faces than for the noncomposite faces. However, when the composite and noncomposite faces were inverted, no difference in response times was found. According to Young et al., in the upright condition, subjects perceived the faces as all of one piece, and this holistic perception interfered with the identification of the top half of the face. Because inversion presumably disrupts holistic perception, no difference in performance was found when the composite faces were turned upside down.

Tanaka and Farah (1993, Experiment 2) directly tested the holistic explanation of the inversion effect using the previously described parts/whole paradigm. In the study phase of the experiment, subjects learned face–name associations to a series of upright or inverted faces. Memory for the faces was tested by displaying the face parts in the same orientation as the study face presented in isolation or in the context of the whole face. A reliable interaction was found such that subjects recognized parts contained in upright faces better in the whole-face test condition than in isolation. In contrast, for faces learned in the inverted orientation, recognition performance was the same in the whole-face and isolated-part test conditions. These results, together with the Young et al. (1987)

findings, suggest that upright faces are processed holistically, whereas inverted faces are processed featurally.

What are the effects of inversion on the processing of configural information in a face? When shown upright faces, subjects seem to be sensitive to configural changes in the features of a face. For example, subjects accurately report changes in the internal spacing of the eyes in upright faces. However, when the same faces are inverted, sensitivity to changes in configural structure are reliably impaired (Rhodes, Brake, & Atkinson, 1993). Similarly, faces made grotesque by exaggerating the spatial distances of the eye and mouth features are judged to be less grotesque when shown in their inverted orientation relative to the upright orientation (Bartlett & Searcy, 1993). These studies suggest that the finely tuned perceptual mechanism that allows humans to encode second-order relational properties in an upright face are disabled when the face is turned upside down (but see Tanaka & Farah, 1991).

The behavioral evidence suggests that inversion disrupts (1) the holistic representation of faces (Farah et al., 1995; Tanaka & Farah, 1993; Young et al., 1987) and (2) the encoding of configural information in faces (Bartlett & Searcy, 1993; Rhodes et al., 1993). However, to date, no studies have examined the effects of inversion on the holistic encoding of configural information. The goal of Experiment 2 was to test the effects of inversion on the holistic encoding of configural information. After subjects studied a series of upright or inverted faces, we tested their memory for faces in which the configuration was different from or the same as the study face. For the upright test faces, we expected to replicate the results of Experiment 1 in which recognition of the feature was best when shown in an old configuration, intermediate in a new configuration, and poorest in isolation. If inversion disrupts holistic encoding of configural information, we expected recognition of the feature to be unaffected by its configural context.

Method

Subjects. Twenty-four students from Oberlin College participated in the experiment. The subjects were tested individually and received course credit for their participation.

Materials. In addition to the set of six faces used in Experiment 1, a new set of six faces was generated. The six study faces were composed of feature exemplars not included in any of the previous faces. Two versions of each study face were created with eyes far apart or close together. Following the method described in Experiment 1, target faces were paired, and foil exemplars were drawn from the target face pair.

Procedure. A procedure similar to the one used in Experiment 1 was employed in Experiment 2. The subjects were instructed that their task was to learn a series of name–face associations. Similar to Experiment 1, a single learning trial consisted of a name shown on a computer screen, followed by a face. Learning was self-paced, and the six study faces were shown five times. After the learning phase, memory for the parts was tested by showing the target part and the foil in isolation, in a new configuration, and in the old configuration. After recognition of one face set was tested, the procedure was repeated in the untested orientation with the other set of faces. The subjects' recognition for the three features (eyes, nose,

and mouth) tested in isolation, in the new configuration, and in the old configuration in the upright or inverted orientation yielded a total of 96 experimental trials. Trials were randomly presented to the subject. Assignment of the face sets to the upright and inverted orientation was counterbalanced across subjects. For half of the subjects, the upright faces were studied and tested first; for half of the subjects, the inverted faces were studied and tested first.

Results and Discussion

As shown in Table 2, in the upright condition, the subjects correctly identified parts presented in isolation, in the new configuration, and in the old configuration on 72%, 80%, and 87% of the trials, respectively. Hence, consistent with the results of Experiment 1, the subjects correctly identified a part best when shown in the old configuration, next best when shown in a new configuration, and poorest when shown in isolation.

A nonparametric Friedman test confirmed the difference in recognition between the isolated-part, new-configuration, and old-configuration test conditions for the upright faces [$\chi^2(2) = 14.771, p < .001$; CI = $\pm 4\%$]. The three conditions were reliably different from each other by Wilcoxon tests ($p < .05$). As shown in Table 2, the subjects improved in their ability to recognize the nose and mouth features in the old-configuration test condition relative to in the new-configuration condition. Direct comparisons revealed that the subjects recognized the spatially unaltered features (nose and mouth) better in the old configuration than in the new configuration [Wilcoxon $T(12) = 3.5, p < .001$]. Thus, altering the spatial relation of one feature (i.e., the eyes) not only impaired its recognition but also impaired the recognition of other features (i.e., nose and mouth) whose absolute spatial locations were not changed.

In contrast to the configural effects found for upright faces, the subjects seemed not to be sensitive to the configural information contained in upside-down faces. The subjects identified parts presented in isolation, in the new configuration, and in the old configuration on 69%, 64%, and 69% of the trials, respectively. Hence, the subjects showed no evidence that they encoded configural information from inverted faces [$\chi^2(2) = 4.738, p > .10$; CI = $\pm 4\%$]. In summary, the results from Experiment 2 demonstrated that configural information influenced the rec-

ognition of facial features from upright faces but had no effect on the recognition of features from inverted faces.

EXPERIMENT 3

The lack of configural effects for inverted faces found in Experiment 2 suggests that upside-down faces are not remembered holistically. However, these results might reflect a difference in the configural encoding of upright versus inverted objects rather than a difference between face and nonface stimuli per se. That is, it is possible that configural information of other upright stimulus objects might also be encoded holistically. To examine this alternative explanation, the holistic encoding of configural information was tested for normally oriented, nonface objects. Houses were selected as a suitable contrast stimulus set to faces because, like faces, houses have identifiable internal features (e.g., door and windows) arranged in a typical configuration (e.g., windows are usually found above the door).

In a previous study using the parts/whole memory paradigm, Tanaka and Farah (1993) found that identification of house parts was the same whether they were presented in isolation or in the context of the whole house. The absence of a parts/whole difference for houses suggests that they are not represented holistically to the same extent as faces. According to the holistic theory, changes in configural information should prove to be *less* disrupting to the identification of house parts than the identification of face parts. To test this prediction, two versions of each house stimulus were constructed (see Figure 3): one version in which the small windows were relatively close together and another version in which they were relatively far apart. After studying a house in either its close-together or far-apart configuration, subjects identified house parts presented in isolation, in a new configuration, and in the old configuration. If subjects encode configural information holistically, their recognition of the parts should be better in the old configuration than in the new configuration. On the other hand, if configural information in houses is not represented holistically, there should be no difference between the recognition in the new-configuration and old-configuration test conditions.

Method

Subjects. Twenty-four students from Oberlin College participated in the experiment. The subjects were tested individually and received course credit for their participation.

Materials. The house stimuli differed according to their door, large-window, and small-window features. Six exemplars for each door, large-window, and small-window feature were generated using the MacPaint graphics program. Following the same method employed for the face stimuli, the six house stimuli were created by selecting one of the six feature exemplars and placing it into a house outline. Each door, large-window, and small-window exemplar was used only once as a target. As shown in Figure 3, two versions of each target stimulus were generated: one version where the spatial separation between the small windows was small (70 pixels) and another version where the spatial separation between the small windows

Table 2
Percent Correct Recognition for Eyes, Nose, and Mouth Features in Upright and Inverted Faces Shown in Isolation, in a New Configuration, and in the Old Configuration

	Features			
	Eyes	Nose	Mouth	M
Upright Faces				
Isolated parts	73	67	77	72
New configuration	78	80	81	80
Old configuration	87	84	91	87
Inverted Faces				
Isolated parts	70	63	75	69
New configuration	67	67	59	64
Old configuration	72	67	67	69

Study House

Isolated Part Test Item

New Configuration Test Item

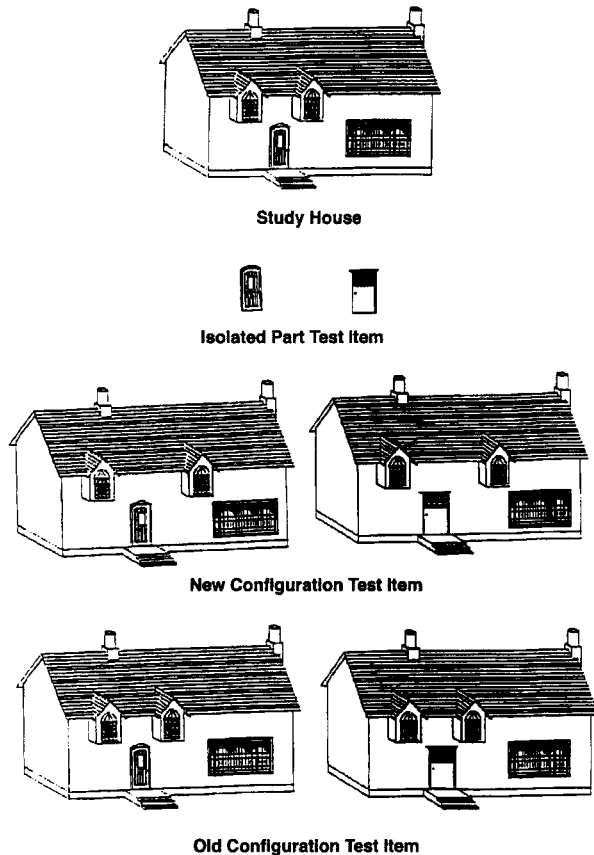**Old Configuration Test Item**

Figure 3. Example of house stimuli used in Experiment 3 with close-together window spacing. In this example, memory for the door feature is tested when shown in isolation, in a new configuration (i.e., far-apart window spacing,), and in the old configuration (i.e., close-together window spacing).

was large (135 pixels). Stimulus foils were created by pairing target stimuli and replacing the target's exemplar with the exemplar from its paired counterpart. For example, if the Jones and Smith houses were paired, the Smith door served as the foil exemplar for the Jones door, and visa versa. Thus, each exemplar appeared an equal number of times as a target and foil.

Procedure. Learning and test trials were blocked according to stimuli type (houses and faces). At the beginning of the learning phase, the subjects were told that they would see a series of faces paired with male names or a series of houses paired with six common surnames (e.g., Jones and Smith). The subjects were informed that their task was to learn the name–house (name–face) pairings. No special instructions concerning the spatial configuration of the stimuli were given. Once the subjects understood the task, the learning phase began. A single learning trial consisted of a name presented on a computer screen for 2 sec followed by a house (face) stimulus. The stimulus remained on the screen until the subject pressed the space bar. Following a 1-sec intertrial interval, the next learning trial was presented. Three of the six to-be-learned stimuli had a small spatial separation, and three had a large spatial separation. Assignment of small and large separations to stimuli were counterbalanced across subjects. Presentation of stimuli was randomized within a learning block.

Immediately following the learning phase, memory for the face and house parts was assessed in a forced-choice recognition test. The subjects were presented with a label of a face part or house part

(e.g., Bob's nose, the Jones's door) for 2 sec followed by a target and foil stimulus. In order to display both stimuli on the screen, the house stimuli were vertically aligned with each house stimulus subtending a visual angle of approximately 9°. Faces were horizontally aligned. The top–bottom locations of the target houses and left–right locations of the target faces were counterbalanced across test trials. The subjects indicated their selection by pressing the corresponding key marked "Top" or "Bottom" for houses and "Left" or "Right" for faces.

Identification of the face and house parts was tested in three experimental conditions: in isolation, in a new configuration, and in an old configuration. In the isolated-part condition, the target and foil part were shown without the surrounding house (or face) context. In the new-configuration condition, part identification was tested in a whole stimulus whose spatial separation was different from the studied stimulus. For example, if a subject studied a house with a small separation between windows in the learning phase, the separation of the target and foil houses would be large in the test phase. In the old-configuration condition, the part was tested with a whole stimulus whose spatial separation was identical to the separation found in the learning phase. The six exemplars, two stimulus types (faces and houses), three test conditions (isolated part, new configuration, and old configuration), and three features yielded a total of 108 experimental trials. Isolated-part, new-configuration, and old-configurations test trials were randomly presented. For half of the subjects, the house stimuli were studied and tested first; for half of the subjects, the face stimuli were studied and tested first.

Results and Discussion

Overall, the subjects correctly identified house parts on 84% of the trials and correctly identified face parts on 78% of the trials. This difference approached a reliable level [Wilcoxon $T(23) = 1.931$, $p < .06$]. As shown in Table 3, the subjects were least accurate when identifying face features shown in the isolated-part condition (70% correct), intermediate in the new-configuration condition (75% correct), and best in the old-configuration condition (81% correct). The difference across the three configural conditions was confirmed by a Friedman test [$\chi^2(2) = 11.646, p < .001$; CI $= \pm 3\%$]. Direct comparisons between the old-configuration, new-configuration, and isolated-part conditions by a Wilcoxon test revealed that the conditions were significantly different from each other ($p < .05$). Moreover, recognition of the nose and mouth features was reliably worse in the new-configuration condition than in the old-configuration condition [Wilcoxon $T(18) = 2.286, p < .02$]. Hence, changing the spatial location of the eye features adversely affected the recognition of nose and mouth features—features whose absolute spatial locations were not changed.

Unlike face recognition, house recognition did not vary across the three configural test conditions. The subjects correctly identified house parts on 83% of the trials when tested in the old configuration, on 84% of the trials when tested in a new configuration, and on 84% when tested in isolation. Thus, recognition of house parts was not affected by the presence (or absence) of configural information [Friedman test, $\chi^2(2) = .029, p > .10$; CI $= \pm 3\%$].

The main result of Experiment 3 was that changes in second-order relational properties affected the recognition of face parts, but not the recognition of house parts. While this finding indicates that faces are processed holistically,

Table 3
Percent Correct Recognition for Face Features (Eyes, Nose, and Mouth) and
House Features (Big Window, Small Window, and Door) Shown in Isolation, in a
New Configuration, and in the Old Configuration for Experiment 3 and Experiment 4

	Faces				Houses			
	Features				Features			
	Eyes	Nose	Mouth	M	Big Window	Small Window	Door	M
Experiment 3								
Isolated parts	66	70	75	70	92	75	84	83
New configuration	79	71	76	75	92	76	85	84
Old configuration	83	76	85	81	91	72	85	83
Experiment 4								
Isolated parts	72	61	75	69	82	78	74	78
New configuration	75	70	78	74	81	81	74	79
Old configuration	87	78	88	85	81	81	76	79

it does not necessarily imply that houses are processed featurally. It is possible that, because the house features used in Experiment 3 were perceptually more salient than face features, the subjects adopted a featural encoding strategy for houses. Thus, the *perceptual discrimination hypothesis* claims that the encoding strategy—featural or holistic—is determined by relative discriminability of the object features. According to this approach, reducing the discriminability of the house features should induce holistic processing as evidenced by a greater part–whole difference. In Experiment 4, this prediction of the perceptual discrimination hypothesis was tested.

EXPERIMENT 4

The learning and testing procedures used in Experiment 4 were the same as the procedures used in Experiment 3. The only difference between Experiments 3 and 4 was that the house features in Experiment 4 were modified to be more perceptually similar to each other. According to the perceptual discrimination hypothesis, decreasing the discriminability of the house features should increase the amount of holistic processing in recognition.

Method

Subjects. Twenty-nine students from Oberlin College participated in the experiment. The subjects were tested individually and received course credit for their participation.

Materials. The house features (big windows, small windows, and doors) used in Experiment 3 were altered such that they were made more perceptually similar to their foils. The face stimuli were composed of the faces used in Experiment 3. Similar to Experiment 3, two versions of each house and face stimulus were generated: one version where the spatial separation between the small windows was small and another version where the spatial separation between the small windows was large.

Procedure. The learning and test procedures were identical to the procedures used in Experiment 3.

Results and Discussion

The subjects recognized the correct face parts on 76% of the trials and recognized the correct house parts on 79% of the trials. The difference between face and house recognition was not reliable by a Wilcoxon test ($p > .10$).

Thus, by decreasing the discriminability of the house features, the overall recognition of the house features was brought to about the same level as that of face features.

As shown in Table 3, configuration affected the recognition of face features and house features differently. In the face recognition condition, 69% of the facial features were correctly identified when shown in isolation, 74% were correctly identified when shown in a new configuration, and 85% were correctly identified when shown in the old configuration. Thus, as configural information improved so did recognition [Friedman test, $\chi^2(2) = 24.87$, $p < .001$; CI $= \pm 4\%$]. Comparisons between the old-configuration, new-configuration, and isolated-part test conditions also showed that they were reliably different from each other by a Wilcoxon test ($p < .05$). Consistent with the predictions of the holistic hypothesis, it was found that changing the spatial position of eye features in the new-configuration test condition reliably impaired recognition of the nose and mouth features relative to their recognition in the old-configuration test condition [Wilcoxon test, $T(25) = 3.175, p < .001$].

In contrast to recognition of face features, the subjects' recognition of house features did not vary across the configural test conditions. Recognition of house parts was 78%, 79%, and 79% for the isolated-part, new-configuration, and old-configuration test conditions, respectively. Hence, for house features, configuration had no effect on recognition [$\chi^2(2) = 0.983, p > .10$; CI $= \pm 3\%$].

To summarize the main results of Experiment 4, it was found that reducing the discriminability of the house features failed to induce holistic processing. When recognition of isolated parts of face and house features was comparable, recognition of faces demonstrated holistic processing, whereas recognition of houses did not. As shown in Table 3, recognition of the isolated eyes (72%) and mouth (75%) features was roughly equivalent to the recognition of the isolated small-window (74%) and door (78%) features of the house. Despite the comparable levels of recognition in the isolated-part condition, there was a striking difference between recognition of the face and house features in the configural test conditions. Whereas

recognition of the eye and mouth features improved when second-order relational information was preserved, recognition of the house features was not affected by the presence (or absence) of second-order relational information. These results suggest that, counter to the perceptual discrimination hypothesis, the perceptual saliency of the object features does not determine the extent to which an object is recognized holistically.

GENERAL DISCUSSION

Holistic recognition has been defined as the difference in recognition when a feature is tested in isolation versus when tested in the context of the whole object (Tanaka & Farah, 1993). The aim of these experiments was to test the effects of configuration on holistic recognition. Face configuration was manipulated by increasing or decreasing the distance between the eyes, thereby preserving its first-order, but not its second-order, relational properties. In Experiment 1, we found that facial features were recognized better when presented in their original configuration than when shown in a new configuration or in isolation. In Experiment 2, we found evidence for configural processing of upright faces, but not inverted faces. Similarly, in Experiments 3 and 4, the subjects were again sensitive to configural information in faces, but they were not sensitive to configural changes in houses. Thus, in four experiments, we found that configuration affected the holistic recognition of features from normal faces but had no effect on recognition of features from inverted faces (Experiment 2) or houses (Experiments 3 and 4).

These experiments were based on the premise that, in holistic recognition, information about the features of a face and information about their configuration are encoded together in the face representation. It follows that changes in one type of information (configuration) should affect memory for the other type of information (feature). Supportive of this prediction, we found that modifying the spatial locations of one feature (i.e., eyes) impaired recognition of other features (i.e., nose and mouth)—features whose spatial locations were not directly changed. This latter result implies that the spatial information of a feature is not defined by its absolute position, but by its position relative to other features in the representation.

The emphasis of these experiments has been on whole-face representations. However, this is not to deny that face representations contain a part structure. Clearly, in our experiments, the subjects performed above chance when recognizing face parts presented in isolation, indicating that the individual face parts were encoded independent of the other features and their configuration. In support of this view, Reinitz and colleagues (Reinitz, Lammers, & Cocharan, 1992; Reinitz, Morrissey, & Demb, 1994) found that subjects will falsely identify a new face as "old" if it is a face composite of features taken from previously learned faces. While our results and the Reinitz findings indicate that faces and other objects, such as houses, are represented as parts, only faces are repre-

sented as wholes (see Farah, 1992, for further discussion of featural and holistic representation).

In contrast to the holistic view of configuration, it has been alternatively suggested that distance between features could be encoded not as a global configuration but as a kind of local feature (Bruce, 1988; Rhodes et al., 1993). For example, the distance between the eyes could be encoded as a spatial value, independent from a description of the eyes themselves. However, this account falls short of explaining the present results for several reasons. First, if the subjects in our experiments encoded the spatial distance between the eyes independent of the information about the eye part itself, it is not clear why recognition of the eye feature should be disrupted in the new configuration. Second, a local coding scheme could not account for the impaired recognition of nose and mouth features—features whose spatial properties were not directly manipulated in the new configuration. Of course, it is possible that the combinations of spatial distances contained in the face could be independently determined (e.g., the distance between the eyes and nose, the distance between the eyes and mouth, etc.), but such an encoding process would be very close to the type of interactive encoding characteristic of holistic processing (Carey, 1992; Sergent, 1984; Tanaka & Farah, 1993).

In some ways, our findings are similar to previous reports of a *face-superiority* effect, in which the parts of a face are better perceived if presented in the context of a whole face than in the context of a scrambled face (e.g., Homa, Haver, & Schwartz, 1976; Mermelstein, Banks, & Prinzmetal, 1979). However, there are important distinctions that separate the two phenomena. Whereas the parts/whole paradigm evaluates the processes of face memory, the face-superiority task measures the processes of face perception under conditions of threshold vision. Also, in face- and object-superiority effects, the main finding is that the perception of a part in context is as good as, but not better than, recognition of the isolated part (Pomerantz, 1981). In contrast, we found that recognition of parts in the whole face was better than recognition of isolated parts. Third, the superiority effect does not appear to be specific to faces but is a more general phenomenon involving the visual encoding of parts in context, whether those objects are words (Reicher, 1969; Wheeler, 1970), geometric forms (Enns & Gilani, 1988; Weisstein & Harris, 1974), or chairs (Davidoff & Donnelly, 1990). In contrast, the holistic recognition of faces was not found with the other types of stimuli.

In summary, the central goal of the present experiments was to operationally define what is meant by *configural processing*. While past studies have stressed the importance of spatial relations in face processing (Galton, 1879; Haig, 1984; Hosie et al., 1988), few studies have directly addressed the question of whether configuration is encoded independently of or combined with featural information. The holistic approach maintains that featural and configural information interact in the face representation. Accordingly, the holistic position predicts that changes in one type of information should pro-

duce concomitant changes in the other. Consistent with this prediction, results from the present experiments show that changes in a face's configuration affect the recognition of its features.

An important question to consider is whether holistic recognition is special to faces. These studies show that stimulus complexity (Experiment 2), orientation (Experiments 3 and 4), and discriminability (Experiment 4) are not sufficient conditions in themselves to produce holistic encoding. Holistic encoding is likely to emerge when the to-be-recognized objects are highly similar in shape, thereby preventing recognition on the basis of local featural information (Rhodes, 1988). Under these circumstances, it is plausible that recognition is mediated by combining relatively small featural and configural differences into a single holistic representation. Additionally, holistic representations might be necessary when recognition decisions must be made very quickly and at specific levels of abstraction. While face recognition is the paradigmatic case of this kind of recognition (Carey, 1992; Tanaka & Gauthier, 1997), it is conceivable that holistic representations might be involved in other types of expert recognition (Gauthier & Tarr, in press)—for instance, in expert dog (Diamond & Carey, 1986) and bird (Tanaka & Taylor, 1991) recognition. Examining the relation between holistic object representation, task demands, and experience is a promising avenue for future research.

REFERENCES

BARTLETT, J. C., & SEARCY, J. (1993). Inversion and configuration of faces. Cognitive Psychology, 23, 281-316.

BIEDERMAN, I. (1987). Recognition-by-components: A theory of human image understanding. Psychological Review, 94, 115-117.

BRUCE, V. (1988). Recognizing faces. London: Erlbaum.

BRUCE, V., DOYLE, T., DENCH, N., & BURTON, A. M. (1991). Remembering facial configurations. Cognition, 38, 109-144.

CAREY, S. (1992). Becoming a face expert. Philosophical Transactions of the Royal Society of London: Series B, 335, 95-103.

CAREY, S., & DIAMOND, R. (1977). From piecemeal to configurational representation of faces. Science, 195, 312-314.

DAVIDOFF, J., & DONNELLY, N. (1990). Object superiority: A comparison of complete and part probes. Acta Psychologica, 73, 225-243.

DIAMOND, R., & CAREY, S. (1986). Why faces are and are not special: An effect of expertise. Journal of Experimental Psychology: General, 115, 107-117.

ENNS, J. T., & GILANI, A. B. (1988). Three-dimensionality and discriminability in the object-superiority effect. Perception & Psychophysics, 44, 243-256.

FARAH, M. J. (1992). Is an object an object an object? Cognitive and neuropsychological investigations of domain-specificity in visual object recognition. Current Directions in Psychological Science, 1, 164-169.

FARAH, M. J., DRAIN, H. M., & TANAKA, J. W. (1995). What causes the inversion effect? Journal of Experimental Psychology: Human Perception & Performance, 21, 628-634.

GALTON, F. (1879). Composite portraits, made by combining those of many different persons into a single, resultant figure. Journal of the Anthropological Institute, 8, 132-144.

GAUTHIER, I., & TARR, M. J. (in press). Becoming a Greeble expert: Exploring mechanisms for face recognition. Vision Research.

HAIG, N. D. (1984). The effect of feature displacement on face recognition. Perception, 13, 104-109.

HOFFMAN, D. D., & RICHARDS, W. A. (1984). Parts of recognition. In S. Pinker (Ed.), Visual cognition (pp.65-96). Cambridge, MA: MIT Press.

HOMA, D., HAVER, B., & SCHWARTZ, T. (1976). Perceptibility of schematic face stimuli: Evidence for a perceptual Gestalt. Memory & Cognition, 4, 176-185.

HOSIE, J. A., ELLIS, H. D., & HAIG, N. D. (1988). The effect of feature displacement on the perception of well-known faces. Perception, 17, 461-474.

LOFTUS, G. R., & MASSON, M. E. J. (1994). Using confidence intervals in within-subject designs. Psychonomic Bulletin & Review, 1, 476-490.

MERMELSTEIN, R., BANKS, W., & PRINZMETAL, W. (1979). Figural goodness effects in perception and memory. Perception & Psychophysics, 26, 472-480.

POMERANTZ, J. R. (1981). Perceptual organization in information processing. In M. Kubovy & J. R. Pomerantz (Eds.), Perceptual organization (pp. 141-180). Hillsdale, NJ: Erlbaum.

REICHER, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. Journal of Experimental Psychology, 81, 275-280.

REINITZ, M. T., LAMMERS, W. J., & COCHRAN, B. P. (1992). Memory-conjunction errors: Miscombination of stored stimulus features can produce illusions of memory. Memory & Cognition, 20, 1-11.

REINITZ, M. T., MORRISSEY, J., & DEMB, J. (1994). Role of attention in face encoding. Journal of Experimental Psychology: Learning, Memory, & Cognition, 20, 161-168.

RHODES, G. (1988). Looking at faces: First-order and second-order features as determinants of facial appearance. Perception, 17, 43-63.

RHODES, G., BRAKE, S., & ATKINSON, A. (1993). What's lost in inverted faces? Cognition, 17, 25-57.

SCAPINELLO, K. F., & YARMEY, A. D. (1970). The role of familiarity and orientation in immediate and delayed recognition of pictorial stimuli. Psychonomic Science, 21, 329-331.

SERGENT, J. (1984). An investigation into component and configural processes underlying face perception. British Journal of Psychology, 75, 221-242.

TANAKA, J. W., & FARAH, M. J. (1991). Second-order relational properties and the inversion effect: Testing a theory of face perception. Perception & Psychophysics, 50, 367-372.

TANAKA, J. W., & FARAH, M. J. (1993). Parts and wholes in face recognition. Quarterly Journal of Experiment Psychology, 46A, 225-245.

TANAKA, J. W., & GAUTHIER, I. (1997). Expertise in object and face recognition. In R. L. Goldstone, P. G. Schyns, & D. L. Medin (Eds.), Psychology of learning and motivation (Vol. 36, pp. 83-125). San Diego, CA: Academic Press.

TANAKA, J. W., & TAYLOR, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? Cognitive Psychology, 23, 457-482.

VALENTINE, T., & BRUCE, V. (1986). The effect of race, inversion and encoding activity upon face recognition. Acta Psychologica, 61, 259-273.

WEISSTEIN, N., & HARRIS, C. S. (1974). Visual detection of line segments: An object-superiority effect. Science, 186, 752-755.

WHEELER, D. D. (1970). Processes in word identification. Cognitive Psychology, 1, 59-85.

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.

YOUNG, A. W., HELLAWELL, D., & HAY, D. C. (1987). Configurational information in face perception. Perception, 16, 747-759.

NOTES

1. We will use the term feature to refer to those parts of an object or a face that can be identified on the basis of its natural discontinuities (Biederman, 1987; Hoffman & Richards, 1984).

2. Note that the opposite effect of featural change on configural encoding cannot be directly addressed given that a change in featural information produces a concurrent change in configuration.