Psychological Review
1998, Vol. 105, No. 3, 482–498


0033-295X/98/$3.00

# What Is "Special" About Face Perception?

Martha J. Farah and Kevin D. Wilson
University of Pennsylvania

Maxwell Drain
University of Michigan

James N. Tanaka
Oberlin College

There is growing evidence that face recognition is "special" but less certainty concerning the way in which it is special. The authors review and compare previous proposals and their own more recent hypothesis, that faces are recognized "holistically" (i.e., using relatively less part decomposition than other types of objects). This hypothesis, which can account for a variety of data from experiments on face memory, was tested with 4 new experiments on face perception. A selective attention paradigm and a masking paradigm were used to compare the perception of faces with the perception of inverted faces, words, and houses. Evidence was found of relatively less part-based shape representation for faces. The literatures on machine vision and single unit recording in monkey temporal cortex are also reviewed for converging evidence on face representation. The neuropsychological literature is reviewed for evidence on the question of whether face representation differs in degree or kind from the representation of other types of objects.

Several lines of research have suggested that face recognition is "special." Neuropsychological studies have demonstrated that face recognition can be selectively impaired relative to the recognition of objects of equivalent difficulty, implying that people use different brain areas for face recognition and other types of object recognition (Farah, Klein, & Levinson, 1995). Single unit recordings in monkeys have revealed a population of cells in the temporal cortex that respond selectively to faces, in some cases responding differentially to particular faces, suggesting a role for these cells in face recognition (e.g., see Desimone, 1991, for a recent review). Although cells have also been found that respond to nonface objects, the selectivity and strength of such responses are weaker (Baylis, Rolls, & Leonard, 1985). Developmentally, face recognition appears to have an innate component. At just 30 min of age, infants will track a moving face farther than other moving patterns of comparable contrast, complexity, and spatial frequency (Johnson, Dziurawiec, Ellis, & Morton, 1991). The face inversion effect, discussed in more detail later, provides another indication that face recognition is different from other kinds of object recognition. Whereas most objects are somewhat harder to recognize upside down than rightside up, inversion makes faces dramatically harder to recognize (see Valentine, 1988, for a review).

The question of how face recognition is special has received less attention. How is shape represented for purposes of face recognition, and how does this differ from the representations of shape used for object recognition? Of course, it is possible that there are no differences and that face recognition is equivalent to other forms of visual pattern recognition in terms of the underlying visual information processing involved. Nevertheless, the differences just reviewed in neural implementation, developmental course, and sensitivity to orientation make it reasonable to suspect that differences also exist in the way faces and objects are represented and to inquire into those differences.

## Hypotheses About Face Representation

Most hypotheses about face representation highlight the importance of the overall structure or "gestalt" of faces relative to other kinds of objects that people recognize. This general idea has been subjected to a variety of specific formulations and operationalizations. Here we review a number of these formulations, briefly mentioning the type of experimental paradigm by which each formulation has been operationalized. We also note briefly the weaknesses of each approach, which have motivated continued attempts to frame and test new hypotheses concerning face perception.

Bradshaw and Wallace (1971) formulated the hypothesis that faces are perceived as gestalts in terms of Sternberg's (1969) distinction between parallel and serial processing. They tested the hypothesis that facial features are perceived simultaneously or *in parallel* using an adaptation of the short-term memory scanning paradigm: Each participant's task was to compare sequentially presented faces and judge them same or different. Pairs that were different could differ in one or more (up to a total of seven) of their features, and the question of interest was

Martha J. Farah and Kevin D. Wilson, Department of Psychology, University of Pennsylvania; Maxwell Drain, Department of Psychology, University of Michigan; James N. Tanaka, Department of Psychology, Oberlin College.

This research was supported by National Institutes of Health Grants R01 NS34030, R01 AG14082, K02 AG00756, and R15 HD30433. We thank Don Hoffman and Jim Johnston for helpful comments on an earlier version of this article.

Correspondence concerning this article should be addressed to Martha J. Farah, Department of Psychology, University of Pennsylvania, 3815 Walnut Street, Philadelphia, Pennsylvania 19104-6196.

whether reaction times to these pairs would vary as a result of the number of features that differed. Bradshaw and Wallace found that the number and identity of differing features did affect reaction time, consistent with a serial self-terminating search for differences, and concluded that faces in their experiment were not perceived as unitary gestalts. The Bradshaw and Wallace experiment was one of several conducted in the 1970s using speeded "same"–"different" paradigms in which it was assumed that if the features of faces are processed in parallel, there should be no effect of number of differing features (e.g., Matthews, 1978; Smith & Nielsen, 1970). This seems a questionable assumption, because the greater the number of differing features, the more dissimilar the two faces will be overall and the easier it will be to detect the difference, no matter how they are represented. In addition, Bradshaw and Wallace (1971) themselves pointed out the possibility that their experimental task may have induced special strategies (p. 447).

Rhodes (1988) formulated the issue in terms of what she labeled first versus second order, or *configurational,* features. First order features were taken to be the appearance of relatively discrete facial features labeled with common words such as *eye, nose, chin,* and so on. Second order features were defined by Rhodes as having configurational properties, under which she included spatial relations among first order features and the position of first order features, along with information about face shape. The hypothesis tested by Rhodes was that face perception involves second order features and is hence at least partly configurational. To determine which types of features predict perceived facial similarity, Rhodes used multidimensional scaling of similarity ratings among a set of faces and then regressed a large set of both first and second order features on the scaling solution. Her results indicated that both first and second order features were relevant determinants of facial appearance. Although this work represents a milestone in methodological sophistication, a drawback is that the concept of configuration was not explicitly defined. For example, is eye tilt, which was included in the analysis, a first order shape feature or a second order position feature? From the point of view of recasting the intuitive concept of gestalt face processing into more explicit information-processing terms, the concept of configuration is not a tremendous improvement. In addition, because Rhodes did not conduct similar analyses with patterns other than faces, she did not draw any conclusions about how face perception differs from the perception of other patterns.

Yet another formulation of the idea that overall structure or gestalt is important in face perception comes from Sergent (1984), who applied Garner's (1974) framework of dependence versus independence of stimulus feature processing to faces and suggested that the perception of facial features shows *dependency* or mutual influence. She analyzed participants' performance in a speeded matching task and a similarity rating task and found that the effects of combinations of features could not always be predicted by the effects of the features individually. Furthermore, this conclusion held only for upright faces. A limitation of these results pointed out by Bruce (1988) is that only one facial feature displayed this pattern of nonindependence with the other features, the feature termed "internal space." This refers to how closely grouped the features are toward the center of the face, which is a more relational property than what one normally thinks of as an individual feature. Nevertheless, the finding cannot simply be a trivial consequence of labeling the relations among a set of features as a feature, because the same feature did not interact with the others in inverted faces.

Another well-known attempt at formulating what is special about face perception is in terms of spatial frequency. Harmon (1973) and Ginsburg (1978) demonstrated that *low spatial frequency information* may be particularly important in face recognition and that the high spatial frequencies may be of only marginal additional use. Because low spatial frequencies represent the large-scale structure of the face, this hypothesis provides another explicit version of the general idea that faces are represented as gestalts. Again, Bruce (1988) provided a useful critique of this hypothesis, including a review of subsequent empirical work that challenges the special role of low spatial frequencies in face perception using filtered images.

The final hypothesis discussed here is based on an analogy with the word superiority effect in reading (Reicher, 1969; Wheeler, 1970; this hypothesis was suggested to us by Don Hoffman, personal communication, 1995). Perhaps faces are represented in terms of parts and wholes to the same degree as other visual patterns, but the whole-level representations are particularly important in encoding the part representations.

The foregoing hypotheses are different, but they have in common an emphasis on the overall structure of the face, above and beyond its more local featural information. The hypothesis of Bradshaw and Wallace (1971), that face perception involves parallel perception of features, captures the idea that all parts of the face are perceived simultaneously. However, this is a weaker notion of overall structure or gestalt than the hypothesis of Sergent (1984), in which the parts are not only perceived together but influence one another so that, in effect, the "whole is more than the sum of its parts." Rhodes's (1988) concept of configurational information captures yet another way in which face perception might transcend local feature perception, by involving nonlocal features consisting of relations among local features. This is distinct from the parallel processing hypothesis, which does not broach the topic of relations among features being processed in parallel, and is at least formally distinct from the interdependence hypothesis, which concerns the interactions among features rather than the features themselves. The spatial frequency hypothesis describes the difference between local features and overall structure in terms of the different spatial scales at which such information is typically available. The facilitation of part encoding by whole context emphasizes the importance of overall structure in deriving a part-based representation. The hypothesis presented shortly is yet another attempt to capture the notion of gestalt face representation in explicit terms from cognitive psychology. Before presenting our hypothesis, we review one last alternative, one that is quite distinct from the rest and that has become perhaps the best-known hypothesis about the distinctive nature of face recognition.

Diamond and Carey (1986) distinguished between what they termed *first order relational information,* which is sufficient to recognize most objects, and *second order relational information,* which is needed for face recognition. They defined these terms explicitly, and their meanings appear to be quite distinct from Rhodes's (1988) similar terminology. First order relational information consists of the spatial relations of the parts of an

object with respect to one another. In contrast, second order relational information exists only for objects whose parts share a general spatial configuration, and it consists of the spatial relations of the parts relative to the prototypical arrangement of the parts. For a face, the first order relational information would include the spatial relations among the eyes, nose, and mouth, for example. The second order relational information would include the spatial locations of these parts relative to the prototypical arrangement of eyes, nose, and mouth. Diamond and Carey also suggested that the use of prototypes and second order relational information is not unique to face recognition but underlies all "expert" recognition of objects with prototypical spatial configurations. In support of their hypothesis, they demonstrated that dog recognition by dog experts showed an inversion effect comparable in magnitude to the face inversion effect, whereas nonexperts, who had developed prototypes for human faces but not for dogs, showed only a face inversion effect.

Although Diamond and Carey's (1986) results show that pronounced inversion effects are not necessarily limited to faces and that expertise may be an important factor in determining the mode of encoding, these results do not directly address the role of first versus second order relational information per se in face recognition. Diamond and Carey's data do not reveal the types of visual information processing that their dog experts were using when recognizing dogs or the types that all of their participants were using when recognizing faces.

Tanaka and Farah (1991) conducted a direct test of the hypothesis that second order relational information is particularly sensitive to inversion and that this sensitivity underlies the face inversion effect. They reasoned that the strongest and most direct test of such a hypothesis would involve nonface stimuli and would consist of varying the relative importance of first and second order relational information for stimulus recognition while holding other aspects of the stimuli and task constant. To this end, they taught participants to identify dot patterns that either shared a common configuration (each pattern having been generated from a prototype by small changes in dot position) or did not. In two experiments, they obtained a moderate inversion effect for the dot patterns, but there was no difference between the two types of patterns. They concluded that relatively greater reliance on second order relational information does not necessarily result in greater sensitivity to pattern inversion. The implication of this finding for the face inversion effect and the nature of face recognition more generally is that the face inversion effect is probably not caused by reliance on second order relational information, and therefore the underlying difference between face and object recognition is probably not the degree to which they rely on second order relational information.

## Holistic Face Representation: Evidence of Minimal Part Decomposition in Memory Representations of Faces

Our alternative hypothesis about the difference between face and object recognition concerns the degree of part decomposition used in representing the two types of stimuli. In many current theories of object recognition, stimulus shape is represented in terms of explicitly represented parts, such that parts are represented as shapes in their own right (e.g., Biederman, 1987; Hoffman & Richards, 1985; Marr, 1982; Palmer, 1975). Our hypothesis is that face recognition differs from other types of recognition in that it involves relatively little part decomposition. For example, according to most theories of vision, recognizing a particular house involves explicitly representing at least some of the parts of the house, such as the door, window, front steps, and so forth, whereas, according to our hypothesis, recognizing a particular face does not involve (or involves to a lesser extent) explicit representations of the eyes, nose, and mouth. Instead, we hypothesize that faces are recognized primarily as undifferentiated wholes. Note that we do not deny the possibility of a mixed population of representations, some of which are holistic and some of which feature explicitly represented parts. Indeed, people's ability to recognize and distinguish isolated parts of faces, even if it is less proficient than their corresponding ability with whole faces, suggests that they must have access to explicit representations of facial parts under at least some circumstances. Such part representations might bear a hierarchical relation to people's whole face representations, analogous to the relation between letter and word representations, or might simply constitute an independent population of representations. Our claim is that, to the extent that this is true, face recognition involves disproportionately more holistic representations than the recognition of other types of patterns.

We previously tested this hypothesis with two types of experiments. The first was based on an approach developed by cognitive psychologists in the 1970s to the question of which portions of a visual pattern are psychologically real parts and which are not. Bower and Glass (1976) demonstrated that some portions of a pattern provided an effective retrieval cue for drawing the pattern from memory and others did not. This distinction corresponded to whether the portions were "good" parts according to Gestalt principles. Reed (1974) found that participants were more likely to be able to verify that pattern fragments that were "good" parts were contained in a mentally imaged pattern. Palmer (1977) obtained ratings of the goodness or naturalness of different ways of parsing patterns and showed that participants were better able to recognize that a pattern fragment came from a previously studied pattern if it was independently rated a good or natural part.

These studies all demonstrate that when a portion of a pattern corresponds to a part in the natural parse of the pattern by the visual system, it will be better remembered. They thus provide an assay for the degree to which a portion of a pattern is treated as a psychologically real part by the viewer. In our initial experiments, we relied most directly on Palmer's (1977) approach, reasoning that if a portion of a pattern is explicitly represented as a part in the visual representation of the stimulus that underlies recognition, then it should be identified more accurately when presented in isolation from the rest of the pattern than when it does not have the status of a part in the pattern representation. In contrast to Palmer's (1977) research, which focused on memory for geometric patterns, ours involved realistic drawings.

Tanaka and Farah (1993) taught participants to identify faces and various contrasting classes of nonface stimuli and then assessed the degree to which the parts of these stimuli were explicitly represented in participants' memories. For example, in one experiment, participants learned to name a set of faces (e.g., Joe or Larry), as well as a set of houses (e.g., Bill's house or Tom's house). They were then given two-alternative forced-

choice tests of the identity of isolated parts (e.g., "Which is Joe's nose?" or "Which is Bill's door?") or whole patterns in which the correct and incorrect choices differed only by a single part (e.g., "Which is Joe?" [when confronted with Joe and a version of Joe with the alternative nose from the isolated part test pair] or "Which is Bill's house?" [when confronted with Bill's house and a version of Bill's house with the alternative door from the isolated test pair]). It was found that, relative to their ability to recognize whole faces and houses, subjects were impaired at recognizing parts of faces as compared with parts of houses. Could the difference have been caused by the nature of the parts themselves? No, because the same pattern of results was obtained when faces were compared with scrambled faces and inverted faces, whose parts are identical. These results are consistent with the hypothesis that during the learning and subsequent recognition of the houses, scrambled faces, and inverted faces, participants explicitly represented their parts, whereas during the learning and subsequent recognition of the intact upright faces, they did not, or they did so to a lesser extent.

Tanaka and Sengco (1997) showed that these results should not be interpreted simply in terms of a part-based representation in which, for faces, the configuration of parts is particularly important. If this were the case, changes in configuration would affect overall face recognition, but so long as individual parts are explicitly represented, this manipulation should not affect recognition of the individual parts per se. Testing this prediction by comparing upright faces with inverted faces and houses, they again found evidence of holistic coding of upright faces.

This operationalization of holistic representation has also been used with members of two special populations. Tanaka, Kay, Grinnel, and Stansfield (in press) have shown that children as young as 6 years old show a disadvantage for isolated parts of faces, whether tested for their long-term memory for faces in the same way as the adult participants described earlier or tested with an immediate memory paradigm in which a pair of test faces or face parts is presented immediately after the study face. Farah, Tanaka, and Drain (unpublished data, described in Farah, 1996) tested a prosopagnosic participant (i.e., an individual who has lost the ability to recognize faces) on short-term memory for faces presented in a normal format or "exploded" so that the parts of the face were presented separately. Whereas normal participants performed better with the normal faces, presumably because they could encode them as wholes, the prosopagnosic participants performed roughly equivalently whether given the opportunity to encode the faces as wholes or forced to encode them as parts. This result is consistent with the hypothesis that the face recognition ability that has been lost in prosopagnosia involves the representation of faces as wholes.

Recent work by Moscovitch, Winocur, and Behrmann (1997) complements these findings by showing that a patient whose face recognition was disproportionately spared, relative to object recognition, could not recognize photographs of faces that had been cut into separate parts. The results of this and numerous other experiments with this patient led the authors to conclude that face and object recognition differ in their reliance on part representations, although their operationalization of "part" appears to include arbitrary fragments as well as structural components in a shape hierarchy (e.g., see Experiment 14).

Another way in which we have tested the holistic representation hypothesis is by determining whether it could explain the face inversion effect (Farah, Tanaka, & Drain, 1995). If face recognition differs from other forms of object recognition by the use of relatively undecomposed or holistic representations, then perhaps the face inversion effect results from the use of holistic, or non-part-based, representation. In the first experiment, we taught participants to identify random dot patterns and later tested their recognition of the patterns either upright or inverted. Half of the patterns learned by participants were presented in a way that encouraged parsing the pattern into parts: Each portion of the pattern corresponding to a part was made from a distinctive color, so grouping by color defined parts. The other half of the patterns learned were presented in all black, and the test stimuli for all patterns were presented in black. When participants had been induced to see the patterns in terms of parts during learning, their later performance in identifying the patterns showed no effect of orientation. In contrast, when they were not induced to encode the patterns in terms of parts, they showed an inversion effect in later recognition. In a second experiment, we manipulated participants' encoding of faces and then tested their ability to recognize the faces upright and inverted. They were induced to learn half of the faces in a partwise manner (in the "exploded" format described earlier), whereas the other half of the faces to be learned were presented in a normal format. All faces were tested in a normal format. For the faces that were initially encoded in terms of parts, there was no inversion effect. In contrast, faces encoded normally showed a normal inversion effect. These results suggest that what is special about face recognition, by virtue of which it is so sensitive to orientation, is that it involves representations with relatively little or no part decomposition.

## Are Perceptual Representations of Faces Holistic?

Each of these earlier lines of research was relevant to testing the hypothesis that faces are stored in a relatively holistic form in memory. Strictly speaking, they do not address the question of whether faces are perceived holistically. This question is of interest because some of the evidence reviewed earlier suggests that the "special" status of faces extends to the visual representations that are initially constructed during perception. For example, patients who are impaired at face recognition appear to perceive faces abnormally (Farah, 1990). In addition, the face inversion effect can be obtained in tasks that are free of any long-term memory component (e.g., matching of sequentially presented faces; Valentine, 1988). These considerations suggest that the memory representations of faces, studied in our previous research, are holistic because faces are initially encoded holistically during perception. However, the hypothesis that faces are perceived holistically has not been tested directly. That was the goal of the present experiments.

In these experiments, we assessed the degree of part decomposition on-line, during the perception of faces, using two types of experimental paradigms. In the first, we measured the relative availability of part and whole representations by requiring participants to compare single features of simultaneously presented pairs of faces and observed the influence of irrelevant features on their ability to judge the similarity or difference of the probed

feature. For example, they might be asked whether two faces have the same or different noses. To the extent that participants have explicit representations of the separate features of a face, they should be able to compare them with one another. To the extent that they do not have explicit representations of these features, but rather, only a holistic representation of the entire face, they should suffer cross talk from irrelevant features when judging the probed feature. The amount of cross talk with upright faces can be compared with that with inverted faces to measure the relative availability of parts and wholes in the representations of the two kinds of stimuli. In the subsequent three experiments, we explored the effect on face perception of masks composed of face parts or whole faces. As Johnston and McClelland (1980) reasoned in their experiments on word perception, to the extent that masks contain shape elements similar to those used in representing the stimulus, the mask will interfere with perception of the stimulus. The effects of part and whole masks on the perception of upright faces were compared with their effects on the perception of words, inverted faces, and houses.

## Experiment 1

The degree to which people explicitly represent the parts of faces during face perception was investigated with a *same–different* matching paradigm in which particular parts of the face were to be compared. On each trial, a pair of faces was briefly presented, one beside the other, followed immediately by the name of a facial part. Participants were to decide whether the two faces shared the same exemplar of the named part (e.g., the same nose). To the extent that the parts of the face are explicitly represented in immediate visual memory as units of shape in their own right, participants should be able to compare them with one another independently of the other separately represented face parts. To the extent that participants have only a unitary, undecomposed representation of each face, they should be able to judge only the overall similarity or difference between the faces, and their judgment of the similarity of any individual part will be contaminated by the amount of similarity between the other parts.

Of course, face representations are unlikely to be either pure holistic representations with no explicit part-level representations (especially in the context of the present task's demands) or pure collections of parts with no explicit whole-level representation. The present experiment was therefore designed to compare the relative contributions of part and whole representations with the perception of upright and inverted faces. That is, the prediction was neither that faces are perceived only as wholes nor that nonfaces are perceived only as parts; rather, a differential weighting of parts and wholes for faces relative to nonfaces was predicted. Because the hypothesis that face perception is holistic pertains only to the perception of upright faces, inverted faces provide control stimuli that are geometrically identical to upright faces (except for orientation), and therefore have the same partwise and holistic similarity relations, but are predicted to have less holistic perceptual representations.

### Method

*Participants.* Twenty-four undergraduate students at Carnegie Mellon University participated in exchange for course credit. All participants had normal or corrected-to-normal vision.

*Materials.* Six faces were created with the Mac-A-Mug (Macintosh) program. All of the internal features of the faces (eyes, nose, and mouth) were distinctive, and all of the external features (facial outline, chin, ears, and hair) were the same. Each of these faces was presented side by side with an identical copy of itself and with copies in which one or more features differed. Specifically, each face was paired with a face differing in eyes only, nose only, mouth only, eyes and nose, eyes and mouth, nose and mouth, and eyes, nose, and mouth. For each of the six original faces, there were 8 face pairs, 1 composed of an identical pair and 7 composed of pairs with one or more features changed as just described. This resulted in 48 different face pairs. Three of the original six faces appeared to the left of their different versions, and three appeared to the right. Each face measured 10.5 × 14.5 cm and subtended about 12° × 16.5° at the average participant viewing distance.

*Procedure.* Stimuli were presented, and responses collected, on a Macintosh II computer. On each trial, after the offset of a central fixation dot, a pair of faces was presented for 1 s, followed immediately by one of the words *eyes, nose, mouth,* or *all.* Participants were instructed to judge whether or not the named parts of the two faces were identical or, if probed with *all,* to judge whether the two faces were identical. Although we have no theoretically motivated predictions for the results of the *all* probe condition, we included it to weaken the implicit task demand to view the faces as collections of unrelated distinct parts (a task demand that might increase the likelihood of an artifactual null result). Participants responded using the computer keyboard, pressing "z" for *same* and "/" for *different.* Response latencies were not recorded. Approximately 12–16 practice trials were given.

The part probes followed face pairs that were either identical or differed by either just the named part or all parts. This resulted in equal numbers of trials on which the correct answer was *same* and *different* for each part probe, along with 72 trials with part probes. The *all* probe followed all pairs, resulting in fewer *same* than *different* trials by a 1:7 ratio, along with 48 trials with *all* probes. The combined 120 trials were presented once with the faces upright and once with the faces inverted; half of the participants performed the task with the upright faces first, and half performed it with the inverted faces first. There was a 5-min break between the two blocks of trials.

### Results and Discussion

The results of the part probe conditions are of primary interest, because they are relevant to confirming or disconfirming the holistic perception hypothesis. Thus, they are discussed first. As can be seen in Table 1, the compatibility of the probed and irrelevant parts had an effect on response accuracy, and the effect was larger for upright faces than for inverted faces. For upright faces, when the irrelevant parts were compatible with the probed part (i.e., the probed part was the same and the irrelevant features were also the same, or the probed part was different and the irrelevant features were also different), participants achieved an average rate of 74.6% correct; when the irrelevant parts were not compatible with the probed part (i.e., the probed part was the same and irrelevant parts were different, or the probed part was different and the irrelevant parts were the same), participants achieved only 61.8% correct. The corresponding means for the inverted faces were 66.6% and 62.2% correct. This pattern of results was predicted by the hypothesis that face perception is holistic. Figure 1 shows the means of the

Table 1
*Percentage Correct Facial Matching in Experiment 1 for Upright and Inverted Faces: Same or Different Probed Parts and Compatible or Incompatible Nonprobed Parts*

| | Upright | | Inverted | |
|---|---|---|---|---|
| Probed feature | Irrelevant features compatible | Irrelevant features incompatible | Irrelevant features compatible | Irrelevant features incompatible |
| Same | 91.5 | 71.6 | 86.6 | 74.7 |
| Different | 57.6 | 51.9 | 46.5 | 49.6 |

compatible and incompatible conditions for upright and inverted faces. Figure 2 shows the effect of compatibility for upright and inverted faces.

A planned comparison was carried out to test the prediction that compatible irrelevant parts will facilitate performance relative to incompatible parts and that this effect will be greater for the upright faces than for the inverted faces. The corresponding weights for the upright and inverted faces with compatible and incompatible irrelevant parts were set at .67, −.67, .33, and −.33, respectively, yielding a value of $F(1, 23) = 35.10$, $p < .0001$.

A repeated measures analysis of variance was performed on participants' percentage correct rates in each of the part probe conditions of interest; the following crossed variables were used: probed part (same or different), irrelevant part (compatible or incompatible with probed feature), and orientation (upright or inverted). The hypothesis that face perception is relatively holistic predicts that the compatibility of the irrelevant parts will have a larger effect on judgments about upright faces than on judgments about inverted faces. The corresponding interaction, between compatibility and orientation, was significant, $F(1, 23) = 6.85$, $p < .02$. In addition, there were three significant main effects and one other significant interaction. *Same* responses were more accurate than *different* responses (81.1% vs. 51.4% correct), $F(1, 23) = 106.94$, $p < .001$; performance was more accurate with compatible than incompatible irrelevant parts (70.6% vs. 62% correct), $F(1, 23) = 20.80$, $p < .001$; upright faces were compared more accurately than inverted faces (68.2% vs. 64.4% correct), $F(1, 23) = 4.23$, $p < .05$; and the effect of compatibility was greater for *same* trials than for *different* trials (89.1% and 73.2% correct for compatible and incompatible *same* trials and 52.1% and 50.8% correct for compatible and incompatible *different* trials), $F(1, 23) = 29.05$, $p < .001$. No other effects were significant ($p$s > .1 in all cases). Separate analyses of variance testing the simple effects of compatibility on upright and inverted faces revealed a significant effect for the upright faces, $F(1, 23) = 31.92$, $p < .0001$, and a borderline significant effect for inverted faces, $F(1, 23) = 3.79$, $p = .064$.

Examination of the means in Table 1 suggests that participants were more likely to respond "same" than "different" in this experiment, and indeed accuracy on *different* trials was at chance. We therefore repeated the critical planned comparison with only the data from *same* trials and confirmed that the predicted pattern of results was obtained with this subset of the data, $F(1, 23) = 52.88$, $p < .0001$.

In the *all* probe condition, participants were to respond "same" only if all of the parts of the faces matched. Mean percentage correct rates for the conditions of interest are shown in Table 2. The only significant effect was that of response, with

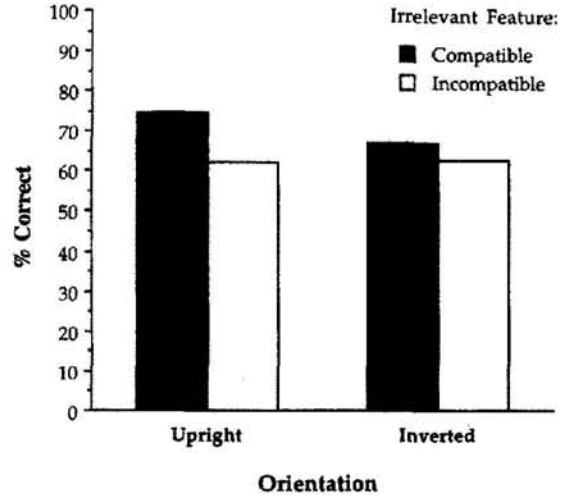


*Figure 1.* Percentage correct facial part matching in Experiment 1 for upright and inverted faces as a function of the compatibility between the probed parts and the nonprobed, or irrelevant, parts.

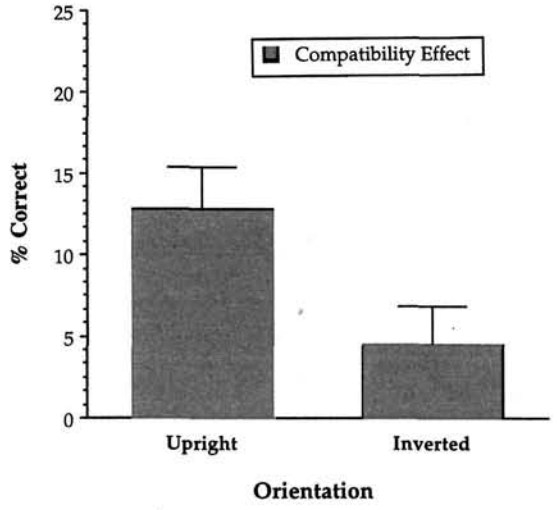


*Figure 2.* Effect of compatibility for upright and inverted faces in Experiment 1.

Table 2
*Percentage Correct Whole Face Matching in Experiment 1 for Upright and Inverted Faces: Same or Different Faces*

| Whole face | Upright | Inverted |
|---|---|---|
| Same | 80.6 | 82.6 |
| Different | 69.1 | 62.8 |

*same* responses more accurate than *different* responses (81.6% vs. 66.0% correct), $F(1, 23) = 15.04$, $p < .001$. No other effects were significant ($ps > .1$).

The prediction of the holistic face perception hypothesis was borne out by the results of this experiment. We reasoned that if perceived faces were represented as undifferentiated whole shapes, without explicitly represented parts, then it would be difficult for participants to compare parts independent of the whole. We found that this was indeed the case when the difficulty of comparing parts of upright faces was compared with the difficulty of comparing the same parts of the same faces inverted, which do not engage face-specific representations (or do so to a lesser extent). Of course, this result is susceptible to the alternative explanation that both upright and inverted faces are represented to the same degree in terms of parts but that, for some reason, the individual parts are less accessible when faces are upright. However, in the absence of an independent reason to believe that there is differential part accessibility, rather than differential representation of parts, with upright faces, the holistic perception hypothesis provides the most straightforward account of these data.

This experiment tested the holistic perception hypothesis with the representations of faces in immediate visual memory. Participants could begin their comparisons of the faces as soon as they read the probe words, which appeared at the same instant that the faces disappeared. In Experiment 2, we manipulated perceptual encoding per se using different kinds of pattern masks, thereby studying the roles of part and whole representations in the initial construction of face representations during perception. This increases the number of qualitatively different experimental paradigms that have been used to test the hypothesis of holistic face representation and, in particular, helps to rule out the alternative hypothesis that parts are represented but merely less available for access during explicit comparisons. It also provides the most direct test of the holistic representation hypothesis as applied to perception.

## Experiment 2

One way of stating the holistic representation hypothesis is that, in the course of constructing a perceptual representation of a face, it is not necessary to construct explicit representations of its parts. This formulation of the issue highlights its similarity to an issue addressed by Johnston and McClelland (1980) in their research on word recognition. They hypothesized that word recognition involves a hierarchical recognition process in which letters are explicitly recognized before words. They tested this hypothesis by assessing the relative effects of different kinds of masks on word perception. Although they interpreted their results using a detailed quantitative model, the qualitative gist of their reasoning was as follows: If the explicit representation of letters is a necessary stage leading up to word recognition, then masks made up of letters should have a more detrimental effect on word recognition than masks made up of letter fragments. Their results supported this prediction.

We used a simple version of the Johnston and McClelland paradigm to address the question of whether face recognition requires explicit representation of facial parts. If faces are recognized as a whole and part representation plays a relatively small role in face recognition, then a mask made of face parts should be less detrimental than a mask consisting of a whole face. In this experiment, we used words as the contrasting nonface stimulus set, because Johnston and McClelland's earlier work had suggested that they are recognized via part representations.

Participants performed a sequential same–different matching task in which a stimulus was presented briefly, followed by a mask, a blank interstimulus interval, and a second stimulus. By varying the nature of the mask, we could assess the effects of interfering with part and whole representations on the quality of perception of the first stimulus. Part and whole word masks consisted of nonwords and words, respectively, as in the original experiments of Johnston and McClelland. These masks were composed of letters not present in the word stimuli, and nonword masks were scrambled versions of the word masks. Examples are shown in the Appendix. By analogy, part and whole face masks were composed of facial parts not used in the stimulus faces, and part face masks were scrambled versions of whole face masks, as shown in Figure 3. The question of whether the part face masks are equivalent to the part word masks for purposes of this experiment is discussed in the *Results and Discussion* section.

### Method

*Participants.* Sixteen students from Carnegie Mellon University and the University of Pennsylvania were paid for their participation in this study. All participants had normal or corrected-to-normal vision. Five were replaced because their performance in one or more conditions was below 55% correct or above 95% correct.

*Materials.* We created a set of 36 four-letter words using 14 target letters (*c, d, e, f, h, j, l, m, o, p, s, t, u,* and *z*) in 18-point Courier font. These words were 3 cm long and subtended an angle of about 3.5° at the average participant viewing distance. Each of the words was paired with a copy of itself and with another word. The other word differed by a single letter (e.g., *most* was changed to *must*), thus producing 36 same-word pairs and 36 different-word pairs. The changed letter occurred equally often in each of the four letter positions. Nine 4-letter word masks, or whole masks, were created from 11 mask letters (*a, b, g, i, k, n, r, v, w, x,* and *y*). As in Johnston and McClelland (1980), *q* was not used in any of the target or mask words. Part masks were created by scrambling the positions of the letters within each of the whole masks, resulting in nine whole masks and nine part masks. Examples are shown in the Appendix.

A set of 36 faces was created with the Mac-A-Mug software. The faces measured 10.5 × 14.5 cm and subtended about 12° × 16.5° at the average participant viewing distance. Each of these faces was paired with a copy of itself and with a similar-appearing but different face to create 36 same-face and 36 different-face pairs. As in the previous experiment, faces shared the same facial outline (hair, ears, chin, and shoulders); only the internal features (eyes, nose, and mouth) differed

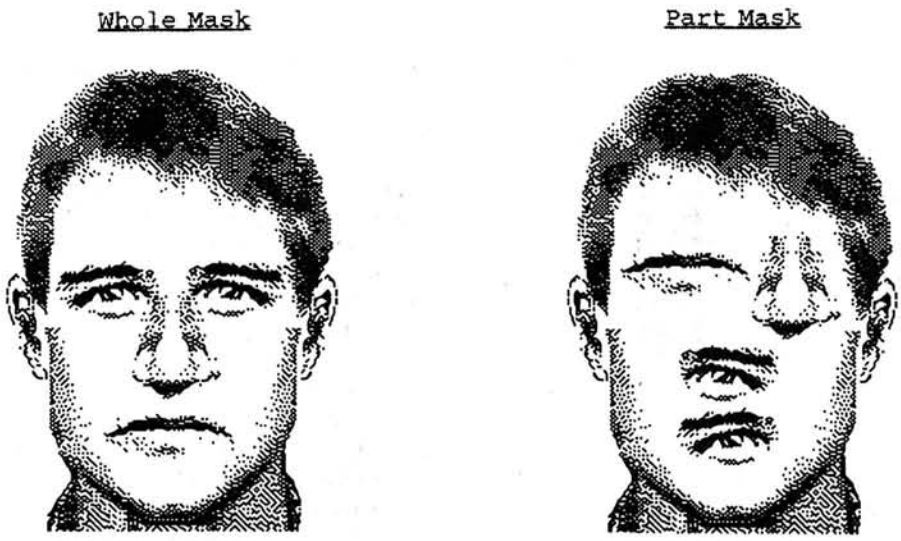


*Figure 3.* Examples of whole and part masks for face stimuli.

between faces. Whereas *different* word pairs differed by only one part, *different* face pairs differed by all three internal parts so as to be roughly equivalent in difficulty. Subsequent experiments provide evidence that the nature of *different* pair construction is not responsible for the results to be reported. Nine additional faces, different from those used in the stimuli pairs, were used as whole masks. As with the word masks, part masks for faces were created by scrambling the positions of the parts within the face (e.g., placing the mouth in the normal position of the nose, the eyes in the normal position of the mouth, and the nose in the normal position of the eyes). An example is shown in Figure 3.

For each set of stimuli, words and faces, the 36 basic items were repeated four times, twice paired with their corresponding *different* version and twice paired with themselves for *same* trials. This resulted in 144 word trials and 144 face trials, which were blocked. Whole and part masks were randomly assigned to trials. Trial order was randomized with the restriction that no more than three consecutive trials could include the same type of mask or require the same response.

*Procedure.* A Macintosh II computer was used to present stimuli and record responses. The contrast on the Macintosh's monitor was set to approximately one half full brightness and two thirds full contrast. All trials began with the presentation of a fixation dot for 500 ms, followed by the first stimulus of a pair. After an interstimulus interval of 100 ms, a mask was exposed for 300 ms. After a 2-s delay, the second stimulus of the pair was presented. In setting exposure durations for the two types of stimuli, we were forced to choose between equating accuracy and equating exposure duration. Because sensitivity to manipulations of perceptual difficulty is known to depend on overall performance levels (with the greatest sensitivity at levels of performance that are intermediate between poor and excellent), we chose to equate accuracy at the expense of equating exposure duration. (Note that both accuracy and exposure duration were equated for upright and inverted faces in subsequent experiments.) The exposure duration for the first and second word stimuli was 17 ms, and the duration for first and second face stimuli was 200 ms. Participants responded using the computer keyboard, pressing "z" for *same* and "/" for *different*. Both responses and response latencies were recorded. After each response, a "Ready?" prompt appeared, and participants, when ready, initiated the next trial by pushing the space bar. Twelve practice trials, with equal numbers of *same*, *different*, part, and whole mask trials, were included, and feedback concerning accuracy was given. Half of the participants performed the block of face trials first, and the other half performed the block of word trials first. There was a 5-min rest break between blocks.

### *Results and Discussion*

We focus our attention on the analysis of accuracy, because error rates are high for conventional reaction time analysis. We therefore begin our discussion of the results with the accuracy measures. As can be seen in Table 3 and Figure 4, whole masks interfered more with perception of faces than part masks (73.1% vs. 77.6% correct), but there was little difference apparent in word perception (77.0% vs. 77.8% correct for the corresponding conditions). Figure 5 shows the effect for type of mask on word and face matching accuracy.

A planned comparison was carried out to test the prediction that whole masks would be more disruptive than part masks and that this difference would be greater for face perception than word perception. The corresponding weights for the whole and

Table 3
*Face and Word Matching in Experiment 2 With Part or Whole Masks: Same or Different Stimuli*

| | Faces | | Words | |
|---|---|---|---|---|
| Mask type | Same | Different | Same | Different |
| | Percentage correct | | | |
| Part | 80.7 | 74.5 | 75.7 | 79.9 |
| Whole | 77.6 | 68.6 | 74.5 | 79.5 |
| | Response time (ms) | | | |
| Part | 1,123.6 | 1,100.5 | 1,011.4 | 1,228.0 |
| Whole | 1,184.0 | 1,146.1 | 1,042.5 | 1,225.4 |

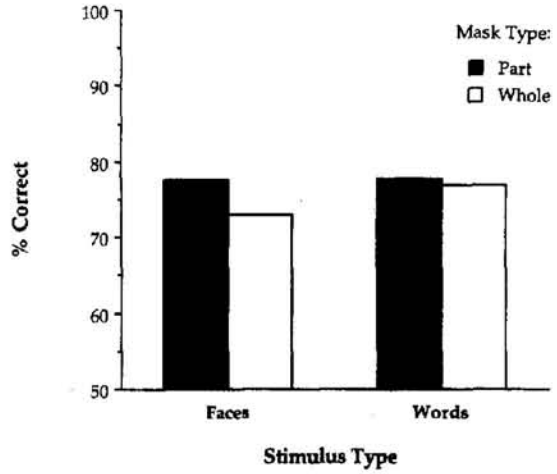


*Figure 4.* Percentage correct face and word matching in Experiment 2 with whole and part masks.

part face mask and whole and part word masks were set at −.67, .67, −.33, and .33, respectively, yielding a value of $F(1, 15) = 11.58$, $p < .005$.

A repeated measures analysis of variance was conducted on participants' number correct performance in each of the conditions of interest, and the following crossed variables were used: stimulus type (word or face), mask type (part or whole), and response type (same or different). There was a significant interaction between stimulus type and response, with higher accuracy for *different* than for *same* word pairs and the opposite trend with faces, $F(1, 15) = 6.31$, $p < .05$. Participants achieved average rates of 75.1% and 79.7% correct for *same* and *different* word pairs and 79.2% and 71.5% correct for the corresponding face pairs. The interaction between stimulus type and mask type narrowly missed the .05 significance level, $F(1, 15) = 4.20$, $p = .058$. Finally, the effect of mask type was of borderline significance, with whole masks being more detrimental to performance, $F(1, 15) = 3.94$, $p = .066$. No other effects were significant ($ps > .1$ in all cases). Separate analyses of variance were carried out to test the simple effects of mask type on face and word matching. The effect on face perception was significant, $F(1, 15) = 6.58$, $p < .05$, and the effect on word perception was not, $F(1, 15) = 0.29$, $p > .1$.

Response times were also analyzed after removal of response times from incorrect trials and those that were more than 3 standard deviations above the mean of other reaction times in the same cell for the participant. The same planned comparison was performed on the response time data, and a significant result was obtained, $F(1, 15) = 5.89$, $p < .05$. With part and whole face masks, participants responded in an average of 1,112 and 1,165 ms; with part and whole word masks, the corresponding means were 1,120 and 1,134 ms. In a repeated measures analysis of variance with the same variables used in the analysis of accuracy data, only the interaction between stimulus type and response was significant; participants responded to *same* and *different* word pairs in 1,027 and 1,227 ms and to *same* and *different* face pairs in 1,154 and 1,123 ms, $F(1, 15) = 11.26$, $p < .005$. No other effects were significant ($ps > .1$ in all cases). As with the accuracy data, separate analyses of variance were performed on the reaction time data for the word and face conditions to assess the simple effects of mask type on each type of stimulus. Neither effect was significant ($ps > .1$ in both cases).

Although there was a trend for the perception of both types of stimuli to be more impaired by whole masks than by part masks, this trend was much more pronounced for faces than for words. This is consistent with the hypothesis that there is less need for the parts of the face to be explicitly represented during face perception than there is for the parts of words to be represented during word perception. We now consider some alternative explanations for this result.

The face and word stimuli differed from one another in a variety of ways other than their identities as faces and words. For example, the words were smaller than the faces. As noted in the *Materials* section, the way in which *different* stimuli were derived from *same* stimuli differed. To obtain comparable levels of accuracy for faces and words, we used different exposure durations. Perhaps the greater susceptibility of faces to whole masking results not from a basic difference in the nature of face and word representation but, rather, from their larger size, more distributed differences, or longer exposures. Other differences between the two sets of stimuli in the present experiment pertain to the nature of their part masks. Whereas the parts of the word part masks were perfectly superimposed on the parts of the word they masked, the parts of the face part masks were not. For example, even when a mouth is centered on a nose, some of the nose is left unmasked. Also, whereas each of the letters in the word part masks occurs in a realistic position (e.g., there are four-letter words that begin with the letter *a*, as shown in the first part word mask in the Appendix), the eyes, noses, and mouths of the face part masks occur in unrealistic positions. Perhaps the relatively greater effect of whole masks for faces is due to the incomplete masking of the face part masks or their more artificial nature relative to the word part masks.

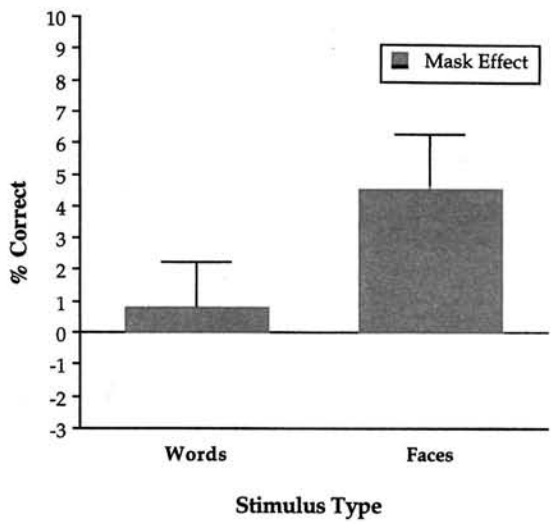


*Figure 5.* Effect of part versus whole masking on accuracy of matching faces and words in Experiment 2.

Table 4
*Upright Face, Inverted Face, and Word Matching in Experiment 3 With Part or Whole Masks: Same or Different Stimuli*

| | Upright faces | | Inverted faces | | Words | |
|---|---|---|---|---|---|---|
| Mask type | Same | Different | Same | Different | Same | Different |
| | | | Percentage correct | | | |
| Part | 79.0 | 80.3 | 78.4 | 72.1 | 72.8 | 88.2 |
| Whole | 73.1 | 77.8 | 80.5 | 68.4 | 71.2 | 89.0 |
| | | | Response time (ms) | | | |
| Part | 1,054.5 | 1,029.2 | 1,094.6 | 1,022.1 | 932.0 | 1,084.7 |
| Whole | 1,066.6 | 1,060.0 | 1,080.2 | 1,101.1 | 992.2 | 1,132.6 |

These possibilities can all be addressed by comparing the effects of part and whole masks on the perception of upright and inverted faces. Inverted faces are equivalent to upright faces in their size, in the nature of their "different" stimuli, and in their exposure duration. They are also subject to the same amount of overlap with inverted part masks as upright faces are with upright part masks, and the inverted part face masks are similarly unrealistic.

## Experiment 3

In this experiment, we replicated Experiment 2 with words and upright faces and also included inverted faces. If the differences between the effects of part and whole masks on face and word perception are due to the kinds of low-level differences between the face and word stimuli discussed earlier, then the effects of masking inverted faces should be similar to the effects of masking upright faces. In contast, if the greater effect of whole masking on upright faces is due to face-specific perceptual representations, then inverted faces should not be disproportionately sensitive to whole masks.

### *Method*

*Participants.* Twenty-four new participants were recruited for this study. All were undergraduate students from the University of Pennsylvania who were paid for their participation. Participants had normal or corrected-to-normal vision. Seven were replaced because their performance in one or more conditions was either below 55% correct or above 95% correct.

*Materials.* The materials were the same as in the previous experiment, with the exception of a new block of trials created by inverting the same face stimuli and masks used in the upright face condition.

*Procedure.* The procedure was the same as that used in the first experiment, with the addition of a third block of trials for the inverted face condition. The order of blocks was counterbalanced over participants so that each block occurred equally often in each ordinal position. In addition, participants completed 20 practice trials for each block, and we used their performance during practice to adjust the brightness of the monitor so as to equate as much as possible the difficulty of the three blocks.

### *Results and Discussion*

Table 4 and Figure 6 show the mean accuracy of participants in the conditions of interest. As in the previous experiment, whole masks interfered more with perception of faces than part masks (75.6% vs. 79.7% correct), but there was again little difference apparent in word perception (80.1% and 80.5% correct for the corresponding conditions). In addition, inverted faces failed to show a large difference between whole and part masks (74.5% and 75.3% correct, respectively). This is consistent with the hypothesis that the difference observed with upright faces is indicative of a form of face-specific representation. Although upright and inverted faces share low-level perceptual features, inverted faces do not engage specialized face perception mechanisms or do so to a lesser extent than upright faces. Figure 7 shows the effects of mask type in the three conditions of interest.

A planned comparison was carried out to test the prediction that whole masks are more disruptive than part masks and that this difference will be greater for upright face perception than for inverted face or word perception. The corresponding weights for the whole and part upright face masks, whole and part inverted face masks, and whole and part word masks, derived from the ratios of the means of Experiment 2, were −.74, .74, −.13, .13, −.13, and .13, respectively, yielding a value of $F(2, 46) = 9.21$, $p < .005$.

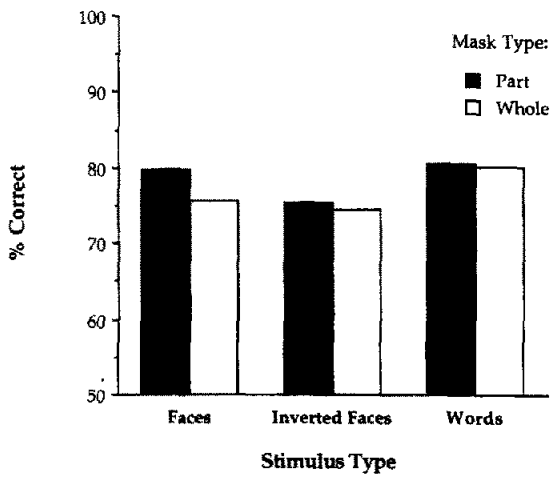


*Figure 6.* Percentage correct upright face, inverted face, and word matching in Experiment 3 with whole and part masks.

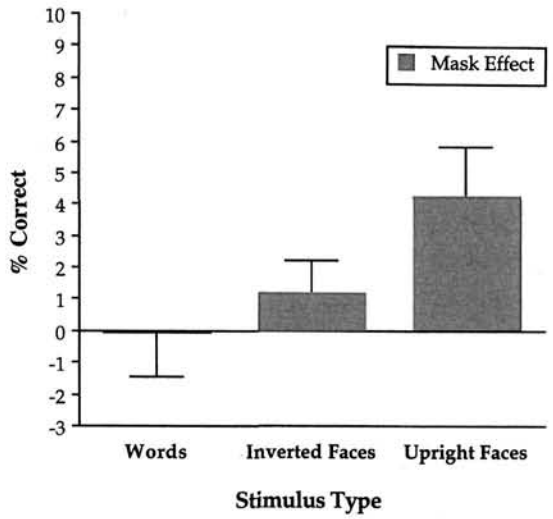


*Figure 7.* Effect of part versus whole masking on accuracy of matching faces, words, and inverted faces in Experiment 3.

A repeated measures analysis of variance was conducted on participants' number correct performance in each of the conditions of interest; the following crossed variables were used: stimulus type (word, inverted face, or upright face), mask type (part or whole), and response type (same or different). There was again a significant interaction between stimulus type and response, although the form of the interaction differed from that observed in Experiment 2, with lower accuracy for *same* than *different* word pairs (72.0% vs. 88.6% correct) and faces (76.1% vs. 79.0% correct) and inverted faces showing a pattern of higher accuracy for *same* than *different* pairs (79.5% vs. 73.0% correct), $F(1, 15) = 32.34$, $p < .001$. We cannot offer any hypotheses to explain this pattern or the difference between the present outcome and the outcome of Experiment 2. However, note that the response times showed a different pattern (described subsequently). Mask type had a significant effect, with higher accuracy for part than whole masks (78.5% vs. 76.7%), $F(1, 23) = 8.95$, $p < .01$. Finally, two effects were of borderline significance: *Different* responses were more accurate than *same* responses, $F(1, 23) = 3.16$, $.05 < p < .1$, and there was a trend toward a three-way interaction among stimulus type, mask type, and response type, $F(2, 46) = 2.63$, $.05 < p < .1$. No other effects, including the overall interaction between mask type and stimulus type, were significant ($ps > .1$ in all cases). Separate analyses of variance were carried out to assess the simple effects of mask type on each of the stimulus types. Mask type had a significant effect on face matching accuracy, $F(1, 23) = 7.04$, $p < .02$, but not on word matching or inverted face matching ($ps > .1$ in both cases).

Response times were analyzed as before. Initially, a planned comparison on the six relevant means was carried out, as with the accuracy data from this experiment, but this failed to produce significant results, $F(2, 46) = 2.41$, $p > .1$. A repeated measures analysis of variance was also carried out, with the same variables used in the analysis of accuracy data. Participants were faster with part than with whole masks (1,036 vs. 1,072 ms), $F(1, 23) = 4.40$, $p < .05$. Also, there was an interaction between stimulus type and response type, with *same* words being particularly fast (1,061 and 1,045 ms for *same* and *different* faces, 1,087 and 1,062 ms for *same* and *different* inverted faces, and 962 and 1,109 ms for *same* and *different* words), $F(1, 23) = 10.45$, $p < .001$. There was a trend of borderline significance for faster responses to *same* than to *different* trials, $F(1, 23) = 3.65$, $.05 < p < .1$, which was opposite to the borderline significant trend in accuracy for better performance on *different* trials. No other effects were significant ($ps > .1$ in all cases). Separate analyses of variance on matching reaction time for faces, words, and inverted faces failed to reveal any significant effects ($ps > .1$ in all cases).

In sum, the predicted pattern of results was obtained in the accuracy data of this experiment. Specifically, the perception of upright faces was more disrupted by the use of a whole mask than by the use of a part mask. In contrast, neither words nor inverted faces showed this pattern. The difference in efficacy of part and whole masking for upright and inverted faces rules out the alternative hypothesis that part face masks are ineffective because of imperfect superposition of patterned regions in the mask and stimulus.

## Experiment 4

The previous two experiments used the relative effects of part and whole masks to measure the degree of holistic representation in face perception and compare it with the degree of holistic representation in the perception of words and inverted faces. The goal of Experiment 4 was to extend this contrast to a kind of concrete object viewed in a normal orientation. We chose houses as our contrast objects.

On the basis of neuropsychological dissociations among disorders of face, word, and object recognition, we have previously hypothesized that objects are intermediate between faces and words in their reliance on holistic representations (Farah, 1991). In the present experiment, we therefore predicted that the difference in efficacy between part and whole masks would be larger for faces than for houses and larger for houses than for words.

### *Method*

*Participants.* To resolve the smaller differences predicted to be found between faces and houses than between faces and words or faces and inverted faces, we increased by 50% the number of participants included in this study. Thirty-six undergraduate students from the University of Pennsylvania were paid for their participation. Participants had normal or corrected-to-normal vision. Six were replaced because their performance in one or more conditions fell below 55% correct or above 95% correct.

*Materials.* The word and face materials from Experiment 2 were used again. In addition, a set of 36 houses was created with architectural design software for Macintosh. Each of these houses was paired with a copy of itself and with a similar-appearing but different house to create 36 *same* and 36 *different* pairs. All houses shared the same external frame, but the internal features (door, bay window, and second story windows) differed between houses. Nine additional houses, different from those used in the stimulus pairs, were used as whole masks. As a means of creating part masks, the locations of the windows and door were scrambled, with each part in another part's position. This procedure is exactly analogous to the procedure for making part masks for the faces. Figure 8 shows typical whole house and part house masks.

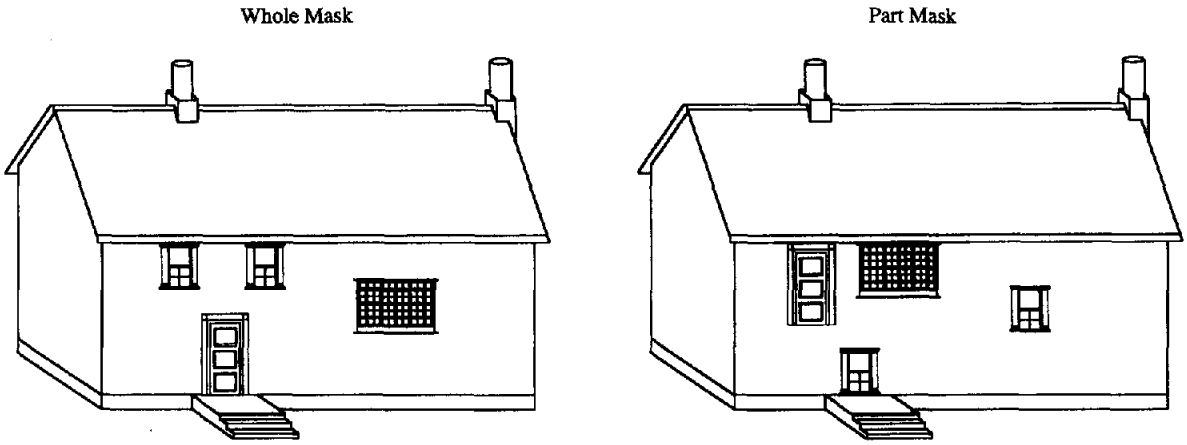


*Figure 8.* Examples of whole and part masks for house stimuli.

*Procedure.* The procedure was the same as that used in Experiment 3, with houses replacing inverted faces.

### *Results and Discussion*

Table 5 and Figure 9 show the mean accuracy of participants in the conditions of interest. As in the previous experiments, whole masks interfered more with perception of faces than part masks (75% vs. 78.5% correct), and there was again little difference apparent in word perception (79.2% and 78.0% correct for the corresponding conditions). In addition, houses showed an intermediate-sized difference between whole and part masks (83.2% vs. 85% correct). This is consistent with the hypothesis that face perception is more holistic than house perception and house perception is more holistic than word perception. Figure 10 shows the mask effect for the three types of stimuli.

A planned comparison was carried out to test the prediction that whole masks are more disruptive than part masks and that this difference will be greatest for face perception, intermediate for house perception, and smallest for word perception. The corresponding weights for the whole and part upright face masks, whole and part house masks, and whole and part word masks, derived from the mean ratios of Experiment 2 with the assumption that houses would be intermediate between faces and words, were −.567, .567, −.333, .333, −.100, and .100, respectively, yielding a value of $F(1, 70) = 15.82$, $p < .0005$.

A repeated measures analysis of variance was conducted on participants' number correct performance in each of the conditions of interest, and the following crossed variables were used: stimulus type (word, house, or face), mask type (part or whole), and response type (same or different). There was a significant effect of stimulus type, with highest accuracy for houses (84.1%), followed by words (78.6%) and faces (76.7%), $F(2, 35) = 13.29$, $p < .0001$. There was also a significant effect of mask type, with slightly better performance with part masks than whole masks (80.5% vs. 79.1% correct), $F(1, 35) = 7.10$, $p < .02$. Overall, performance on *different* pairs was more accurate than performance on *same* pairs (81.5% vs. 78.1% correct), $F(1, 35) = 5.32$, $p < .05$. The interaction of interest, between stimulus type and mask type, was significant, $F(2, 70) = 6.77$, $p < .005$. There was also a significant interaction between stimulus type and response type, with *same* trials less accurate than *different* trials for words (70.3% vs. 86.9% correct) and *same* trials more accurate for both houses (86.9% vs. 81.3% correct) and faces (77.2% vs. 76.3% correct), $F(2, 70) = 33.10$, $p <$

Table 5
*Face, House, and Word Matching in Experiment 4 With Part or Whole Masks: Same or Different Stimuli*

| | Faces | | Houses | | Words | |
|---|---|---|---|---|---|---|
| Mask type | Same | Different | Same | Different | Same | Different |
| | | | Percentage correct | | | |
| Part | 78.5 | 78.6 | 87.4 | 82.5 | 70.2 | 85.7 |
| Whole | 75.9 | 74.0 | 86.3 | 80.1 | 70.4 | 88.1 |
| | | | Response time (ms) | | | |
| Part | 1,207.0 | 1,106.8 | 1,187.9 | 1,092.4 | 1,105.0 | 1,199.9 |
| Whole | 1,190.7 | 1,155.3 | 1,208.5 | 1,103.3 | 1,173.4 | 1,245.7 |

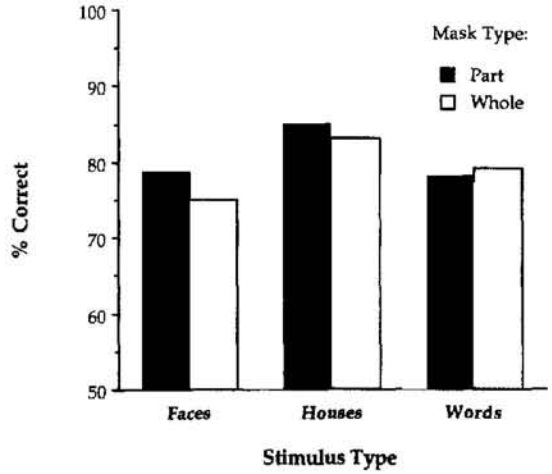


*Figure 9.* Percentage correct face, house, and word matching in Experiment 4 with whole and part masks.

.0001. No other effects were significant ($ps > .1$ in all cases). As usual, separate analyses of variance were performed to assess the simple effects of mask type on each type of stimulus. Both face and house perception were significantly affected by mask type, $F(1, 35) = 12.11$, $p < .002$, and $F(1, 35) = 4.56$, $p < .05$, respectively, but word perception was not ($p > .1$).

The data from just the house and face conditions were also analyzed separately to allow a direct comparison between the effects of part and whole masking on faces and on another type of concrete object in a normal orientation. A planned comparison was carried out with the following weights for part and whole masked faces and houses, respectively: .67, −.67, .33, and −.33. This yielded $F(1, 35) = 29.83$, $p < .0001$. A repeated measures analysis of variance was also carried out, revealing significant effects of stimulus type, $F(1, 35) = 34.76$, $p < .0001$, and mask type, $F(1, 35) = 11.86$, $= p = .002$. There was a borderline significant effect of response type, $F(1, 35) = 3.29$, $p = .08$. Finally, the interaction between stimulus type and mask was also of borderline significance, $F(2, 35) = 3.15$, $p = .08$. No other effects were significant ($ps > .1$ in all cases).

Response times were analyzed as before. Initially, a planned comparison on the six relevant means was carried out, as with the accuracy data from this experiment, $F(2, 70) = 2.30$, $p > .1$. A repeated measures analysis of variance was also carried out, with the same variables used in the analysis of accuracy data. There was a significant effect of mask type, part masks allowing faster responses than whole masks (1,150 vs. 1,179 ms), $F(1, 35) = 5.15$, $p < .05$. There was also a significant interaction between stimulus type and response type, with *same* responses faster than *different* responses for words (1,139 vs. 1,223 ms) and the opposite for both faces (1,199 vs. 1,131 ms) and houses (1,198 vs. 1,098 ms), $F(2, 70) = 11.02$, $p < .0001$. No other effects were significant ($ps > .1$ in all cases). Separate analyses of variance revealed no significant effects of mask type for face and house perception ($ps > .1$ in both cases) and a borderline significant effect of mask type for word perception, $F(1, 35) = 3.71$, $p = .06$.

## General Discussion

The present experiments add to a growing body of evidence in cognitive psychology suggesting that faces are represented holistically (i.e., with little or no part decomposition) relative to objects and patterns other than faces. Previous research has compared face representation with the representation of scrambled faces, inverted faces, and houses and assessed the role of parts versus holistic representation in both long-term and short-term memory for faces. It has also tested the generality of holistic face representation across developmental stages and showed that an individual with a neurological impairment in terms of face recognition did not benefit from the opportunity to represent faces holistically. Finally, previous research has also obtained an inversion effect with dot patterns, but only if they were initially encoded holistically, and shown that the face inversion effect can be eliminated for face stimuli if the faces were initially encoded in terms of separate parts. In sum, memory for faces has been shown to be holistic in the context of a number of different experimental paradigms and relative to a number of different comparison stimuli.

In the present studies, we showed that faces are represented more holistically than other stimuli in immediate perceptual memory and during perception. Comparison stimuli included inverted faces, words, and houses. In the remainder of this article, we discuss the relation of our hypothesis to previous claims about face representation and to perspectives on face representation derived from neurophysiology and computational vision. Also, we consider the question of how uniquely "special" is face representation.

### *Holistic Face Representation and Earlier Claims*

As we noted earlier, our hypothesis has much in common with earlier ones about face representation, in that it gives special importance to the overall structure or gestalt of the face relative to local features. However, each of the hypotheses is reasonably

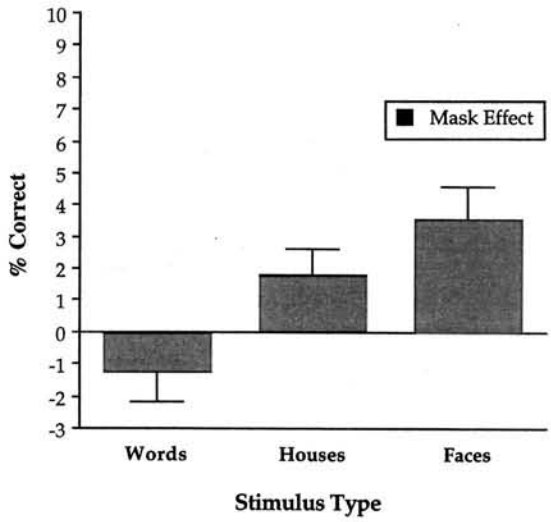


*Figure 10.* Effect of part versus whole masking on accuracy of matching faces, houses, and words in Experiment 4.

distinct in terms of the way in which information about overall structure is represented. In the case of our hypothesis, faces are represented holistically, which we define as meaning without explicitly representing (or relying to a lesser degree on explicit represention of) the local features themselves. In the framework of structural descriptions, we hypothesize that there is relatively little part decomposition for faces relative to other objects or patterns.

An undecomposed representation is essentially a template. In measuring the fit of a stimulus to a template representation, it is the overall best fit that is important rather than some summation of the fits of particular regions of the input pattern with explicitly defined parts of the internal representation. This is not to say that different regions of the template might not be differentially weighted in computing fit; rather, their heavier contribution to the fit equation is not a function of part identity per se but spatial location within the overall pattern. Consistent with this, Tanaka and Farah (1993) found that the eye region was more heavily weighted than others in people's judgments of facial identity only when the face was upright and intact and that eyes per se showed no special weighting in the inverted and scrambled faces conditions.

How does this hypothesis differ from earlier claims, and how well does it account for the data marshalled in support of earlier claims? The contrast between the idea of holistic face representation and the parallel processing hypothesis of the 1970s resides in the role of part-based representation: Whereas part-based representation is relegated to a lesser or possibly even nonexistent role according to the holistic face perception hypothesis, the parallel processing hypothesis maintains that the representation of faces is as part based as the representation of other objects but that these psychologically real parts are processed in parallel rather than serially. As for the evidence showing serial processing of parts, it is possible that this results from a strategy specific to the experimental task, as Bradshaw and Wallace (1971) themselves suggested.

The claim put forth by Rhodes (1988), that configurational features are important in face representation, differs more subtly from our hypothesis. Representations that include configurational features such as distances between first order features will behave in many ways like holistic face representations, in that a holistic representation of a face with particular features in a particular spatial arrangement will contain the information coded in the configurational feature for that spatial arrangement. However, the difference is that the features, first order and configurational, are psychologically real or explicit according to Rhodes's hypothesis, whereas in ours they are not. One could, of course, extract such features from a holistic representation, and in this sense holistic representations implicitly contain both first order and configurational features. To make an analogy with early vision, the retinal image implicitly contains information about edges without explicitly representing them. The finding that shared first and second order features are predictive of participants' similarity ratings does not imply that participants explicitly represent them. To continue the analogy, similarity, as measured by overlap of retinal images, will be predicted by similar edge maps. Therefore, the results of Rhodes (1988) are not inconsistent with holistic face representation.

Sergent's (1984) claim that the features of faces are not processed independently is essentially a claim about the stimulus properties of faces that are predictive of human sorting and similarity-rating behavior, following Garner's (1974) taxonomy of stimulus properties. Her conclusion puts constraints on possible psychological representations but is not itself a specific claim about representation. As with the Rhodes (1988) findings, the holistic representation hypothesis accommodates these findings naturally, because facial features are not hypothesized to be independent units of representation and would therefore not be expected to combine independently.

Like the holistic representation hypothesis, the spatial frequency hypothesis is an explicit claim about the representations underlying face recognition. The two hypotheses are also similar in that the spatial frequency spectrum of a pattern is not decomposed in terms of the spatially delimited parts of the pattern (e.g., eyes and nose) but is matched with other candidate patterns holistically. However, the hypotheses differ in that the spatial frequency hypothesis, as put forth by Harmon (1973) and Ginsburg (1978), emphasizes the contributions of low spatial frequencies to face perception, whereas the holistic representation hypothesis makes no distinction between different frequency bands in face recognition. The latter hypothesis is therefore able to accommodate (although it does not predict) the finding that disconfirmed the spatial frequency hypothesis, that different ranges of spatial frequency are critical to face perception in different task contexts.

The idea that face representation involves parts but that the derivation of part-based representations is sensitive to top-down support from whole representations, as in the word superiority effect, is similar to the holistic representation hypothesis in its emphasis on the whole-level representation. One point of difference from our hypothesis is the assumption that part-level representations are necessarily computed during face perception and recognition. A more critical difference is that word superiority and face superiority effects are limited to threshold viewing conditions (Homa, Haver, & Schwartz, 1976; Mermelstein, Banks, & Prinzmetal, 1979) and do not manifest themselves in the full range of tasks in which holistic face representation has been found.

In sum, although there is considerable similarity and overlap between the predictions of earlier hypotheses and the holistic face representation hypothesis, the hypotheses themselves are distinct. The holistic face perception hypothesis is consistent with the findings of earlier research and, in addition, has generated a number of new predictions about memory for faces and perception of faces that have been confirmed in experiments reviewed and reported here.

Can the results of our experiments be accounted for by the earlier hypotheses? Given that most of the hypotheses include explicit, psychologically real representations of parts (parallel processing, first order and configurational features, nonindependently processed features, and even the second order relational hypothesis of Diamond & Carey, 1986, for which no independent empirical support exists), they cannot easily account for the findings reported here: specifically, participants' disproportionately poor ability to compare parts within whole faces in immediate visual memory and their disproportionately good ability to perceive a face that has been masked by face parts. The spatial frequency hypothesis, which does not include explicit

representations of parts, can account for these findings. However, as noted earlier, it has suffered direct disconfirmation in research using filtered face images.

### *Face Recognition by Monkeys and Machines: Converging Evidence*

A relatively new source of evidence on face representation is single cell recordings in monkey temporal cortex (see Desimone, 1991, for a review). Some neurons in this area respond selectively to faces, relative to other patterns, and some even respond differentially to different faces. Moreover, this selectivity in responding is maintained over changes in size, position, and contrast, consistent with a primary role for these cells in face recognition. A discrepant finding concerns the effect of lesions to the superior temporal sulcus, which did not abolish the ability of monkeys to recognize faces despite the high concentration of face-selective cells in this area (Heywood & Cowey, 1993). However, the face-selective cells in this particular region of temporal cortex are tuned more sharply to emotional expression than facial identity, whereas the cells in more inferior areas of temporal cortex are tuned more sharply for identity than expression (Hasselmo, Rolls, & Baylis, 1989). It is therefore reasonable to suppose that the inferior temporal cortex is indeed the locus of face recognition in monkeys. Given the high degree of similarity between the human and monkey visual systems, the so-called "face cells" of monkeys may provide additional clues to the nature of human face recognition.

The holistic face representation hypothesis predicts that face cells should function essentially as templates relative to a normalized stimulus pattern. That is, scrambling the features of a face should not just reduce a cell's response but should abolish it, even though there remains partial similarity between the intact and the scrambled face at the level of features. In contrast, deleting a feature should not have a dramatic effect on the cell's response, because only one region of the pattern has been changed. Both of these predictions are borne out by recordings from face cells (Desimone, Albright, Gross, & Bruce, 1984).

Evidence concerning a role for part representations in the face perception of monkeys has been discussed by Perrett, Mistlin, and Chitty (1987), who pointed out that the temporal cortex contains cells responsive to eyes and mouths as well as faces. They suggested that these facial feature cells may provide the input to face cells, so that face representations are built up from representations of face parts. Although it is indeed tempting to conjecture that the part cells provide input to the face cells, two considerations weigh in favor of caution before accepting this interpretation of the part cells. First, Desimone (1991) has pointed out that the selectivity of the part cells for face parts per se has been less well established than the selectivity of face cells for faces. That is, the possibility remains that these cells may represent more elementary visual attributes such as dark spots surrounded by bits of white rather than eyes. Second, only eye and mouth cells have been reported, and these parts of the face convey expressions that are important social cues for primates. In the absence of nose cells, chin cells, and so forth, it seems more likely that the eye and mouth cells form part of a system for nonverbal communication rather than facial identity recognition. The hypothesis that the part cells are the input to the face cells could be tested by analysis of response latencies. Although the finding that the earliest part responses occurred earlier than the earliest face responses would be ambiguous, the reverse finding would decisively rule out the hierarchical hypothesis.

Several computer systems have been developed for face recognition, and the types of representations they use may provide insights into the computational pressures toward different ways of representing faces. Although early systems used representations in which facial features were explicitly represented for recognition (e.g., Goldstein, Harmon, & Lesk, 1972; Kanade, 1977), Yuille (1991) pointed out that such representations are extremely difficult to extract from a gray scale image of a face. His "deformable template" approach uses facial features only as anchor points at which the image and stored template are brought into register; the overall fit is what determines recognition.

Turk and Pentland (1991) sought an efficient representation for faces on which to base an automatic face recognition system. Using principal-components analysis of gray scale images of faces, they found that a population code of whole faces, rather than facial features, best captured the differences among faces in a concise format. In their system of "eigenfaces," a given face is represented by weights denoting the overall similarity of the face to an ensemble of other whole faces. A related way of identifying efficient codes for representing patterns is by forcing an artificial neural network to represent patterns using limited numbers of neuronlike units and analyzing what the individual units come to represent. Such networks have been trained to classify gray scale images of faces by identity and expression (Fleming & Cottrell, 1990) as well as by sex (Golomb, Lawrence, & Sejnowski, 1991). In all of these cases, the units in the networks' hidden layer, which represent the face in a compressed form, generally correspond to whole faces rather than facial features.

In sum, recent work in computational vision has favored holistic representations of faces. It would be of great interest to compare the performance of a given system with faces and with some large and relatively homogeneous-appearing set of nonface objects. Would "eigenchair" representations be more efficient than representations based on a set of "eigenseats," "eigenlegs," and so on? Such a comparison could potentially illuminate the computational basis for holistic face representations.

### *Holistic Representation: Unique to Faces?*

We began this article with the assertion that face recognition is "special" and went on to pose the question of how it is special, in terms of the shape representations used in recognition. There are two possible interpretations of the word *special* in this context. Faces could be special in degree or special in kind. In closing, we attempt to address the question of whether face recognition is the extreme end of a continuum of part-based to more holistic representation or whether holistic representation is confined, categorically, to faces. The evidence currently at hand does not allow a decisive answer to this question. Nevertheless, some clues are available.

The comparison among faces, houses, and words in Experiment 4 suggests that houses are intermediate between faces

and words in their susceptibility to part masks. This finding is consistent with the idea of a continuum of representation, with faces the most holistic, words the most part based, and objects such as houses intermediate. We predicted this pattern of results on the basis of previously observed patterns of association and dissociation among visual recognition impairments after brain damage. Whereas face, object, and word recognition are all pairwise doubly dissociable after brain damage (i.e., for any two of these abilities, patients exist for whom one is impaired and the other preserved), it is unclear whether all three-way combinations of ability and deficit can occur. In a 1991 review of 99 published cases, Farah found no unambiguous cases of intact object recognition with impaired face and word recognition or of impaired object recognition with intact face and word recognition (see Rumiati, Humphreys, Riddoch, & Bateman, 1994, and Farah, 1997, for an updated discussion). What type of underlying organization would impose these constraints on patterns of co-occurrence among visual recognition impairments?

The simplest solution involves two underlying representational abilities, one that is essential for face recognition, useful for object recognition, and not used for word recognition and another that is essential for word recognition, useful for object recognition, and not used for face recognition. The work of Johnston and McClelland (1980) with normal participants, discussed earlier, suggests that word recognition requires the construction of a part-based (specifically, letter-based) representation. Work with an alexic patient suggests that a problem underlying selective impairments in visual word recognition is an impairment in representing multiple parts, be they letters in a word or complex nonorthographic stimuli (Farah & Wallace, 1991). Taking these findings together with the evidence of holistic face representation, a plausible inference is that the two representational abilities uncovered in the analysis of co-occurrence correspond to holistic and part-based representation. More specifically, the ability to represent complex wholes with little or no part decomposition may be the ability required for face recognition, useful for object recognition, and not needed for word recognition, and the ability to represent a number of distinct parts may be the ability required for word recognition, useful for object recognition, and not needed for face recognition. This interpretation of the data suggests that faces are special in degree, not in kind. Specifically, it suggests that they constitute an extreme case of stimuli that rely on holistic shape representation but are not necessarily discontinuous from other types of objects in their reliance on holistic representation.

## References


Baylis, G. C., Rolls, E. T., & Leonard, C. M. (1985). Selectivity between faces in the responses of a population of neurons in the cortex in the superior temporal sulcus of the monkey. *Brain Research, 342,* 91–102.

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 reintegrative power of picture fragments. *Journal of Experimental Psychology, 2,* 456–466.

Bradshaw, J. L., & Wallace, G. (1971). Models for the processing and identification of faces. *Perception & Psychophysics, 9,* 443–448.

Bruce, V. (1988). *Recognizing faces.* Hove, England: Erlbaum.

Desimone, R. (1991). Face-selective cells in the temporal cortex of monkeys. *Journal of Cognitive Neuroscience, 3,* 1–8.

Desimone, R., Albright, T. D., Gross, C. D., & Bruce, C. (1984). Stimulus-selective responses of inferior temporal neurons in the macaque. *Journal of Neuroscience, 4,* 2051–2062.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General, 115,* 107–117.

Farah, M. J. (1990). *Visual agnosia: Disorders of object recognition and what they tell us about normal vision.* Cambridge, MA: MIT Press/Bradford Books.

Farah, M. J. (1991). Patterns of co-occurrence among the associative agnosias: Implications for visual object representation. *Cognitive Neuropsychology, 8,* 1–19.

Farah, M. J. (1996). Is face recognition 'special'? Evidence from neuropsychology. *Behavioral Brain Research, 76,* 181–189.

Farah, M. J. (1997). Distinguishing perceptual and semantic impairments affecting visual object recognition. *Visual Cognition, 4,* 199–206.

Farah, M. J., Klein, K. L., & Levinson, K. L. (1995). Face perception and within-category discrimination in prosopagnosia. *Neuropsychologia, 33,* 661–674.

Farah, M. J., Tanaka, J. N., & Drain, M. (1995). What causes the face inversion effect. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 628–634.

Farah, M. J., & Wallace, M. A. (1991). Pure alexia as a visual impairment: A reconsideration. *Cognitive Neuropsychology, 8,* 313–334.

Fleming, M., & Cottrell, G. W. (1990). Categorization of faces using unsupervised feature extraction. *Proceedings of IJCNN-90, 2,* 65–70.

Garner, W. R. (1974). *The processing of information and structure.* Potomac, MD: Erlbaum.

Ginsburg, A. (1978). *Visual information processing based on spatial filters constrained by biological data.* Unpublished doctoral dissertation, Cambridge University, Cambridge, England.

Goldstein, A. J., Harmon, J. D., & Lesk, A. B. (1972). Identification of human faces. *Proceedings of the IEEE, 59,* 748–760.

Golomb, B. A., Lawrence, D. T., & Sejnowski, T. J. (1991). *SexNet: A neural network identifies sex from human faces.* San Mateo, CA: Morgan Kaufmann.

Harmon, L. D. (1973). The recognition of faces. *Scientific American, 227,* 71–82.

Hasselmo, M. E., Rolls, E. T., & Baylis, G. C. (1989). The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. *Behavioral Brain Research, 32,* 203–218.

Heywood, C. A., & Cowey, A. (1993). The role of "face-cell" area in the discrimination and recognition of faces by monkeys. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences, 335,* 31–38.

Hoffman, D. D., & Richards, W. (1985). The parts of recognition. *Cognition, 18,* 65–96.

Homa, D., Haver, B., & Schwartz, T. (1976). Perceptibility of schematic face stimuli: Evidence for a perceptual Gestalt. *Memory & Cognition, 4,* 176–185.

Johnson, M. H., Dziurawiec, S., Ellis, H. D., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition, 40,* 1–19.

Johnston, J. C., & McClelland, J. C. (1980). Experimental tests of a hierarchical model of word identification. *Journal of Verbal Learning and Verbal Behavior, 19,* 503–524.

Kanade, T. (1977). *Computer recognition of human faces.* Basel, Switzerland: Birkhauser Verlag.

Marr, D. (1982). *Vision.* San Francisco: Freeman.

Matthews, M. L. (1978). Discrimination of identikit construction of faces: Evidence for a dual processing strategy. *Perception & Psychophysics, 23,* 153–161.

Mermelstein, R., Banks, W., & Prinzmetal, W. (1979). Figural goodness effects in perception and memory. *Perception & Psychophysics, 26,* 472–480.

Moscovitch, M., Winocur, G., & Behrmann, M. (1997). What is special about face recognition? Nineteen experiments on a person with visual object agnosia and dyslexia but normal face recognition. *Journal of Cognitive Neuroscience, 9,* 555–604.

Palmer, S. E. (1975). The effects of contextual scenes on the identification of objects. *Memory & Cognition, 3,* 519–526.

Palmer, S. E. (1977). Hierarchical structure in perceptual representation. *Cognitive Psychology, 9,* 441–474.

Perrett, D. I., Mistlin, A. J., & Chitty, A. J. (1987). Visual neurones responsive to faces. *Trends in Neuroscience, 10,* 358–364.

Reed, S. K. (1974). Structural descriptions and the limitations of visual images. *Memory & Cognition, 2,* 329–336.

Reicher, B. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimenal Psychology, 81,* 275–280.

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

Rumiati, R. I., Humphreys, G. W., Riddoch, M. J., & Bateman, A. (1994). Visual object recognition without prosopagnosia or alexia: Evidence for hierarchical theories of object recognition. *Visual Cognition, 1,* 181–225.

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

Smith, E. E., & Nielsen, G. D. (1970). Representations and retrieval processes in short-term memory: Recognition and recall of faces. *Journal of Experimental Psychology, 85,* 397–405.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica, 30,* 276–315.

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 Experimental Psychology, 46,* 225–245.

Tanaka, J. W., Kay, J. B., Grinnell, E., & Stansfield, B. (in press). Face recognition in young children: When the whole is greater than the sum of its parts. *Visual Cognition.*

Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory & Cognition, 25,* 583–592.

Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience, 3,* 71–86.

Valentine, T. (1988). Upside-down faces: A review of the effects of inversion upon face recognition. *British Journal of Psychology, 79,* 471–491.

Wheeler, D. D. (1970). Processes in word recognition. *Cognitive Psychology, 1,* 59–85.

Yuille, A. L. (1991). Deformable templates for face recognition. *Journal of Cognitive Neuroscience, 3,* 59–70.

## Appendix

### Examples of Whole and Part Masks for Word Stimuli

| Whole masks | Part masks |
|---|---|
| brag | arbg |
| rank | nkar |
| baby | abyb |
| king | nikg |
| bark | abkr |
| gang | ngag |
| ring | gnri |
| wing | nigw |
| wink | wkni |

Received March 24, 1995
Revision received November 26, 1997
Accepted December 26, 1997 ■