Journal of Experimental Psychology:
Human Perception and Performance
1995, Vol. 21, No. 4, 901–913

0096-1523/95/$3.00

Facial Organization Blocks Access to Low-Level Features: An Object Inferiority Effect

Satoru Suzuki and Patrick Cavanagh
Harvard University

The current study investigated the influence of a low-level local feature (curvature) and a high-level emergent feature (facial expression) on rapid search. These features distinguished the target from the distractors and were presented either alone or together. Stimuli were triplets of up and down arcs organized to form meaningless patterns or schematic faces. In the feature search, the target had the only down arc in the display. In the conjunction search, the target was a unique combination of up and down arcs. When triplets depicted faces, the target was also the only smiling face among frowning faces. The face-level feature facilitated the conjunction search but, surprisingly, slowed the feature search. These results demonstrated that an object inferiority effect could occur even when the emergent feature was useful in the search. Rapid search processes appear to operate only on high-level representations even when low-level features would be more efficient.

Recent studies using visual search paradigms have demonstrated rapid processing for intermediate- to high-level scene attributes such as layout of surfaces in three dimensions (He & Nakayama, 1992), orientation of three-dimensional (3D) objects (Enns & Rensink, 1990, 1991), shadows (Rensink & Cavanagh, 1993), and stimulus familiarity (Wang, Cavanagh, & Green, 1994). These studies challenge the idea that the search process has rapid access to only low-level image features[1] such as luminance, color, line orientation, size (e.g., Treisman & Gelade, 1980), and curvature (Wolfe, Yee, & Friedman-Hill, 1992). They suggest that the visual search process can access different levels of image representation during rapid pattern discrimination.

In the present series of studies, we systematically investigated the relative influence of a low-level local feature and a high-level emergent feature on visual search by using curvature of arcs (upward curving or downward curving) and facial expression (smiling or frowning) as the respective features. The reasons why we chose facial organization for our high-level feature are twofold. First, it has been shown that the perception of facial organization is very rapid—rapid enough to be a plausible participant in determining the visual search performance. Specifically, the detection threshold of a schematic face has been demonstrated to be less than 40 ms (in terms of stimulus-to-mask interval), which is significantly lower than that for a feature-matched meaningless pattern, namely a jumbled face (Gorea & Julesz, 1990; Purcell & Stewart, 1988).

Second, Mermelstein, Banks, and Prinzmetal (1979) have shown that a shape discrimination task is slowed when the target shape is embedded as the nose in a face. Their results suggest that the global organization impedes the discrimination of the parts. They argued that a gestalt such as a face is first processed holistically to extract global features such as facial expressions. Thus, target discrimination based solely on the individual parts is slowed by the presence of a global feature because it takes an extra step to tease apart the local features from the whole. In the study by Mermelstein et al., the facial context was never informative for the discrimination. Although they did not test stimuli where global features could contribute to the performance, they did propose that if emergent features did distinguish the target from the distractors, then a global context would facilitate the target discrimination. Their claim was tested in our experiments in which the separated and combined influences of the global emergent feature (facial expression) and the local feature (curvature) were examined.

A set of visual search experiments was conducted to compare the search efficiency for the low- and high-level features alone (local curvature feature or global facial feature) with that obtained when both levels of features were available and when both discriminated the target from the distractors. To separate the effects of the two levels of features, we constructed two types of visual search based on the local curvature—feature search and conjunction search—and combined them with the presence or the absence of the global facial feature. We displayed triplets of

Satoru Suzuki and Patrick Cavanagh, Department of Psychology, Harvard University.

This research was supported by U.S. Air Force Office of Scientific Research Grant 91-0169. We are grateful to Jody Culham, Qinqin Wang, James Intriligator, and Ronald Rensink for many insightful comments on earlier versions of this article.

Correspondence concerning this article should be addressed to Satoru Suzuki, who is now at the Perception Laboratory, Department of Industrial and Management Systems Engineering, Arizona State University, Tempe, Arizona 85287.

[1] The low-level image features refer to those image features that are believed to be processed at an early stage of visual processing, namely in the primary visual cortex, by the neurons whose receptive fields show response tunings for simple image features such as luminance–color contrast, line orientation, spatial frequency, motion direction, and binocular disparity.

upward- and downward-curving arcs (up and down arcs) as search items. In Figure 1, each quadrant shows an example of a stimulus array containing one target and five distractors. For convenience in this figure, the target (composed of one down arc and two up arcs) is located at the right-most location in each array. In the experimental trials, however, the target could occupy any of the six locations, and the number of the distractors varied from one to five.

Figure 1c shows the curvature feature search. In this condition, the target contains the unique down arc, whereas the distractors contain only up arcs. This condition provides an estimate of the efficiency of the search based on the curvature feature alone in the absence of the facial organization. In Figure 1d, the arcs within each triplet have been rearranged to form a schematic face. The target is now a smiling face, and the distractors are frowning faces. In this condition, both the local curvature feature and the emergent facial feature are available to discriminate the target from the distractors. Thus, this condition provides the search efficiency when the two levels of features are combined.

Although it is easy to create stimuli in which curvature alone varies or in which both curvature and facial expression vary, it is much harder to create stimuli in which only the facial expression varies, as this requires a stimulus with a distinguishing high-level feature but no distinguishing low-level feature. Although this is logically impossible, we can approximate this requirement by rendering the search based on the low-level feature very difficult, and we have chosen a conjunction search for this reason. Figures 1a and 1b show the curvature-by-location conjunction search with and without facial organization. The conjunction search itself (Figure 1a) is quite difficult (e.g., Treisman & Gelade, 1980) because it is the configuration of arcs that identifies the target rather than the presence or the absence of a unique arc. The target has one down arc and two up arcs, whereas the distractors have one up arc and two down arcs. In Figure 1b, both the global facial feature and the curvature-by-location conjunction are available to discriminate the target. However, because of the difficulty of the search based on the conjunction cue, we expected that the facial feature would trigger the responses on most trials. Note that we are pre-supposing that the presence of facial organization significantly speeds the difficult conjunction search, a constraint that we were able to test. If our data meet this requirement, we can regard the performance in the conjunction search with facial organization (Figure 1b) as a measure of the efficiency of search based on the facial feature alone.

At least three patterns of results are possible on the basis

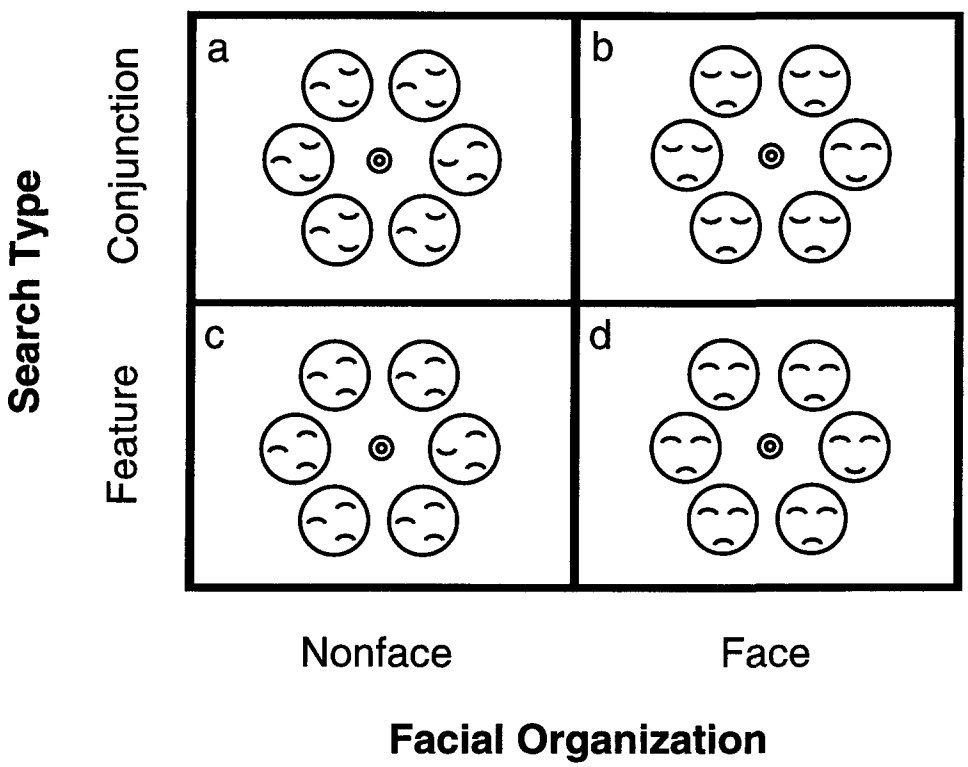

Figure 1. The 2 × 2 design used in Experiment 1. The search type (conjunction vs. feature) is crossed with the presence–absence of the facial organization (nonface vs. face). An example of the stimulus array is shown for each of the four conditions. The right-most pattern in each stimulus array is the target. Panel a shows conjunction search (curvature by location) without facial organization. Panel b shows conjunction search with facial organization (target is smiling whereas distractors are frowning). Panel c shows feature (curvature) search without facial organization. Panel d shows feature search with facial organization.

of three different models of the relative efficiency of the two levels of image representation and how accessible they are to the search process during rapid visual search. The first model assumes that processing proceeds sequentially from low- to high-level features so that the high-level features are not available until processing of the low-level features is completed. If this were the case, the search based on the local curvature feature (Figure 1c) would be as fast as (or faster than) the search based on the global facial feature (Figure 1b). Furthermore, the addition of the facial organization (Figure 1d) would make no difference in the performance of the curvature feature search because the search would still be based on the more rapidly accessible curvature feature.

The second model assumes that the search process would have equal access to both levels of image representation so that it could take advantage of whichever level of processing that turns out to be more efficient. If this were the case, the search based on the two levels of features combined (Figure 1d) would be at least as fast as, if not faster than, the search based on the more efficient of the two features by virtue of the target differing from the distractors by two distinctive features instead of just one.

The last model assumes that the search process has access to only the high-level global representation during rapid search. If this were the case, so long as the stimulus arcs are organized into faces, the constituent curvature features would be unavailable to the search process, just as a surface-level representation has been shown to render the low-level image features ineffective during rapid search (He & Nakayama, 1992). Thus, even if both levels of features were present in the image to discriminate the target from the distractors (Figure 1d), the search rate would be determined by the efficiency of the global face-level processing. If the global facial feature were more efficient than the local curvature feature, the facial organization would facilitate the curvature search. Alternatively, if the local curvature feature were the more efficient one, then the facial organization would end up slowing the curvature search.

Note that Mermelstein et al. (1979) proposed that as long as the global feature distinguishes the target from the distractors, it should facilitate the target discrimination based on local features. They would thus predict that the global facial feature could only facilitate the curvature search.

We tested these predictions in Experiment 1 by conducting four visual search tasks corresponding to Figures 1a–1d. In addition, the conjunction search conditions (Figures 1a and 1b) were replicated using upside-down faces instead of jumbled faces in Experiment 2 (Figure 5, as shown later) to show that the effects of facial organization exist above and beyond the difference in local geometry. In Experiment 3, the feature search tasks (Figures 1c and 1d) were replicated with the bounding circles removed to see if the effect of facial organization depended on the presence of facial contour (Figure 8, as shown later). Furthermore, in this experiment, the target was an up arc or a down arc for any given trial, and the participant was specifically instructed to always look for an oddly curved arc in the display. This modification was made to discourage the participant from consciously adopting a face strategy, such as looking for a smiling face, to demonstrate that the effects of facial organization are not contingent on such a conscious strategy.

Experiment 1

We conducted four visual search tasks corresponding to Figures 1a–1d to examine how the two levels of attribute (curvature and facial expression) would interact in determining the visual search performance. We varied the size of the stimulus array from one to six. We used the slopes of search functions (search rates in milliseconds per item) as the measure of task difficulty.

Method

Participants. Ten undergraduate and graduate students from Harvard University participated in the experiment. All but 1 participant were naive as to the purpose of the experiment. All participants had normal or corrected-to-normal vision and were tested individually.

Apparatus and stimuli. All the experimental stimuli that were used are shown in Figures 1a–1d. The hexagonal array subtended a visual angle of 2°10′ in diameter and each individual stimulus 50′ in diameter. The horizontal and vertical distance between the arcs within a triplet was 25′. Pixel maps of the up and down arcs are shown in Figure 2. Each arc subtended 17′ in width and 4′ in height. A Macintosh IIcx computer was used to display the stimuli and to control the experiment.

Each visual search task consisted of an equal number of target-present and target-absent trials. The stimuli were presented at the six vertices of the hexagonal array; at each location, however, the stimulus was positioned randomly within the 8′ × 8′ square region about the vertex. In the target-present trials, the target appeared equally likely at any of the six locations. The distractor positions were chosen randomly from the five remaining locations, and each number of distractors (zero to five) was displayed with equal probability. The target-absent trials were basically the same as the

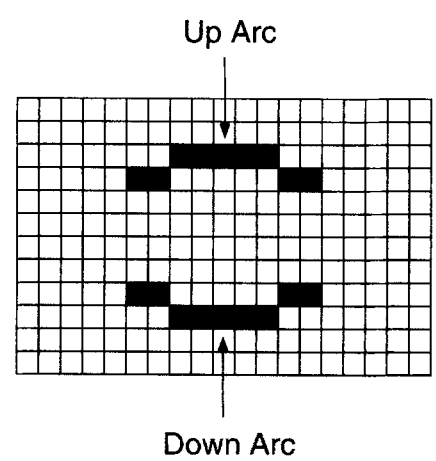

Figure 2. Depiction of the stimulus arcs (up arc and down arc) used in Experiment 1, shown in a grid of pixel-by-pixel squares. Each pixel subtended about 2′ of visual angle from a 70-cm viewing distance.

target-present trials, except that one to six distractors were displayed with equal probability.

There was a total of 144 experimental trials for each task to ensure that every possible combination of stimulus parameters appeared twice across the 144 trials: 2 repetitions of Target Present–Absent (2) × Number of Stimuli (6) × Target Location (6). Because the last factor, target location, did not apply to the target-absent trials, 12 repetitions were used to match the number of the target-present trials.

Procedure. The participant was seated in front of the computer screen at a distance of 70 cm. Prior to each trial, the participant fixated on the bull's-eye at the center of the screen. A short beep signaled the beginning of a trial. The stimulus array appeared after a 1-s delay and stayed on until the participant made a response. The task was to press one key (with the right index finger) when the target was present and another key (with the left index finger) when the target was absent. The participants were shown the target and distractor patterns prior to each task; no verbal descriptions of the stimuli were given to inform the participants of the resemblance of the stimuli to faces in the face conditions. The participants were encouraged to respond as quickly as possible while keeping incorrect responses to a minimum. A short, low-pitch beep was given as feedback whenever an incorrect response was made.

Practice trials were given prior to each task until the participant became familiar with the stimuli and the procedure of the experiment. The practice sessions lasted until at least 90% correct performance was reached. Twenty practice trials were also added at the beginning of each task.

The experimental order of the four tasks for half of the participants was as follows: (a) feature search without facial organization, (b) conjunction search without facial organization, followed by a 15-min break, (c) feature search with facial organization, and (d) conjunction search with facial organization. For the other half of the participants, the order was as follows: (a) conjunction search without facial organization, (b) feature search without facial organization, followed by a 15-min break, (c) conjunction search with facial organization, and (d) feature search with facial organization. The face conditions were tested after the nonface conditions to minimize the possible face interpretation of the meaningless patterns. The possible confound from the order effect was examined in Experiment 2 (control experiment) and in Experiment 3, in which task orders were fully counterbalanced across participants.

Results

In Figure 3, mean response times for correct responses are plotted as a function of the number of stimuli for the target-present trials. Because linear regression accounts for 94%–96% of the total variance due to the stimulus set size for each search function, we performed the subsequent analyses on the slopes (search rates) of the search functions.

Figure 4 shows the feature and conjunction search rates for target-present trials with and without facial organization. The target-absent trials (not shown) had the same trends as the target-present trials except that their search rates were about twice as high for all conditions, an outcome which is indicative of serial searches. The error rates were low (2%–3%) for all conditions. The statistical analyses were performed on the target-present trials because only in those trials was the target actually discriminated against the background of distractors. Furthermore, the performance on the target-absent trials tended to be less reliable because of the variation in the individual participants' criterion for verifying the absence of the target.

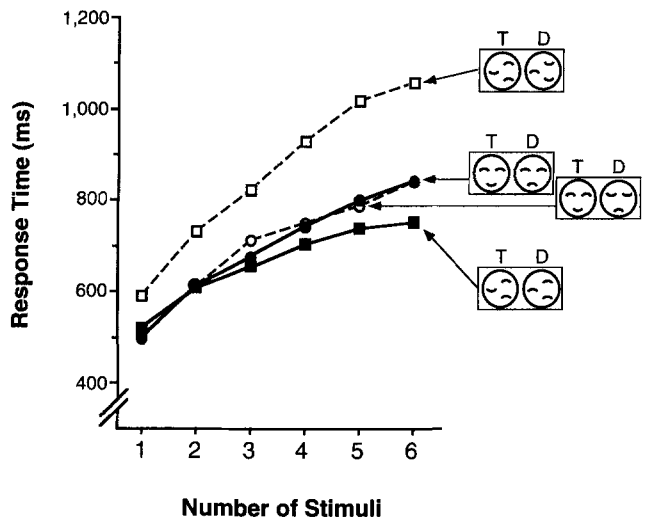

Figure 3. Mean search functions for the four experimental conditions for target-present trials (correct responses only) in Experiment 1: from top to bottom, conjunction search (curvature by location) without facial organization (open squares), feature (curvature) search with facial organization (filled circles), conjunction search with facial organization (open circles), and feature search without facial organization (filled squares). Shown beside each search function are the corresponding target (T) and distractor (D) patterns. Mean search functions for target-absent trials (not shown) had the same pattern except that the slopes were about twice as steep for each condition.

We performed a two-way analysis of variance on the search rates. Overall, the feature search was faster than the conjunction search, $F(1, 9) = 15.68$, $p < .005$. There was no significant main effect of the facial organization, $F(1, 9) = 0.60$, *ns*. As seen in Figure 4, the interaction between the

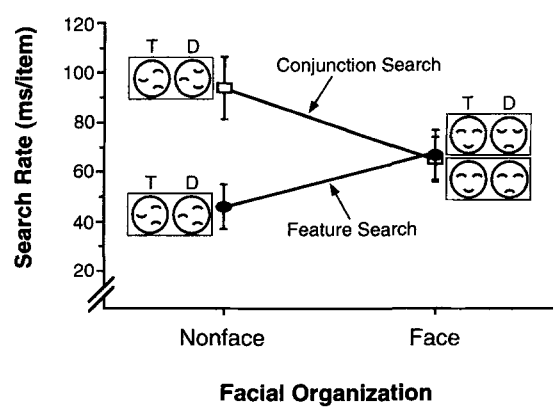

Figure 4. Mean search rates for target-present trials. Shown next to each point on the graph are the corresponding target (T) and distractor (D) patterns. Mean search rates for target-absent trials (not shown) had the same pattern except that they were about twice as steep for each condition. The vertical bars represent ±1 *SE*.

search type and the presence–absence of facial organization was significant, $F(1, 9) = 15.13$, $p < .005$.

We now consider the effects of the search type and the facial organization separately. As for the effect of the search type, we found that without the facial organization (left half of Figure 4), the feature (curvature) search was much faster than the conjunction search (95 ms/item vs. 46 ms/item, respectively), $F(1, 9) = 24.21$, $p < .001$, as was expected on the basis of earlier findings by Treisman and others (e.g., Treisman & Gelade, 1980). An interesting phenomenon was observed in the presence of the facial organization, as can be seen in the right half of Figure 4. The feature and conjunction searches no longer differed in search rate (67 ms/item vs. 65 ms/item, respectively), $F(1, 9) = 0.04$, *ns*. Note that the target–distractor similarity has also been considered to be an important factor in determining search efficiency; a search is faster when the target and the distractor are more dissimilar (e.g., Duncan & Humphreys, 1989). One way to estimate the target–distractor similarity is to measure their physical overlap. For our stimuli, this metric gave the interesting result that the target and the distractor were more similar in the feature (curvature) search (only one arc differed; see Figure 1c) than in the conjunction search (all arcs differed; see Figure 1a), incorrectly predicting that the conjunction search would be faster. Thus, in our task, the physical similarity between target and distractor did not appear to be an important factor in determining search efficiency.

Looking next at the effects of the facial organization, we found that the presence of facial organization significantly facilitated the conjunction search (see Figure 4; 95 ms/item vs. 65 ms/item), $F(1, 9) = 6.25$, $p < .05$. This replicates the well-known phenomenon of object superiority (e.g., Enns & Rensink, 1990, 1991; Weisstein & Harris, 1974). An unexpected phenomenon was found in the feature (curvature) search conditions (see Figure 4), where the presence of the facial organization significantly slowed the search (46 ms/item vs. 67 ms/item), $F(1, 9) = 8.35$, $p < .025$.

It could, however, be argued that the effect of the facial organization was due to some particular local alignment of the arcs independent of the "faceness" of the face patterns. Specifically, the two identical arcs within each stimulus triplet were arranged horizontally in the face patterns (Figures 1b and 1d) but vertically in the meaningless patterns (Figures 1a and 1c). To examine this point, we designed a control experiment.

Experiment 2

It is well-known that upside-down faces evoke much weaker impressions of faceness (Carey, 1982; Thompson, 1980) and are more prone to feature-based perception (Sergent, 1984; Young, Hallwell, & Hay, 1987). The detection threshold has also been found to be significantly higher for upside-down faces (Purcell & Stewart, 1988).

In this experiment, we substituted an upside-down version of the face patterns for the meaningless patterns (see Figure 5). The local alignment of the arcs was now identical for the face and upside-down–face triplets. Only the conjunction search conditions were replicated, because the demonstration of a significant advantage of the faces over the upside-down faces in this condition would be sufficient to rule out the argument based on the local alignment of the arcs. Furthermore, in this experiment, the participants were tested alternately for the upright-face and upside-down–face conditions for a total of four blocks with complete counterbalancing (twice as many trials per participant as in Experiment 1) to examine the possibility that the effect of the facial organization found in Experiment 1 might have been confounded by an order effect.

Method

Participants. Four undergraduate and graduate students from Harvard University participated in the experiment. These individuals did not participate in Experiment 1 and were naive as to the purpose of the experiment. All participants had normal or corrected-to-normal vision and were tested individually.

Apparatus and stimuli. They were basically identical to those in Experiment 1 except that the upside-down version of the face patterns was substituted for the nonface patterns (Figure 5).

Procedure. The procedure was basically identical to Experiment 1 with one exception. The number of trials per task was doubled; the participants were given two blocks per task. Each block consisted of 144 trials, so that every possible combination of stimulus parameters appeared twice within each block (see the *Method* section in Experiment 1). We alternated the upright-face task (Figure 5b) with the upside-down–face task (Figure 5a) over the course of the four blocks. Thus, half of the participants began with an upright-face block, so the order was (a) upright, (b) upside-down, (c) upright, and (d) upside-down. The other half of the participants started with a block of upside-down faces, the order being (a) upside-down, (b) upright, (c) upside-down, and (d) upright. If the facilitating effect of the facial organization found in the previous experiment was due to the fact that the face conditions were tested after the nonface conditions, this counterbalancing would eliminate the effect of facial organization.

Results

The search functions for the target-present trials are shown in Figure 6. The superior performance in the upright-face condition is apparent. Because linear regression accounts for more than 99% of the total variance due to the stimulus set size for both conditions (upright and upside-down faces), we subsequently analyzed the search rates. Mean search rates for the target-present trials are plotted as a function of the face orientation in Figure 7. The target-absent trials (not shown) had the same pattern as the target-present trials except that their search rates were higher. The overall error rates were low (3%–5%). We performed the statistical analyses on the search rates for the target-present trials as we did in Experiment 1.

The search in the upright condition was faster than that in the upside-down condition for all participants (85 ms/item vs. 141 ms/item, respectively), $T(3) = 2.89$, $p < .05$. (A one-tailed t test was used, because this was a replication of the face superiority effect found in Experiment 1.) Because

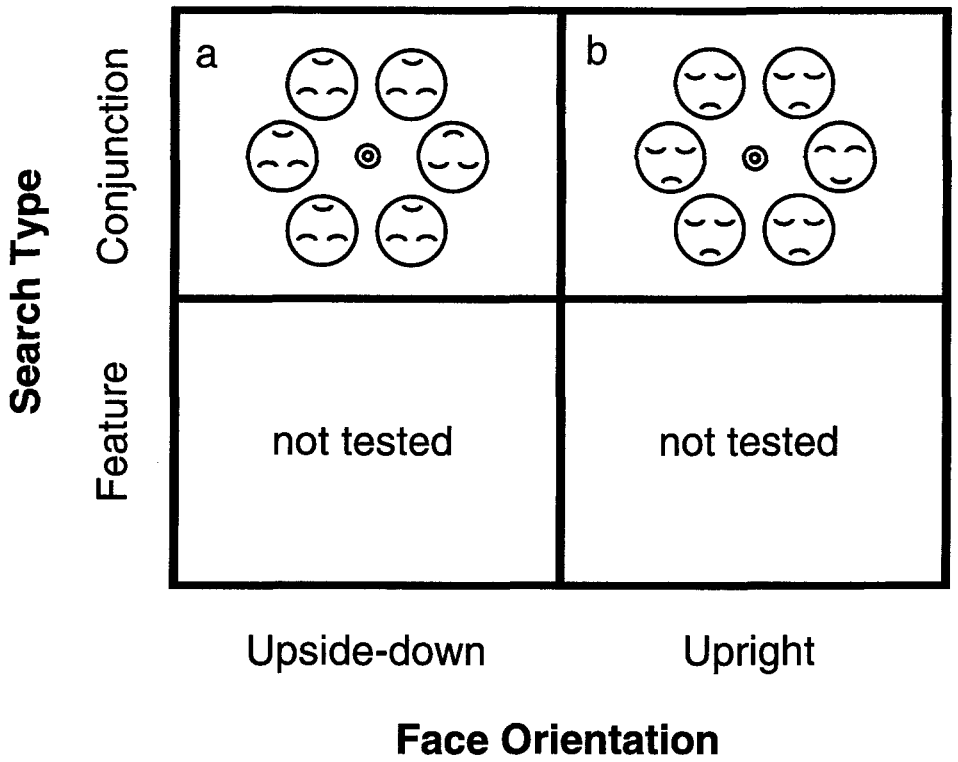

Figure 5. The stimulus arrays used in Experiment 2. The right-most pattern in each stimulus array is the target. Note that the facial expressions seem weaker in the upside-down faces.

nothing but the faceness of the stimuli was manipulated by inverting the face patterns and because the trial orders were completely counterbalanced, we concluded that neither the differences in the local grouping nor testing order was responsible for the effects of facial organization demonstrated in Experiment 1.

In Experiment 1, we designated curvature of arcs (up or down) as a low-level local feature and facial expression (smile or frown) as a high-level emergent feature. We then investigated how the two levels of features influenced visual search performance. We found the expected advantage of feature (curvature) search over conjunction search (curvature by location) for the nonface stimuli (left half of Figure 4). We also found the expected advantage of facial organization over meaningless organization in the difficult conjunction search (see Figure 4). Our critical finding is that the facial organization slowed the feature search (see Figure 4); our interpretation is that the presence of the global face-level feature precluded access to the local curvature feature even though the latter could be processed more efficiently to discriminate the target.

Our results for the feature (curvature) search condition with nonface stimuli, however, require some discussion because the search rate was 46 ms/item, which is much too high to be considered a rapid feature search (e.g., Wolfe et al., 1992, found a rate of 1.4 ms/item and 11.3 ms/item for positive and negative trials, respectively, when a leftward-curving arc was searched for among rightward-curving arcs). Our search rate was high partly because we counted each triplet of arcs within the bounding circle as an item. In the curvature search task, the number of items to be searched was really the total number of individual arcs, and there were three times as many arcs as there were stimulus triplets. If we consider each arc as an item, then the search rate was about 15 ms/arc. This was still somewhat higher than the rate typically reported for a fast feature search (less than 10 ms/item) in the literature. We suspected that the bounding circles surrounding each triplet could account for the remaining slowness of our curvature search. It is plausible that a fast feature search relies on the grouping of distractors by their similarity in low-level attributes. The bounding circle around each triplet may have prevented this process by forcing the local grouping within each triplet.

The next experiment was designed to explore this possibility as well as to replicate and extend the most interesting results found in Experiment 1, that the presence of facial organization slowed the low-level curvature search. We intended to see if facial organization would slow the curvature search even when the search was rendered a typical parallel search.

Experiment 3

In this experiment, we conducted the curvature search task under the face and nonface conditions as we did in

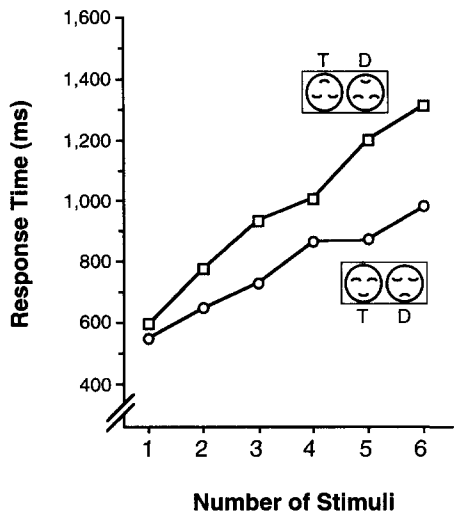

Figure 6. Mean search functions for the two experimental conditions for target-present trials (correct responses only) in Experiment 2: from top to bottom, conjunction search (curvature by location) with upside-down facial organization (open squares) and conjunction search with upright facial organization (open circles). Shown beside each search function are the corresponding target (T) and distractor (D) patterns. Mean search functions for target-absent trials (not shown) had the same pattern except that the slopes were steeper for each condition.

Experiment 1 but without the circular contours surrounding the triplets of arcs (see Figure 8). The individual arcs were made slightly larger and were given higher curvature to increase their discriminability. Another modification was that the target could be down or an up arc with equal probability from trial to trial, so the participant's task was to search for an oddly curved arc instead of always looking for a down arc as in Experiment 1. In the face condition, the target arc was embedded in a smiling face if it was a down arc (Figure 8d) and in a frowning face if it was an up arc (Figure 8b). Therefore, the participant could not always look for just a smiling face as was possible in Experiment 1. In addition, the participants were given verbal instructions to look for an oddly curved arc in both the nonface and face conditions. The latter modifications were made to reduce the likelihood that the participant might consciously use a face strategy.

A pilot study had shown that the aforementioned modifications indeed rendered the curvature search a typical parallel search with a flat search function when facial organization was not present. The slowing of the curvature search by facial organization, if observed under this condition, would be strong evidence in support of the idea that the search process has an obligatory access to the face-level representation even if the lower-level feature could support a fast parallel search.

Method

Participants. Ten graduate students and postdoctoral fellows from Harvard University participated in the experiment. All were naive as to the purpose of the experiment and did not participate in Experiments 1 or 2, except for 1 individual who participated in Experiment 1. All participants had normal or corrected-to-normal vision and were tested individually.

Apparatus and stimuli. The stimuli that were used are shown in Figures 8a–8d. The hexagonal array subtended a visual angle of 3°18′ in diameter. Pixel maps of the up and down arcs are shown in Figure 9. Each arc subtended 18′ in width and 7′ in height; each triplet of arcs was confined within a 54′ × 54′ square region. The horizontal and vertical distance between the arcs within a triplet was 32′. As shown in Figures 8a–8d, the target triplet always contained a unique up (Figures 8a and 8b) or down arc (Figures 8c and 8d) equally likely in the target-present trials. In the target-absent trials, the whole display contained only the up arcs or the down arcs with equal probability. As with the previous experiments, a Macintosh IIcx computer was used to display the stimuli and to control the experiment.

The triplets were presented at the six vertices of the hexagonal array; at each location, however, the triplet was positioned randomly within the 36′ × 36′ square region about the vertex. In the target-present trials, the triplet that contained the target arc appeared equally likely at any of the six locations. The positions for the distractor triplets were chosen randomly from the five remaining locations, and each number of distractor triplets (one, three, or five) was displayed with equal probability. The target-absent trials were basically the same as the target-present trials except that two, four, or six distractor triplets were displayed with equal probability.

There was a total of 72 experimental trials per block for each of the two tasks (nonface and face) to ensure that every possible combination of stimulus parameters appeared once within each block: Target Arc Type (2) × Target Present–Absent (2) × Number of Triplets (3) × Target Triplet Location (6). Because the last factor, target triplet location, did not apply to the target-absent trials, 12 repetitions were used to match the number of the target-present trials.

Procedure. The basic procedure was identical to the previous experiments. The stimuli were displayed until the participant made a response. The participant was instructed to always search for an oddly curved arc (up among down or down among up) and to respond as quickly as possible while keeping errors to a minimum.

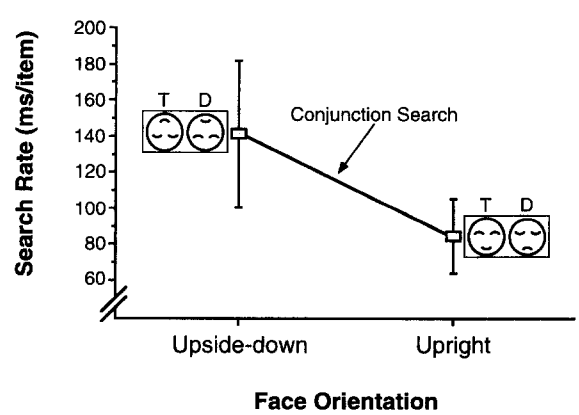

Figure 7. Mean search rates for target-present trials. Shown next to each point on the graph are the corresponding target (T) and distractor (D) patterns. Mean search rates for target-absent trials (not shown) had the same pattern except that they were steeper for each condition. The vertical bars represent ±1 *SE*.

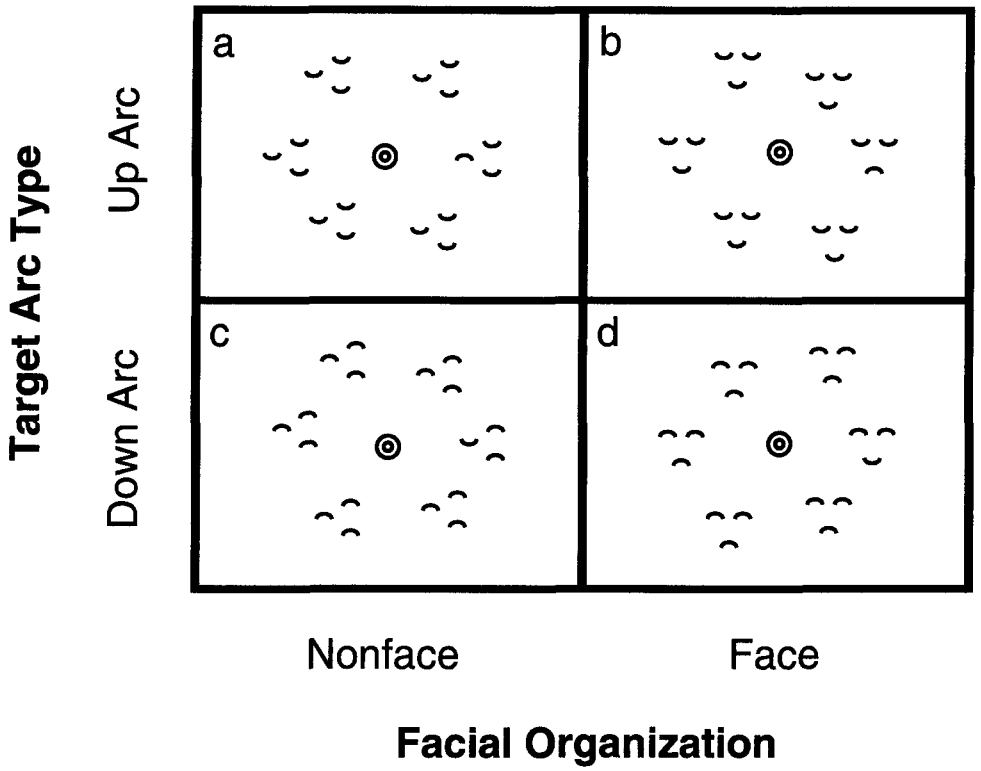

Figure 8. The stimulus arrays used in Experiment 3. The right-most triplet in each stimulus array contains the target arc. Note that the bounding circles have been removed. Panels a and c show feature (curvature) search without facial organization when the target is an up arc (a) and when the target is a down arc (c). Panels b and d show feature search with facial organization when the target is an up arc (b) and when the target is a down arc (d). The spacings between arcs are drawn approximately to scale for the actual display of a typical trial with 18 arcs.

Thirty practice trials were given prior to each block, and an additional 20 practice trials were also given at the beginning of each block.

Each participant was tested for two blocks of each condition for a total of 288 (72 × 4) trials. The nonface and face blocks were tested alternately, and the two possible orders were counterbalanced across participants. Half of the participants started with the nonface block, the order being (a) nonface, (b) face, (c) nonface, and (d) face. The other half of the participants started with the face block, the order being (a) face, (b) nonface, (c) face, and (d) nonface.

Results

Mean response times for the target-present trials are plotted as a function of the number of arcs in the display in Figure 10. The overall error rate was 4.5% for the nonface condition and 7.5% for the face condition; thus, there was no apparent speed–accuracy trade-off. The target-absent trials (not shown) had the same trend as the target-present trials with no apparent steepening of the slopes, an outcome which is indicative of parallel searches. The search rates are plotted in Figure 11 for the target-present trials. (Linear regression accounts for more than 94% of the total variance due to the stimulus set size for both conditions.) Note that these search rates are quite low and are within the so-called parallel search range at least for the nonface condition (3 ms/arc). As in the previous experiments, we performed statistical analyses on the target-present trials.

The search rate was significantly higher for the face condition than for the nonface condition (11 ms/arc vs. 3 ms/arc, respectively), $F(1, 9) = 5.66$, $p < .05$. Note that the search rate of 3 ms/item for the nonface condition satisfies the behavioral criterion to be regarded as a fast parallel search. Furthermore, as is apparent in Figure 10, the overall response time was much slower in the face condition than in the nonface condition. This difference was significant for all display sizes: The response times were 798 ms (face) versus 685 ms (nonface), $F(1, 9) = 18.20$, $p < .005$, for Display Size 6; 891 ms (face) versus 701 ms (nonface), $F(1, 9) = 48.38$, $p < .001$, for Display Size 12; and 928 ms (face) versus 718 ms (nonface), $F(1, 9) = 21.10$, $p < .005$, for Display Size 18.

Indeed, facial organization slowed the curvature search even under the condition in which bounding circles were removed and the curvature search rate was in the range considered to be parallel, that is, the response time was independent of the display size. Not only do these results confirm the results from Experiment 1, but they further demonstrate the strength of facial organization in preempting the constituent low-level features. The results have

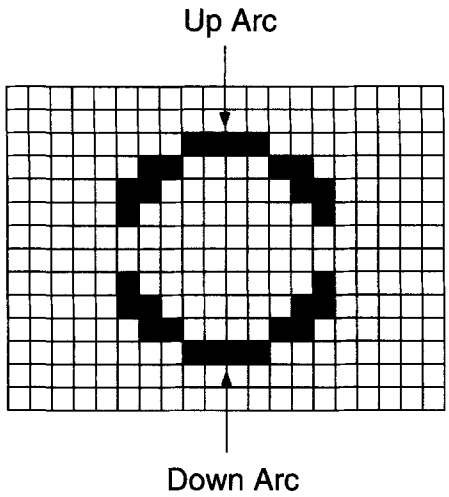

Figure 9. Depiction of the stimulus arcs (up arc and down arc) used in Experiment 3, shown in a grid of pixel-by-pixel squares. Each pixel subtended about 2′ of visual angle from a 70-cm viewing distance.

demonstrated that facial grouping can even override the grouping of arcs by similarity based on a low-level feature such as curvature. Note that the average distance between a pair of adjacent arcs was identical for the face and nonface conditions (30′ within and 1°12′ between triplets; see Figure 8). Thus, there was no difference in the local proximity of arcs between the two conditions.

Discussion

In the present experiments, we investigated the relative influence of low-level (curvature) and high-level (face) representations on visual search when either one or both were available in the image to discriminate the target. In particular, we expected that the results would tell us whether the search process has prior access to the local curvature-level representation or to the global face-level representation.

The main result (see Figure 4) of these studies is that the facial organization speeded the curvature-by-location conjunction search but slowed the curvature feature search. The first result indicates that a global representation such as facial expression can be an effective feature in visual search. The second result is rather surprising in that the curvature feature search was slower when the target was distinctive in both curvature and facial expression rather than when the target was distinctive only in curvature. The predictions from the first two models discussed in the introduction clearly contradict this second result. The first model assumed the sequential processing of image features from lower to higher level representations. Consequently, the high-level facial feature should not have affected the speed of the low-level curvature search because the curvature feature would be available to the search process before the facial feature. The second model assumed parallel processing of the curvature and facial features. In this case, the curvature feature search should have been facilitated, if anything, by the addition of the global facial feature because the search process could have taken advantage of the more efficient of the two features.

The unexpected results are explained by the third model that assumed that the search process would make prior access to the global face-level representation regardless of how rapidly the local curvature feature could be processed. This model's specific prediction depended on the relative efficiency (processing speed) of the global facial feature and the local curvature feature. If the global facial feature were processed faster, its addition would speed the curvature feature search, whereas if it were processed more slowly, its addition would then slow the curvature feature search. Our results indicate that the global facial feature was processed no faster (or slightly more slowly) than the local curvature feature. Specifically, our estimate of the search rate for the local curvature feature alone was 46 ms/item (the curvature feature search without facial organization; Figure 4, lower left). We estimated the search speed for the global facial feature to be 65 ms/item (the conjunction search with facial organization; see Figure 4). As described in the introduction, we can use the conjunction search with facial organization to estimate the processing speed for the global facial feature alone because the facial organization significantly speeded the conjunction search (65 ms/item vs. 95 ms/item; see Figure 4), indicating that the search speed in this condition was mediated principally by the global feature.

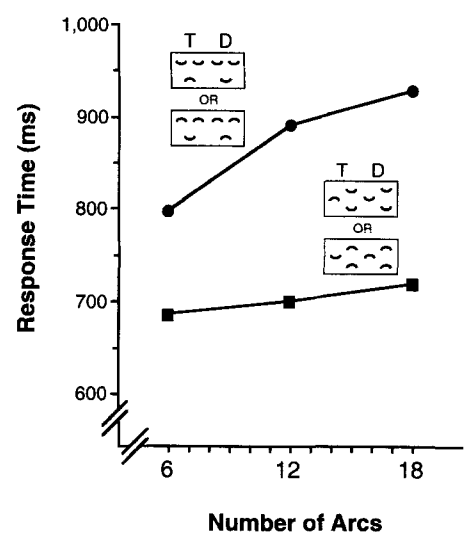

Figure 10. Mean search functions for the two experimental conditions for target-present trials (correct responses only) in Experiment 3: from top to bottom, feature (curvature) search with facial organization (filled circles) and feature search without facial organization (filled squares). Shown beside each search function are the two pairs of corresponding stimulus triplets. For each pair, the triplet containing the target (T) arc is shown on the left, and the triplet containing only the distractor (D) arcs are shown on the right. The top pair contains an up-arc target, whereas the bottom pair contains a down-arc target. Mean search functions for target-absent trials (not shown) had the same pattern.

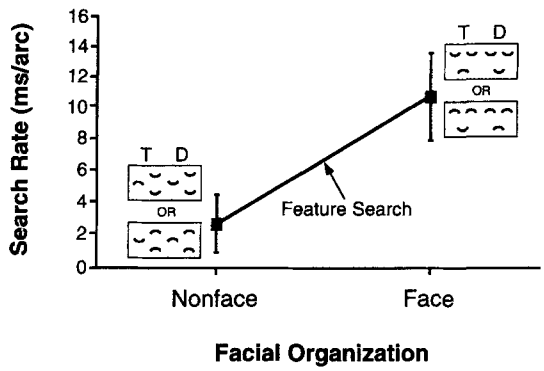

Figure 11. Mean search rates for target-present trials. Shown beside each search function are the two pairs of corresponding stimulus triplets. For each pair, the triplet containing the target (T) arc is shown on the left, and the triplet containing only the distractor (D) arcs are shown on the right. The top pair contains an up-arc target, whereas the bottom pair contains a down-arc target. Mean search rates for target-absent trials (not shown) had the same pattern. The vertical bars represent ±1 *SE*.

Thus, for our stimuli, the processing speed for global facial feature (65 ms/item) was similar or perhaps slightly slower than that for the local curvature feature (46 ms/item). This difference approached statistical significance, $F(1, 9) = 4.04$, $p < .08$. Given the slower processing of the global facial feature, the third model predicted that its addition to the curvature feature search could only slow it down. This prediction was found, as curvature search was significantly slowed when the global facial feature was added to the stimuli: The search rate increased from 46 ms/item to 67 ms/item (see Figure 4).

An important implication of this model is that if stimuli reach the "search" level in the form of complex gestalts, such as faces, the search is obliged to operate on those representations even though the lower levels of coding, for example, the curves within the face, could offer much faster processing. The lower level features have been preempted. The critical result from Experiment 1 was confirmed and extended in Experiment 3, demonstrating that facial organization slowed curvature search even if the contour circles were removed and the search times were rendered nearly independent of the number of arcs in the display.

With reference to Mermelstein et al.'s (1979) predictions, we have revised and extended their general claim of emergent feature precedence in two ways. First, emergent features are not necessarily more efficient than their constituent local features for shape discrimination. Second, the search process operates on an emergent feature (if present) regardless of its efficiency for a given target discrimination task and even if the search based on the constituent local feature could be fast and parallel (Experiment 3).

The effects of familiar organizations such as 3D objects, words, and faces on pattern detection and discrimination performances have been the focus of many studies. Those studies typically compare the performances on familiar patterns with those on feature-matched meaningless patterns. The question that is asked is whether global structures make patterns more detectable and discriminable than the sum of the constituent low-level features. As far as detectability is concerned, the existing evidence unanimously indicates that familiar organizations such as 3D objects, words, and faces are detected more robustly (the signal level required to detect the whole is less than that required to detect the constituent parts). The detection threshold (in terms of stimulus onset asynchrony [SOA] between stimulus and mask) for a 3D line drawing is lower than that for a feature-matched meaningless drawing (Purcell & Stewart, 1991). The detectability of a word is higher than that of a random letter string (Doyle & Leach, 1988). Faces are detected at lower thresholds (in terms of SOA) than upside-down and jumbled faces (Gorea & Julesz, 1990; Purcell & Stewart, 1988). These studies demonstrate that perceptual encoding is more robust for familiar patterns than for feature-matched meaningless patterns.

Pattern discrimination performance under poor viewing conditions is also generally higher for familiar patterns, provided their global features can be used for discrimination. The classification thresholds (in terms of SOA) are lower for 3D objects than for feature-matched meaningless patterns in the 3D-object versus meaningless-pattern classification task (Purcell & Stewart, 1991). The identification of the location and the orientation of a line segment flashed at near-threshold SOA is more accurate when the line is embedded in a coherent 3D line drawing (Weisstein & Harris, 1974; Williams & Weisstein, 1978; but see Klein, 1978; McClelland, 1978; Weisstein & Harris, 1980, for the effect of different masks, and Earhard, 1980; Earhard & Armitage, 1980; Weisstein & Harris, 1980, for the effect of the location of fixation point). The accuracy of identifying a letter at near-threshold SOA improves if it is a part of a word—word superiority effect (Reicher, 1969; Wheeler, 1970). The discrimination accuracy under poor viewing conditions is also higher for words than for single letters or meaningless letter sequences (Prinzmetal, 1992). Finally, faces are classified at lower thresholds (in terms of SOA) than jumbled faces in the face versus jumbled-face classification task (Purcell & Stewart, 1988). These results are consistent with the aforementioned hypothesis that familiar patterns are encoded more robustly. Specifically, discrimination under near-threshold viewing conditions should be more accurate if it were based on the more robustly encoded object representation rather than on the noisier encoding of local features. The robust encoding of familiar patterns has also been suggested by a memory study. Tanaka and Farah (1993) have shown that reinstatement of the encoding context helped to recognize facial parts only when the parts were remembered in face contexts but not when they were remembered in jumbled or upside-down face contexts. The results imply that when placed in face contexts, the subtle differences in the shape of facial parts are encoded more robustly as the more salient differences in the global impression of faces.

So far, we have reviewed the tasks that critically depend on the stimulus encoding process, such as detection and discrimination at near-threshold viewing conditions and rec-

ognition memory, and have seen that familiar patterns are encoded more robustly. Would familiar structures also speed the discrimination process under suprathreshold conditions? When response time is measured in pattern discrimination, the stimuli are usually presented for 1–2 orders of magnitude longer than the threshold SOA. Thus, in this case, the speed of pattern discrimination would be determined mostly by the speed of processing of encoded stimuli rather than by the strength of their encoding. Klein (1978) used the same set of stimuli that were originally used by Weisstein and Harris (1974) and first replicated the object superiority effect in discrimination accuracy in a near-threshold viewing condition. However, when Klein measured response time in a suprathreshold version of the task, he found that orientation discrimination was even slower when the target line was embedded in a 3D line drawing than when it was presented among a set of unconnected line segments that were the constituent line segments of the 3D drawing. This finding was replicated by Widmayer and Purcell (1982).

These seemingly paradoxical results are consistent with our third model that assumed that global representation dominates during speeded pattern discrimination, obscuring or preempting the lower level representations of the constituent parts; in other words, only global representation is "visible" to the rapid discrimination process. Even if the encoding of global feature representation is always more robust than that of local feature representation, the processing speed of global features may not necessarily be greater than that of local features. As for the stimuli used by Klein (1978) and Widmayer and Purcell (1982), the processing speed of the local target–line orientation would have been greater than that of the 3D structures as a whole. In an extreme example where global structures are irrelevant to the discrimination task, the presence of a global structure slows the discrimination of the local constituent features as expected (Mermelstein et al., 1979). In those studies, however, not only did the object and the meaningless versions of the stimuli differ in the global structure, but they also differed in lower level image features such as the number of terminators, closure, connectivity, and feature proximity. In our studies, all of these features were equated between the face and nonface conditions and showed that the global facial organization speeded visual search based on the curvature feature when it was a difficult conjunction search but slowed it when it was an easy feature search.

In summary, previous studies on detection and discrimination at near-threshold viewing conditions, recognition memory, and speeded discrimination at suprathreshold conditions as well as our current studies converge onto the following principles: First, familiar global structures are always encoded more robustly—they are detected and discriminated at lower thresholds because their representations are less degraded under poor viewing conditions. The robust encoding, however, does not necessarily lead to higher speed of processing. Depending on the particular stimuli used, low-level features could be processed more rapidly than global features or vice versa. Because our data suggest that the rapid search process makes initial access to the global-level representation (our third model), object superiority in response time occurs only if the global features of the stimulus set are more rapidly processed for target discrimination than the constituent low-level features, whereas object inferiority occurs if the low-level features are more rapidly processed (see Figure 4).

We next consider a possible neural mechanism subserving the robust (noise resistant) encoding of familiar global configurations. In the case of faces, a possible neural substrate might be found in the temporal lobe. Some cells in this area (more specifically, in the superior temporal sulcus and the inferior temporal cortex) have been identified to respond selectively to faces with response specificity for such facial characteristics as identity (Hasselmo, Rolls, & Baylis, 1989; Perret et al., 1984), gaze direction (Perret et al., 1985), facial expression (Hasselmo et al., 1989), and face shape (Malcolm & Yamane, 1992). These cells generally do not respond to meaningless patterns, and a majority of them either do not respond or respond significantly less to facial parts or jumbled faces (Perret, Rolls, & Caan, 1982).

More recently, Rolls and his colleagues showed that when the stimuli were flashed for only 16 ms, all the face-selective cells they studied made greater or equal responses to faces than to jumbled faces (Rolls, Tovee, Purcell, Stewart, & Azzopardi, in press). They further showed that the cell activity stayed above its spontaneous level for 300–400 ms after the brief stimulus presentation and the mask that the response extinguished. The latter result rules out an attractive hypothesis that the face superiority effects may be due to the face-selective cells' activity (i.e., face-level image processing) being impenetrable to the mask that neither has face-level representation nor activates the face-selective cells (see Purcell & Stewart, 1991). Alternatively, Rolls et al. pointed out that the total number of cells in the temporal lobe activated by faces should be much greater than those activated by jumbled faces on the basis of the numerous recording studies of face-selective cells to date; they suggested that this greater population of cells responding to faces underlies the face superiority effects. One should, however, take this type of neural explanation with due caution because researchers are still far from understanding how the responses of the face-selective cells relate to the perceptual effects of facial organization.

Besides demonstrating both the face superiority and inferiority effects using the same paradigm (visual search) and the same set of stimuli (up and down arcs), the present experiments have gone further to examine the strength of facial organization by showing that even an elementary image feature such as curvature can be preempted by facial organization, and this occurs even without the facial contours where the curvature features alone can support a fast parallel search. It is reasonable to assume that the mechanism of fast curvature search stems from the grouping of the distractors by similarity in curvature. Our results from Experiment 3 then imply that the local grouping of arcs by facial organization overrides the low-level grouping of arcs by curvature.

Our results, however, do not tell us what specific geo-

metric features in the facial organization engage the global face processing. We arranged three arcs on the vertices of an upside-down equilateral triangle to form our schematic faces, although this particular arrangement may not have been necessary. For example, to test the importance of vertical symmetry, one could laterally shift the mouth off the axis of symmetry and see how sensitive the effects of facial organization would be to the central placement of the mouth. Similarly, one could vary the eye-to-eye and eye-to-mouth distances to see what ratio would give the strongest and weakest effects of facial organization.

Finally, we point out that it is rather remarkable that the global face processing has been found to be engaged by the minimalist image of a face (a triangular arrangement of three arcs) used in our experiments. Using detailed gray-scale (8-bit) faces with potentially confounding nonfacial cues removed, Stewart, Purcell, and Skov (1993) reported search rates comparable to ours when a smiling face was the target among angry faces and vice versa. Display size was varied between four and nine faces, and the search rates ranged from 42 ms to 70 ms per face for target-present trials and from 101 ms to 149 ms per face for target-absent trials. This suggests that increased faceness does not necessarily improve the discriminability between smiling and angry faces. In fact, caricatures of familiar faces have been shown to be identified more quickly than the veridical line drawings where each caricature was generated by exaggerating all metric differences between a face and a norm (Rhodes, Brennan, & Carey, 1987). Thus, it may be the case that the differences between smiling and frowning faces have actually been amplified in our minimalist drawings.

In the present studies, we have demonstrated the strength of facial organization in grouping the constituent low-level features into a unitized face representation. Facial organization renders the constituent features inaccessible to the rapid search process. The unitizing effect of facial organization is so strong that merely arranging three arcs in a facial organization defeats the global grouping of the nontarget arcs and disrupts the parallel search for the oddly curved arc. These results are compatible with our daily experience of imagining faces in many inanimate objects, invariably in the fronts of cars and trucks and occasionally in rocks, clouds, and landscapes. Wherever we look, someone seems to be looking back.

References

Carey, S. (1982). Face perception: Anomalies of development. In S. Strauss (Ed.), *U-shaped behavioral growth* (pp. 169–190). New York: Academic Press.

Doyle, J. R., & Leach, C. (1988). Word superiority in signal detection: Barely a glimpse, yet reading nonetheless. *Cognitive Psychology, 20,* 283–318.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96,* 433–458.

Earhard, B. (1980). The line-in-object superiority effect in perception: It depends on where you fix your eyes and what is located at the point of fixation. *Perception & Psychophysics, 28,* 9–18.

Earhard, B., & Armitage, R. (1980). From an object-superiority effect to an object-inferiority effect with movement of the fixation point. *Perception & Psychophysics, 28,* 369–376.

Enns, J. T., & Rensink, R. (1990). Sensitivity to three-dimensional orientation in visual search. *Psychological Science, 1,* 323–326.

Enns, J. T., & Rensink, R. (1991). Preattentive recovery of three-dimensional orientation from line drawings. *Psychological Review, 98,* 335–351.

Gorea, A., & Julesz, B. (1990). Context superiority in a detection task with line-element stimuli: A low-level effect. *Perception, 19,* 5–16.

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.

He, Z. J., & Nakayama, K. (1992). Surfaces versus features in visual search. *Nature, 359,* 231–233.

Klein, R. (1978). Visual detection of line segments: Two exceptions to the object superiority effect. *Perception & Psychophysics, 24,* 237–242.

Malcolm, P. Y., & Yamane, S. (1992, May 29). Sparse population coding of faces in the inferotemporal cortex. *Science, 256,* 1327–1331.

McClelland, J. L. (1978). Perception and masking of wholes and parts. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 210–223.

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

Perret, D. I., Rolls, E. T., & Caan, W. (1982). Visual neurons responsive to faces in the monkey temporal cortex. *Experimental Brain Research, 47,* 329–342.

Perret, D. I., Smith, P. A. J., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., & Jeeves, M. A. (1984). Neurons responsive to faces in the temporal cortex: Studies of functional organizations, sensitivity to identity and relation to perception. *Human Neurobiology, 3,* 197–208.

Perret, D. I., Smith, P. A. J., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., & Jeeves, M. A. (1985). Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proceedings of the Royal Society of London, Series B, 223,* 293–317.

Prinzmetal, W. (1992). The word-superiority effect does not require a T-scope. *Perception & Psychophysics, 51,* 473–484.

Purcell, D. G., & Stewart, A. L. (1988). The face detection effect: Configuration enhances detection. *Perception & Psychophysics, 43,* 355–366.

Purcell, D. G., & Stewart, A. L. (1991). The object detection effect: Configuration enhances perception. *Perception & Psychophysics, 50,* 215–224.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology, 81,* 274–280.

Rensink, R., & Cavanagh, P. (1993). Processing of shadows at preattentive levels. *Investigative Ophthalmology & Visual Science, 34,* 1288.

Rhodes, G., Brennan, S., & Carey, S. (1987). Identification and ratings of caricatures: Implications for mental representations of faces. *Cognitive Psychology, 19,* 473–497.

Rolls, E. T., Tovee, M. J., Purcell, D. G., Stewart, A. L., & Azzopardi, P. (in press). The responses of neurons in the temporal cortex of primates, and face identification and detection. *Experimental Brain Research.*

Sergent, J. (1984). An investigation into component and configural

process underlying face perception. *British Journal of Psychology, 75,* 221–242.

Stewart, A. L., Purcell, D. G., & Skov, R. (1993, June). Another look at face in the crowd: An anger inferiority effect. Poster presented at the convention of the American Psychological Society, Chicago.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology, 46A,* 225–245.

Thompson, P. (1980). Margaret Thatcher: A new illusion. *Perception, 9,* 483–484.

Treisman, A., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology, 12,* 97–136.

Wang, Q., Cavanagh, P., & Green, M. (1994). Familiarity and pop-out in visual search. *Perception & Psychophysics, 56,* 495–500.

Weisstein, N., & Harris, C. S. (1974, November 22). Visual detection of line segments: An object superiority effect. *Science, 186,* 752–755.

Weisstein, N., & Harris, C. S. (1980). Masking and the unmasking of distributed representations in the visual system. In C. S. Harris (Ed.), *Visual coding and adaptability* (pp. 317–363). Hillsdale, NJ: Erlbaum.

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

Widmayer, M., & Purcell, D. G. (1982). Visual scanning of line segments: Object superiority and its reversal. *Bulletin of the Psychonomic Society, 19,* 353–354.

Williams, A., & Weisstein, N. (1978). Line segments are perceived better in a coherent context than alone: An object-line effect in visual perception. *Memory & Cognition, 6,* 85–90.

Wolfe, J. M., Yee, A., & Friedman-Hill, S. R. (1992). Curvature is a basic feature for visual search tasks. *Perception, 21,* 465–480.

Young, A. W., Hallwell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception, 16,* 747–759.

Received November 8, 1993
Revision received August 1, 1994
Accepted August 26, 1994 ■