Three-dimensional Features Facilitate Object Recognition

Muriel Boucart
Hopitaux Universitaires de Strasbourg, INSERM, Strasbourg, France

Jonathan Grainger
CREPCO, Université de Provence, Aix-en-Provence, France

Ludovic Ferrand
Université Paris V René Descartes, Paris, France

A priming paradigm was used to investigate the contribution of local features and global shape information to object recognition. Five types of incomplete forms were used as primes: (1) forms with both maxima (local curvatures) and mid-segments of edges present and aligned on the outline contour; (2) forms similar in global shape to the first version of stimuli, but with misaligned elements; (3) forms with only maxima; (4) forms with only midsegments of edges; and (5) forms containing 3D corner junctions (cf. Experiments 3 and 4). The target was an outline drawing of an object from which the incomplete prime was derived. Subjects were asked to name the target as rapidly as possible. Primes were presented at levels of contrast corresponding to identification thresholds, as well as above and below threshold levels (determined in Experiments 1 and 3). Facilitation effects relative to a neutral (no prime) condition occurred at threshold and above threshold for primes with aligned elements, forms with only maxima, and forms with only midsegments. Priming occurred only above threshold for forms with non-aligned elements. In Experiment 4 the presence of 3D local features increased the magnitude of priming relative to forms with midsegments and to forms with flat corners (in Experiment 2). This result suggests that 3D features facilitate object identification.

Requests for reprints should be sent to M. Boucart, Hopitaux Universitaires de Strasbourg, INSERM (U405), Département de Psychiatrie 1, 67091 Strasbourg cedex, France. Email: boucart@alsace.u-strasbg.fr.

© 1995 Erlbaum (UK) Taylor & Francis
Most recent studies of object processing have been aimed at specifying the physical properties that mediate object recognition. Proposals have either ignored the influence of global shape information or the role of more local features such as the existence of a limited set of volumetric entities ("geons") common to all objects (Biederman, 1987) Global shape information has been defined both in terms of a very coarse level of description based on low spatial frequencies (Ginsburg, 1986; Watt, 1988, 1994) or a raw two-dimensional view (Sanocki, 1993), and in terms of more detailed levels of descriptions, including global contour and attributes of objects (Mervis & Rosch, 1981; Riddoch & Humphreys, 1987; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976; Snodgrass & Vanderwart, 1980). For instance, Rosch et al. (1976) proposed that the rapid classification of visual objects into semantic categories can be explained by a high degree of global perceptual similarity between members of the same category, at least at the basic level of abstraction that constitutes the entry level of semantic representations. Though this proposal holds mainly for typical exemplars of a category (Jolicoeur, Gluck, & Koslyn, 1984; Riddoch & Humphreys, 1987), several studies have shown that picture naming, semantic categorization, and picture matching were affected by global physical similarity (including the global configuration and the number of common parts) between objects (Keltner et al., 1984; Riddoch & Humphreys, 1987; Rosch et al., 1976; Snodgrass & McCullough, 1986; Vitkovitch, Humphreys, and Lloyd-Jones, 1993). For instance, Humphreys, Riddoch, and Quinlan (1988) showed that structurally similar objects were slower to name than were structurally distinct objects. The lower identification time was interpreted as reflecting the competition between a large number of stored representations of objects activated by the presented item, whereas fewer candidates are activated for objects that are not structurally similar to other objects from the same semantic category. Consistent with this account, Vitkovitch et al. (1993) reported a wider range of within-category naming errors for pictures from semantic categories with a high degree of between-item structural similarity (e.g., objects from natural categories) than for pictures containing structurally distinct objects (e.g., non-living things) under speeded naming conditions.

Biederman (1987) proposed a more "locally based" account in which object recognition relies upon the activation of a viewpoint-invariant structural description specifying the local attributes of a form (edges, vertices, and parts) and their relations. Empirical evidence for this account was provided in a naming task involving incomplete pictures. This study showed that deletion of vertices on the
outline of the form was more disruptive for identification than was dele-
tion of local edges at midsections of lines or curves, suggesting that some local
features (vertices or junctions of parts) are more critical than others (edges) for
distinguishing the component parts of objects and their interrelations.

Of greater relevance for the present study is the dissociation between the role
played by global and local shape information in priming studies of object recog-
nition. In a series of experiments designed to investigate the type of physical
information that was responsible for priming effects in a picture-naming task,
Birdman and Cooper (1991) tested the contribution of (1) local features (edges
and vertices), (2) convex or concave components corresponding to the object's
parts, and (3) object models. In a first phase, subjects named a series of incom-
plete pictures of objects. Half of the edges and vertices of each component were
deleted. In a second session approximately seven minutes after the first, subjects
were asked to name a further series of pictures, which were either the same as
those in the first phase, the complementary version composed of the other half
of the edges and vertices of each component of the object presented in the first
phase, or a different model of the objects seen in the first phase (e.g. an elephant
seen from two different viewing angles). In both the first and the second phase
of the experiment, only 50% of the contour of each of the object was presented.
The results showed equivalent facilitation for identical and complementary
pictures and less facilitation relative to these two conditions for a different model
of the object. From this result it was argued that activation of object representa-
tions can be attributed to the explicit presentation of the component parts of the
object rather than to local features (edges and vertices) or the explicit model.

The assumption that objects are parsed into constituent parts as a prelude to
identification was questioned by Baker-Cove and Kosslyn (1993). They investi-
gated the role of part descriptions in object identification by comparing naming
performance for complete outline drawings of objects, for pictures parsed along
"natural" or "unnatural" part boundaries—that is, at the regions of junctions or
arbitrary locations—and for scrambled pictures. In this latter condition all the
object parts were present, but their spatial arrangement was disrupted. They
found that performance was affected by degradation of the global configuration
of the objects: that is, accuracy was lower in the scrambled condition but was not
affected by whether objects were parsed at natural or unnatural locations (i.e.
whether or not the parts were recoverable). However, Baker-Cove and Kosslyn
found that the type of parsing started to affect performance when the exposure
time was reduced (200 msec instead of a display until response) and a mask
followed the display. In this presentation condition unnaturally parsed objects
were named less accurately than were naturally parsed objects. From these
results the authors proposed that object recognition proceeds by matching the
figure as a whole to models stored in memory and that parsing occurs when the
global configuration is degraded. Along the same lines, Sanocki (1993)
proposed that the earlier availability of the global structure of an object may
provide a rough reference to delimit a set of candidate objects among stored representations. Global information is then used to interpret and constrain the analysis of more detailed features that are integrated later. This prediction was tested in a priming study (Sanocki, 1993, Experiment 4) by manipulating the order of presentation of the prime and the target. The target could either precede or follow the prime. In both of these conditions, the prime corresponded either to the global structure of the picture (defined as the external contour) or to local details (defined as internal contours). The target contained both global and local information. The results showed that the degree of priming was affected by the order of presentation of global and local primes. Facilitation was more pronounced when global information preceded the target, and the effect was reversed with an advantage for local information when the prime followed the target, thus suggesting that the activation of a global representation facilitates the processing of local details used to distinguish between object alternatives.

The present study was designed to test the contribution of rough global structure and local features in the activation of stored representations of objects. In contrast to Biederman and Cooper (1991), who demonstrated the role of parts in priming effects but who did not test which local features were important, we investigated the role of maxima corresponding to local curvature in the outline contour and to junctions between object parts and the role of local edges. Biederman (1987) showed that deletion of junctions was more disruptive for object identification than deletion of midsegments of edges, suggesting that junctions help the determination of object parts. We compared priming effects for four types of primes derived from outline drawings of objects: (1) incomplete objects with both maxima (corners) and midsegments of edges present and aligned on the outline contour, (2) the same pictures with midsegments of edges deleted and only maxima present on the outline contour, (3) the same pictures with maxima deleted and only midsegments of edges present, and (4) pictures with both maxima and midsegments of edges present but not aligned on the outline contour of the object. For these versions of the forms, the rough global structure was similar to that of pictures with aligned elements (Version 1). The similarity in global structure was increased by presenting the prime stimuli at contrast thresholds for identification (determined in Experiment 1 for each subject) and for short exposure durations, thus leaving low spatial frequencies, conveying global information, primarily available for visual processing (Breitmeyer & Gass, 1976; Legge, 1978). Thus the manipulation of the physical information present in the prime allows us to compare the efficiency of (1) global shape information, (2) local parts (the components of the object), and (3) local features (maxima on the contour and midsegments of edges) in the activation of stored object representations. Similar performance for pictures with collinear elements and for pictures with non-aligned elements would suggest that the activation of structural representations of objects is mediated by global form information rather than via the component parts of the object, as parts are
more difficult to discriminate in forms with non-aligned elements. A larger priming effect for pictures with collinear elements than for pictures with non-aligned elements would indicate that the identification of the component parts of the object and their spatial relations is more efficient than is global shape information in activating object representations. A larger priming effect for objects having only maxima present than for objects having only midsegments present would suggest that local curvature provide more information about the object structure than do midsegments of edges.

In addition to the manipulation of physical information, our paradigm is procedurally different from that used by Biederman and Cooper (1991). Instead of presenting two series of pictures separated by long delays (long-lag priming), we used a procedure in which prime stimuli were presented very briefly, followed by a mask stimulus and then the target with very short delays (masked priming). Variants of this masked priming paradigm have proved very useful in the study of basic orthographic and phonological processes in visual word recognition (Evett & Humphreys, 1981; Ferrand & Grainger, 1992; Forster & Davis, 1984; Segui & Grainger, 1990), and picture and word naming (Ferrand, Grainger, & Segui, 1991). Also, by varying prime exposure duration and/or contrast, this paradigm has been explored as a tool for uncovering the time course of processing of visual and phonological codes in letter and word perception (cf. Ferrand & Grainger, 1993; Jacobs, Grainger, & Ferrand, in press). It is this masked incremental priming technique that is adapted to the study of pictorial stimuli in the present research. Compared to long-lag priming techniques which may be useful for the study of information retention and decay over relatively long periods of time, we claim that the masked priming paradigm is a more appropriate means of examining the fast developing (and perhaps fast-decaying), highly automatized processes underlying normal object recognition (see Jacobs et al., in press, for a more detailed discussion of the advantages of the incremental priming technique).

GENERAL METHOD

Unless otherwise mentioned, the same method was used for each of the four experiments.

Stimuli. The stimuli were 15 outline drawings of objects (squirrel, bear, rabbit, mouse, tortoise, bird, lorry, plane, car, tractor, sailing-boat, rain, mushroom, pear, strawberry) and 4 types of incomplete form derived from each outline drawing. The incomplete forms had an outline contour made of line segments. For all types of forms, the length of the segments varied between 4 and 8 pixels. In one version, the line segments were collinear (COLL) on the contour, with both the maxima of the outline contour and midsegments of edges represented. Relative to the outline drawing from which the figures were
derived, the mean percentage of contour present was 58.4 (ranging from 56.76% to 60.04% according to the figures). In a second version, only the maxima (max) were drawn on the contour of the objects, and in a third version only the midsegments of edges (mseg) were present. For these two latter types of forms, the mean percentage of contour present, relative to the complete figure from which they were derived, was 51.2 (ranging from 29.41% to 32.99% according to the figures). The number of pixels was exactly the same for pictures with only maxima present and pictures with only midsegments of edges present on the outline contour. The fourth version of incomplete forms (n-col) was derived from the figures whose contour was made of collinear line segments. The collinearity was reduced either by a rotation or by a shift of the line segments ranging from one to three pixels inside or outside the contour relative to forms made of collinear elements. The number of line segments and the number of pixels was exactly the same for forms with collinear and non-collinear elements. Examples of the four different types of forms are displayed in Figure 1 (A, B, C, & D). At a viewing distance of 57 m, the mean angular size of the stimuli was 2.5° horizontally and vertically (ranging from 1.8° to 3°).

Apparatus. The stimuli were displayed on a black-and-white video monitor. There were generated through a 386 DX micro-computer equipped with a VGA (Video-7) graphic card. The screen resolution was 640 × 480 pixels. In a dark room the luminance of the background was set at 34 cd/m². The background was gray. The luminance of the primes was an experimental variable. The luminance of the target was fixed. It was displayed in black on the gray background. A fixation point subtending 0.039° was centrally displayed. Two response keys connected to the computer were used in Experiments 1 and 3. A voice key was used for response in Experiments 2 and 4.

Procedure. The procedure is described separately for each experiment.

EXPERIMENT 1

In Experiment 1 we measured identification thresholds for each subject as a function of the type of incomplete form. This experiment was also designed to test whether identification thresholds were affected by the type of information present in the contour: the collinearity of the elements, the presence of maxima or of midsegments of edges, and the amount of contour present.
FIG. 1. Examples of stimuli used in the four experiments. In Experiments 2 and 4, the complete outline drawing served as target and the following four incomplete forms served as primes: (A) forms with both maxima and midsegments of edges aligned on the contour; (B) forms with only maxima present; (C) forms with only midsegments of edges present; and (D) forms with both maxima and midsegments present on the contour but not aligned. In Experiments 3 and 4, forms with flat corners (B) were replaced by forms containing 3D local features (Y, T, or arrow-like shapes) in addition to maxima (E).
Method

Subjects. The subjects were four students in experimental psychology and two of the authors (MB and LF). The six subjects participated in Experiments 1 and 2. All had normal or corrected-to-normal vision and were well practiced in psychophysical experiments. The six subjects were native French speakers.

Procedure. Before the experiment, subjects were asked to study the complete outline drawings of objects and the four types of incomplete forms derived from the objects. The object names were also presented in order to ensure that each subject would use the same name for a given picture. Identification thresholds were measured by means of an ascending method of limits. The stimuli were only incomplete forms. A picture was centrally displayed for 4 frames (64 msec). After a delay of 64 msec, a mask was centrally presented. The mask was a 100 × 100 pixel matrix with 50% black and 50% white dots randomly distributed. We used three different masks in order to reduce the possibility that a random structure in the dots would mask one picture more than another. An example of a mask is presented in Figure 2. One of the three masks was randomly presented after the picture. For each picture the threshold measurement started with the picture presented at the luminance of the gray background (34 cd/m²). When the subject pressed the response key on the right, the luminance of the stimulus decreased—that is, the stimulus became darker than the background. The luminance decreased linearly by steps of 0.1 cd/m² when the right key was pressed. The luminance of the mask was fixed (black and white dots). The task was to name the picture correctly. An experimenter recorded the response, and when the correct name was given the experimenter recorded the level of luminance, and the subject pressed the left key to start a new threshold measurement with a different picture. If the name was not correct, the subject was told to continue decreasing the luminance.

The 15 stimuli for each type of incomplete form were randomly presented. To ensure that the thresholds were stable, two repetitions were performed for each.

FIG. 2. An example of a mask used in the four experiments and an illustration of the spatial location of the target relative to the prime. The target was displayed randomly half a degree to the left or right of fixation. The prime was centered on the fixation point.
picture on different days. For each session of threshold measurements, the experiment lasted about one hour. Subjects were adapted to the luminance of the background for 10 min before the beginning of each session.

Results
An analysis of variance (ANOVA) was carried out on the identification thresholds both with subjects ($F_1$) and with figures ($F_2$) as random variables for the four types of incomplete forms. The results are graphically presented in Figure 3 for each subject.

There was a significant main effect of type of form, $F_1(3, 15) = 46.02, p < 0.001$, and $F_2(3, 42) = 42, p < 0.001$. The lowest threshold (contrast = 9.91%) was found for forms having collinear elements and more than 50% contour present. This was followed by forms with the same amount of contour present but with non-collinear line segments (contrast = 11.38%). The difference between the two types of stimuli was significant, $F_1(1, 5) = 35.2, p < 0.001$, and $F_2(1, 14) = 37.8, p < 0.001$. The highest threshold was observed both for forms with contours made up of maxima (contrast = 12.95%) and for forms with contours made up of midsegments of edges (contrast = 13.1%). There was no significant difference between these two types of forms (both $F_2 < 1$). Planned comparisons indicated that there was a significant difference in contrast threshold between forms with collinear elements and forms with maxima, $F_1(1, 5) = 74.85, p < 0.001$, and $F_2(1, 14) = 92, p < 0.001$, and forms with midsegments of edges, $F_1(1, 5) = 54, p < 0.001$, and $F_2(1, 14) = 139.6, p < 0.001$. The thresholds were also significantly lower for forms with non-collinear elements than for forms with maxima, $F_1(1, 5) = 52.23, p < 0.001$, and $F_2(1, 14) = 25.33, p < 0.001$, and forms with midsegments of edges, $F_1(1, 5) = 36.8, p < 0.001$, and $F_2(1, 14) = 22.9, p < 0.001$.

Discussion
The results of Experiment 1 show that contrast thresholds for object identification were affected by two parameters: (1) the degree of structure of the pictures and (2) the energy in terms of the surface (the amount of contour) present in the stimuli. The effect of structure appeared in the lower identification thresholds for pictures with contours composed of collinear elements compared to pictures with non-collinear elements. This result cannot be attributed to energy differences given that the two types of forms had exactly the same amount of contour present. Rather, this finding indicates that collinearity improved object identification in facilitating the computation of contour information. Similar effects have been reported by Boscott and Humphreys (1992a, 1992b) in matching tasks on incomplete forms with collinear or non-collinear local contour elements. Collinearity renders external and internal boundaries of the object more discriminable and thus renders the different objects less confusable.
FIG. 3. Contrast thresholds for identification averaged over the 15 pictures for each type of incon- 
pinite bar—collinear elements (COLL), non-collinear elements (N-COLL), maxima (MAX), and 
minima of edges (MIN); and for each subject in Experiments I. The vertical bars represent 
standard deviation values.

460
Identification thresholds were affected by the energy content of the stimuli. Lower thresholds were found for pictures having a large amount of contour present (58.4% on average for pictures with collinear elements and pictures with non-collinear elements) than for pictures with less contour present (31.2% on average for pictures with only maxima present and pictures with only midg- ments present). This result reflects the summation of light over space by the visual system (Hood & Pinkerstein, 1986).

The results show that, except for two subjects (DP and CL) showing oppos- ite trends, thresholds were equivalent for pictures with maxima and pictures with only midg- ments of edges present in their contour, suggesting that the two types of local features did not affect object identification at contrast threshold.

In experiment 2 we tested whether the identification of a complete outline drawing of an object is facilitated by the prior presentation of an incomplete version of the same object with varying types of physical information present in its outline contour.

EXPERIMENT 2

Method

Procedure. The stimuli were the same as those used in Experiment 1. The sequence of stimuli on a trial was as follows: A fixation point was displayed for 500 msec and followed 500 msec later by a prime presented for 64 msec. After an interval of 64 msec, one of the three masks was presented for 64 msec. There was a second delay of 4 frames (64 msec) followed by the target presented for 64 msec. The target was always a complete outline drawing of an object. There were five priming conditions, with the prime stimulus being either one of the four incomplete versions of forms derived from the object presented as target or no prime before the mask (neutral condition). In order to avoid a summation of ‘luminance due to the spatial superimposition of the prime and the target,2 the centre of the target was located randomly 0.5’ to the left or right of the centre of the prime stimulus. An example of this is displayed in Figure 2. Five levels of contrast were used for the prime. These were based on the contrast thresholds obtained in Experiment 1 and were adjusted to individual subject thresholds. A first level of contrast (L1) was below the identification threshold of all stimuli. A second level of contrast (L2) corresponded to the identification threshold of forms with collinear elements. A third level of contrast (L3) corresponded to the threshold obtained for forms with non-collinear elements. The fourth level of contrast (L4) was the mean threshold for figures with maxima and figures with midg- ments of edges. For four subjects (CA, MB, LF and SD, see Figure 3) the

---

2Luminance summation effects resulting in an improvement in accuracy of the identification of a target letter have been reported by Hellige, Walch, Lawrence, & Brosa (1979) when target and mask, which were both letters, overlapped exactly.
threshold was the same or very similar for the two types of forms. For two other subjects (CL and DP), the threshold was lower either for forms with maxima or for forms with midsegments of edges. For the six subjects the mean threshold was taken as a fourth level of contrast (L4). The value chosen as the fifth level of contrast (L5) was above identification threshold for all stimuli. As there was a variability between pictures for each type of incomplete form, the values chosen for the levels of contrast L2, L3, and L4 were the mean values—that is, those corresponding to the thresholds represented in Figure 3 for each subject and for each type of form. The target was always presented in black on gray background. Subjects were asked to name the target object as rapidly as possible. The onset of the target activated the computer clock, which was stopped when the subject’s verbal response triggered a voice key linked to a headset microphone. The intertrial interval was fixed at 2000 msec after the response.

The five levels of contrast, the five conditions of prime (four types of incomplete forms plus the neutral condition), the spatial location of the target relative to the centre of the prime stimulus (left/right), and the 15 different figures of each type of form were randomly and equally represented. There were 750 trials in three blocks of 250 trials each run on different days. Each block lasted approximately 25 min.

Results
Fewer than one percent errors were made to target objects. Response times (RTs) lower than 150 msec and higher than 1500 msec were discarded (less than 1% of the data). Two ANOVAs were conducted on the mean RTs: one with subjects as random factor (F1) and one with figures as random factor (F2). There were three-within subject factors (1): The spatial location of the target relative to the prime (left vs. right), (2) the levels of contrast (L1 to L5), and (3) the type of prime (neutral condition and the four types of incomplete forms).

The mean RT was 658 msec. There was no significant main effect of the spatial location of the target (left, 643 msec; vs. right, 654 msec; both Fs < 1). There was no interaction between the spatial location of the target and the other experimental variables. The results graphically presented in Figure 4 are therefore averaged over each position of the target.

RTs decreased significantly with an increase in contrast (L1, 676 msec; L2, 663 msec; L3, 660 msec; L4, 650 msec; L5, 640 msec), F(4, 20) = 3.02, p < 0.04, and F(2, 56) = 11, p < 0.001. A significant main effect of priming condition was observed, F(4, 20) = 13.9, p < 0.001, and F(2, 56) = 58, p < 0.001. The longest RT was found for the neutral condition (691 msec). Relative to the neutral condition, there was a priming effect for figures with collinear elements (by 92 msec), F(1, 5) = 17.9, p < 0.008, and F(1, 14) = 119, p < 0.001, and for figures with maxima (by 28 msec), F(1, 5) = 7.2, p < 0.043, and F(1, 14) = 42.1, p < 0.001. A facilitation effect, relative to the neutral baseline, was also
observed for forms with midsegments of edges (by 17 msec) and for forms with non-collinear elements (by 29 msec), but these effects were significant only with pictures as the random variable—midsegments, $F_{(1, 14)} = 28.2, p < 0.001$; non-collinear, $F_{(1, 5)} = 4.3, p < 0.09$, and $F_{(1, 14)} = 38.7, p < 0.001$.

Planned comparisons for the different types of forms showed that naming times were significantly faster when the target was preceded by forms with
collinear elements (599 msc) than when it was preceded by forms with non-collinear elements (622 msc). \( F_1(1, 5) = 37.5, p < 0.001 \), and \( F_2(1, 14) = 70.6, p < 0.001 \). Naming times were shorter when the primes were pictures with maxima than when the targets was preceded by pictures with midsegments of edges (663 vs. 674 msc). \( F_1(1, 5) = 8.4, p < 0.03 \), and \( F_2(1, 14) = 5.59, p < 0.031 \). There was a significant difference in naming times between forms with collinear elements and forms with maxima as primes (599 vs. 663 msc), \( F_1(1, 5) = 12.7, p < 0.016 \), and \( F_2(1, 14) = 68.37, p < 0.001 \), and between forms with collinear elements and forms with midsegments of edges as primes (399 vs. 674 msc) \( F_1(1, 5) = 22.62, p < 0.005 \), and \( F_2(1, 14) = 117.4, p < 0.001 \). Forms with non-collinear elements as primes gave rise to faster naming times for the target than forms with midsegments of edges (662 vs. 674 msc), \( F_1(1, 5) = 2.02, \) n.s., and \( F_2(1, 14) = 12.06, p < 0.007 \). No difference was found between forms with non-collinear elements and forms with maxima (662 vs. 663 msc).

The type of prime interacted significantly with the level of contrast, \( F_1(16, 80) = 2.23, p < 0.01 \), and \( F_2(16, 224) = 2.24, p < 0.001 \). There was no priming effect for the four types of forms when the level of contrast was below identification threshold (L1). For forms with collinear elements, there was a priming effect (86 msc) at the level of contrast corresponding to the identification threshold for this type of form—L2: \( F_1(1, 5) = 9.9, p < 0.027 \), and \( F_2(1, 14) = 77.5, p < 0.001 \). There was also a priming effect (71 msc) at the level of contrast just above threshold for this type of form—L3: \( F_1(1, 5) = 6.4, p < 0.057 \), and \( F_2(1, 14) = 46.6, p < 0.001 \)—and for the two other levels of contrast—L4 (148 msc): \( F_1(1, 5) = 15.4, p < 0.011 \), and \( F_2(1, 14) = 58.7, p < 0.001 \); L5 (125 msc): \( F_1(1, 5) = 51.3, p < 0.001 \), and \( F_2(1, 14) = 45.7, p < 0.001 \).

For forms with maxima, a priming effect was found at the level of contrast corresponding to identification threshold for these pictures—L4 (41 msc): \( F_1(1, 5) = 6.45, p < 0.005 \), and \( F_2(1, 14) = 12, p < 0.004 \)—and above the threshold—L5 (52 msc): \( F_1(1, 5) = 7.5, p < 0.04 \), and \( F_2(1, 14) = 21, p < 0.001 \). The same effect was found for primes with midsegments of edges—L4 (46 msc): \( F_1(1, 5) = 7.8, p < 0.038 \), and \( F_2(1, 14) = 7.7, p < 0.014 \); L5 (51 msc): \( F_1(1, 5) = 6.98, p < 0.045 \), and \( F_2(1, 14) = 15.2, p < 0.001 \). There were no significant priming effects for forms with maxima and for forms with midsegments of edges at the levels of contrast L1, L2, and L3 that were below identification threshold for these forms.

For primes with non-collinear elements, the identification threshold was that corresponding to the level of contrast L3. No significant priming effect was found at this level of contrast (L3). There was a priming effect at the two levels of contrast that were above threshold for this type of prime, but this was significant only with pictures as the random variable—L4 (49 msc): \( F_2(1, 14) = 33.8, p < 0.001 \); L5 (40 msc): \( F_2(1, 14) = 15.2, p < 0.001 \).
Discussion

The results show, first, that the naming of an outline drawing of an object was facilitated when it was preceded by a prime corresponding to an incomplete version of the picture relative to a no-prime condition. The facilitation was larger for forms with collinear elements than for forms with the same amount of contour but non-collinear elements. Priming was globally more pronounced for forms with maxima than for forms whose outline contour was composed of midsagittal edges of edges. However, the magnitude of priming was equivalent for the two types of stimuli at levels of contrast corresponding to thresholds for identification and above threshold levels. Second, with the exception of forms having non-collinear elements, target naming was facilitated when the primes were displayed at levels of contrast corresponding to, or greater than, the identification threshold for each type of form but not below contrast threshold. Priming effects occurred only at contrast values above threshold for primes with non-collinear elements.

The faster RTs found when forms were primed by stimuli with collinear elements than for forms primed by non-collinear stimuli, despite the primes having the same amount of energy, suggest that object recognition is not uniquely based on global shape information derived from low spatial frequencies. This type of description might be useful to delimit broad classes of objects that differ in terms of their global shape information but of less use for pictures having a high within-category global shape similarity, as was the case here.

The increase in priming with increases in contrast, observed for forms with collinear elements but not for forms with non-collinear elements, is in favor of a part-based account of object identification. With increases in contrast, object parts and edges conveyed by medium and high spatial frequencies (Glissburg, 1984, 1986) become visible. The visibility of parts was further increased by the alignment of local edges providing information about internal boundaries. Numerous electrophysiological and psychophysical data show that collinearity is a critical physical parameter for early processing of contour, especially for binding processes. For instance, Petechnis and van der Heyd (1991) recorded the activity of cells in area V2 of the cortex of monkeys presented with bars composed of dots. They observed that a misalignment as small as 2 minutes of arc was sufficient to reduce by half the magnitude of the response of cells, and a larger misalignment totally abolished the response. Collinearity and curvilinearity have also been shown to facilitate the segregation of lines composed of discrete elements embedded in textures (Beck, Rosenfeld, & Ivry, 1999; Field, Hayes, & Hess, 1993, Northoff, 1992; see Boccarti, Delord, & Giaccà, 1996 for a review). With non-collinear local edges, component parts could be distinguished on the basis of local densities of elements (see Figure 1), but these local structures were probably less efficient than aligned edges for segmentation of the form into component parts, thus resulting in less priming for these forms.
Interestingly, the increase in contrast did not result in an increase in priming for forms with non-collinear elements. The magnitude of facilitation and the mean naming times did not vary at contrast threshold for identification and above threshold for these forms. This result suggests that when the binding of local contour elements is impaired due to a lack of collinearity, later cognitive processes involving the use of object representations are of little help to compensate for the defect. This result is inconsistent with Sanocki’s (1993) account that rough global shape information is used to guide the interpretation of later integrated local information. With optimal contrast conditions (black on white pictures), Boucart and Humphreys (1992a) showed that semantic information interfered with matching based on physical information for forms with collinear elements but not for forms with non-collinear elements, even though the complete outline drawing from which the two types of forms were derived preceded the incomplete forms. Together with the present data, this result suggests that the activation of stored representations of objects does not help contour integration if easy automatic integration processes are impaired by lack of collinearity.

On average, priming effects were larger for forms with maxima than for forms with midsegments of edges. Nevertheless, the magnitude of priming was equivalent at contrast values corresponding to identification threshold and above threshold for these two types of forms. The result suggests that local corners at the level of junctions did not provide more information than did local edges in delimiting object parts. If objects are stored in the form of volumetric entities, as suggested by Biederman (1987) and Biederman and Gerhardt (1993), local features providing information about three-dimensionality ought to be more efficient than two-dimensional edges in identifying objects and in producing priming. This was examined in Experiments 3 and 4.

Except for forms with non-collinear elements, the priming effects were tied to the identification threshold determined in Experiment 1. Facilitation for the naming of target objects occurred when the primes were presented at contrast values corresponding to threshold or above threshold but not below threshold. This result suggests that priming effects appear when the contrast is high enough to make available the physical information allowing the activation of the smallest number of candidate objects among stored object representations. As mentioned before, low spatial frequencies conveying information about rough global structure are not sufficient to distinguish between objects having a high degree of global shape similarity. Identification of local structures is necessary to reduce the number of candidate objects activated, and the identification of local structures (parts) was impaired in forms lacking collinearity.

In Experiment 3 we examined further the contribution of local features (midsegments of edges and maxima) to the activation of stored object representations. Here we used maxima providing information about three-dimensionality rather than flat corners.
EXPERIMENT 3

Experiment 3 was designed to determine the levels of contrast to be used in the priming task of Experiment 4.

Method

Subjects. The six subjects were members of the medical staff at the Department of Psychiatry of the Hopital Universitaire de Strasbourg. Their age ranged between 25 and 30 years. All of them had normal or corrected-to-normal vision. They were native French speakers. The subjects were not experienced in psychophysical experiments.

Stimuli. Two versions of incomplete forms derived from outline drawings of objects were used as stimuli. In one version, only midsegments of edges were drawn on the outline contour. These stimuli were the same as those used in Experiments 1 and 2. The second version was derived from the forms with maxima used in the two previous experiments. In addition to some maxima, local features providing information about three-dimensionality were drawn at the levels of junction between parts. These local features were Y, T, and arrow-like forms. Examples of the stimuli used are presented in Figure 1(c). Each pair of the two types of stimuli contained the same number of pixels.

Apparatus. The stimuli were displayed on the black-and-white CRT screen of a 256 X 256 micro-computer. The combinations of red, green, and blue signals used to define the levels of luminance were carefully calibrated with a luminance meter (Minolta LS 100) to the same values as those in Experiments 1 and 2. Two keys connected to the computer were used for response.

Procedure. The procedure was exactly the same as that described in Experiment 1. Each picture was first presented at the level of luminance corresponding to that of the background (34 cd/m^2) for 64 msec. It was followed, 64 msec later, by a noise mask presented for 64 msec. Subjects were asked to decrease the luminance of the picture until correct identification by pressing a response key. As in Experiment 1, subjects had been presented with the complete stimuli on a sheet of paper before the experiment. They were tested in two sessions of threshold measurements on separate days. Pictures with midsegments and pictures with 3D features were randomly presented.
FIG. 5. Contrast thresholds for identification averaged over the 13 pictures for the two types of incomplete forms used in Experiments 3 and 4: forms with 3D local features and forms with misalignments of edges. The vertical bars represent standard deviation values.
Results and Discussion

The results are displayed in Figure 5 for each subject. For a same amount of contour present, identification thresholds were significantly lower for forms containing 3D features (contrast = 14.84%) than for forms with midsegments (contrast = 15.44%), F(1, 5) = 17.2, p < 0.009, and F(1, 14) = 11.3, p < 0.004. This result shows that three-dimensional cues facilitate object identification. This could be due either to an earlier detection of the 3D features, because of a larger number of pixels present in 3D features relative to the line segments, or to objects being stored in terms of volumetric entities (Biederman, 1987; Biederman & Gerhardtstein, 1993). We return to this point in the General Discussion. On average, contrast thresholds in Experiment 3 were higher than those observed for figures with maxima (12.95%) and pictures with midsegments (13.1%) in Experiment 2. In addition to possible differences in visual acuity, this result can be explained by the fact that the subjects in Experiment 3 were naive to psychophysical experiments. A high between-stimuli variability was observed in the first session of threshold measurements, though data were more stable in the second session (the standard deviation between pictures was ±0.96% for 3D forms and ±1.12% for forms with midsegments in the second session). For this reason, only the values from the second session of threshold measurement were kept for the priming experiment.

EXPERIMENT 4

Method

Procedure. The procedure was the same as that described in Experiment 2. There were three priming conditions: (1) a neutral (no-prime) condition, where nothing was displayed before the mask, (2) forms composed of midsegments of edges, and (3) forms with 3D local features. The primes were presented for 64 msec. The same temporal interval was used for the no-prime condition. After a delay of 64 msec, a mask (a matrix of black and white dots) was presented for 64 msec. It was followed, 64 msec later, by a target that was the complete outline drawing of the object from which the prime was derived. The target was presented for 64 msec. Subjects were required to name the target picture as quickly as possible. Naming times were recorded by a voice key. Three levels of contrast were used for primes. They were adjusted to each subject's threshold obtained in Experiment 3. One level of contrast (L1) was below identification threshold for all pictures. The second value (L2) corresponded to the identification threshold of pictures with 3D features. The third level of contrast (L3) corresponded to threshold for forms with midsegments. This level of contrast was above threshold for pictures with 3D features. As in Experiment 2, the levels of contrast were averaged over the 15 pictures of each version of incomplete forms.
The 15 pictures of each version of incomplete forms were presented once with the target located on the right and once with the target located on the left (0.5° from fixation). The same procedure was used for the no-prime condition. There were thus 90 trials for each level of contrast, yielding a total of 270 trials. The three priming conditions, the spatial location of the target, and the three levels of contrast were randomly and equally represented.

Results

RTs lower than 150 msec and longer than 1500 msec were discarded (1.52% of the data were eliminated this way). The mean RT was 791 msec. There was no effect of the spatial location of the target (91 msec for both left and right targets). The data, averaged over the spatial location of the target, are presented in Figure 6.

RTs did not vary significantly with the level of contrast (L1, 795 msec; L2, 784 msec; L3, 754 msec, both F's < 1). Performance was affected by the type of prime (no prime, 806 msec; forms with 3D features, 775 msec; forms with midsegments, 792 msec) F(2, 10) = 3.62, p < 0.065, and F(2, 28) = 3.9, p < 0.007. The interaction between contrast and the type of prime was not statistically significant, F(4, 20) = 2.1, n.s., and F(4, 56) = 1.6, n.s., but Figure 6 suggests that performance for the three types of prime differed with contrast.

Relative to the no-prime condition, no priming effect was observed at the level of contrast that was below threshold for the two types of prime (L1: no prime, 794 msec; 3D forms, 793 msec; midsegments, 797 msec). A significant

![Mean response times (in msec) averaged over the spatial location of the target and the different pictures as a function of the type of prime and the level of luminance in Experiment 4. The vertical bars represent standard deviation values.](image)
priming effect was found for forms with 3D features at the level of contrast corresponding to threshold for these forms. RTs for forms with 3D features were 67 msec faster than RTs for the no-prime condition, \( F(1, 5) = 10.3, p < 0.024 \), and \( F_{23} (1, 14) = 9, p < 0.01 \). There was a slight facilitation for forms with midsegments (17 msec, both Fs < 1.6) at this level of contrast. The magnitude of the facilitation for forms with midsegments increased at the level of contrast corresponding to threshold for these forms (27 msec), but the effect was not statistically significant, \( F_{1}(1, 5) = 2.36, \) n.s., and \( F_{23} (1, 14) = 2.7, \) n.s. Figure 6 shows that, on the contrary, RTs increased for forms with 3D features at the level of contrast that was above the threshold for these forms (L2); though a slight facilitation effect remained relative to the neutral condition (24 msec), the priming effect was not statistically significant, \( F_{1}(1, 5) = 1.8, \) n.s., and \( F_{23} (1, 14) = 2.69, \) n.s.

The data from Experiments 2 and 4 were contrasted, taking as factors the three priming conditions (no prime, forms with maxima (or 3D features), and forms with midsegments) and the three levels of contrast: (1) below threshold: L1 for Experiments 2 and 4; (2) at threshold: L4 for the two types of forms in Experiment 2; (3) L2 for forms with 3D features and L3 for forms with midsegments in Experiment 4.

On average, subjects were slower in Experiment 4 than in Experiment 2 (by 107 msec), \( F_{1}(1, 10) = 9.8, p < 0.01 \); and \( F_{23} (1, 28) = 61, p < 0.001 \). This increase in RTs affected the three priming conditions equally—no prime. + 108 msec, \( F_{1}(1, 10) = 9.11, p < 0.003 \); and \( F_{23} (1, 28) = 41.6, p < 0.001 \); forms with 3D features. + 102 msec, \( F_{1}(1, 10) = 8.88, p < 0.01 \); and \( F_{23} (1, 28) = 57.6, p < 0.001 \); forms with midsegments. + 112 msec, \( F_{1}(1, 10) = 9.3, p < 0.012 \); and \( F_{23} (1, 28) = 54.3, p < 0.001 \). The main effect of priming condition was significant, \( F_{1}(1, 10) = 5.9, p < 0.01 \); and \( F_{23} (1, 28) = 16.65, p < 0.005 \), but there was no significant interaction between experiment and priming condition.

We compared the magnitude of priming for forms with flat corners in Experiment 2 and forms with 3D features in Experiment 4 and also for forms: with midsegments of edges in the two experiments at levels of contrast corresponding to identification threshold for each type of picture. The magnitude of facilitation was larger for forms with 3D features (67 msec: 8.25%) than the forms with flat corners (41 msec: 5.8%). The difference was significant for pictures as random variable, \( F_{1}(1, 10) = 2, \) n.s., and \( F_{23} (1, 28) = 4.76, p < 0.038 \). Priming for forms with midsegments of edges was more pronounced in Experiment 2 than in Experiment 4 (46 msec, 6.5% vs. 27 msec, 3%), but the difference was not statistically significant \( F_{1} \) and \( F_{2} < 1.5. \)

The percent of priming was obtained by subtracting the mean RT for the neutral condition and the neutral RT for each type of form, and by dividing the results by the mean RT for the neutral condition.
Discussin

The results show that priming was larger for forms containing 3D local features on their outline contour than for flat forms whose contour was made of midsegm-
ents of edges. Though facilitation increased with the increase in contrast for forms with midsegments, the priming effect was not significant for these stimuli. Also, the comparison of Experiment 2 and Experiment 4 shows that facilitation tended to be more pronounced for forms with 3D features than for forms with flat corners (maxima in Experiment 2) at the level of contrast corresponding to threshold for these stimuli. The advantage for forms with 3D features can be explained either by the fact that 3D junctions facilitate grouping of local struc-
tures (Humphreys, Keules, & Connelly, 1994) and that increase the discriminability of the object's parts and their spatial arrangement, or because objects are stored in memory in the form of volumetric entities, or to both processes. We return to this point in the General Discussion.

Paradoxically, priming decreased with the increase in contrast for forms with 3D features. Examination of individual data showed that this occurred for three subjects, whereas priming increased for three other subjects. A possible expla-
nation for this result could be that, as the contrast increases, small details conveyed by high spatial frequencies become more visible. As some details were common or very similar in different objects (e.g. the wheels in the car, lorry, and train), attention to these local parts due to their higher visibility might have ac-
tracted the representation of an object that did not correspond to the target. In this case the increase in RTs at the highest level of contrast might reflect inhibition arising from wrong object representation being activated by the prime stimulus.

GENERAL DISCUSSION

We have reported four experiments designed to investigate the contribution of global shape information and different types of local features in object identifi-
cation and picture priming. The main results can be summarized as follows: (1) In Experiment 1, contrast thresholds for identification were affected both by the energy content of the stimulus and by their degree of local structure. Lower thresholds were observed for pictures having a large amount of contour present, and, among the latter, thresholds were lower for pictures with collinear elements relative to pictures with misaligned elements. (2) In Experiment 3, identification thresholds were systematically lower for pictures containing 3D local features than for pictures whose outline contour was composed of midsegments of edges. (3) In Experiment 2, facilitation in the naming of a target outline drawing of an object was larger when it was preceded by an incomplete version of the object with aligned elements than when the prime was the same form with non-aligned elements. (4) Priming was equivalent for forms whose contour contained only maxima and for those whose contour contained only midsegments of edges at
levels of contrast corresponding to identification threshold for these forms. (5)
In Experiment 4, forms composed of local features providing information about three-dimensionality led to a larger priming effect than did forms with midseg-
ments of edges and forms composed of flat corners (maxima in Experiment 2) at levels of contrast corresponding to identification threshold for these stimuli.

The effect of the amount of surface present on identification thresholds is consistent with psychophysical data showing that, at sensory levels of processing, the visual system sums luminance over space (see Hood & Finkelstein, 1986, and Oltz & Thomas, 1984, for reviews). Boucart and Bonnet (1990) found that absolute detection thresholds were the same for incomplete forms with collinear elements and for forms with non-collinear elements derived from those with aligned elements. This result indicates that equivalence in the energy content and in global shape information gives rise to equivalent detection thresholds. The present experiments show, however, that the same does not hold for identification thresholds. The effect of collinearity on identification threshold in Experiment 1 shows that rough global shape information (based on low spatial frequencies), which was very similar for the two types of forms, is not sufficient for object recognition (at least with different objects having very similar global shapes). More information is needed about shape boundaries and parts of objects in order to distinguish the different objects. Misalignment impairs the computation of external and internal boundaries (Boucart & Humphreys, 1992a, 1992b; Boucart et al., 1994). In these circumstances, subjects are constrained to rely upon global shape information, and the discrimi-
nation of objects similar in their global shape is more difficult than when more detailed information is available.

Another effect of the degree of structure on object identification appears in the lower identification thresholds for pictures containing 3D local features than for pictures made of midsegments of edges (in Experiment 3). Given that the two types of forms had the same amount of contour present, this effect cannot be attributed to a difference in the energy content of the stimuli, though it is possible that the features themselves (3D corners vs. line segments) did differ in their energy content. In this case, an earlier availability of local characteristics would provide an advantage for later processes. Nevertheless, 3D junctions provide more information than do line segments about the spatial arrangement of the object's component parts, and this could also explain the lower identifi-
cation thresholds obtained for these stimuli. Evidence for the efficiency of 3D features in the spatial organization of local and global configurations was shown in a recent study by Humphreys et al. (1994). These authors compared perfor-
mance in the detection of a target (a corner) in incomplete pictures of cubes made of either 3D corner junctions or flat 2D corners in a visual search para-
digm. They found that performance was affected by the number of distractors present in the display when distractors configured into 2D shapes, whereas RTs did not vary with the number of distractors when the non-target corner junctions
configured into cubes. Moreover, performance with 3D cubes was more robust to misalignments between corner junctions than that with flat 2D forms. These results suggest that 3D features are detected early and constitute an important physical characteristic for perceptual organization of forms (see also Enns & Remsniak, 1991).

Consistent with the threshold data, the larger priming effect for forms with collinear elements than for forms with non-collinear elements suggests that rough global shape information does not efficiently activate stored object representations, because this type of information activates too many candidates. These two types of stimuli differed mainly in terms of the visibility of the component parts of objects, which were more discriminable with aligned edges forming external and internal boundaries. The larger amount of information about the object in collinear stimuli would reduce the number of candidates activated in stored representations and therefore produce a larger priming effect. The larger facilitation for forms with collinear elements is inconsistent with an account of object priming based on the activation of a rough global representation used to delimit a class of candidate objects and guide the processing of more local information (Baker-Cave & Kosslyn, 1993; Sanocki, 1993). Our results suggest, rather, that priming is all the greater when more local information about the object’s structure is available, because the number of candidates activated is then smaller. Global shape information might be efficient when the different objects have distinct global shapes, such as the houses vs. vehicles used in Sanocki’s (1993) study. This was not the case in the present study.

The fact that priming effect increase as a function of the increase in contrast supports a part-based account of object recognition (Biederman, 1987; Biederman & Cooper, 1991), because local parts become more and more visible with increases in contrast. The advantage for forms with collinear elements is also consistent with Srivinas (1993) proposal that the magnitude of priming depends on the physical similarity between prime and target. Object component parts were as discriminable in forms with collinear elements as they were in complete outline drawings. Srivinas (1993) found that priming was larger when the pictures presented at study and test were physically identical, relative to physically similar forms, regardless of whether or not studied items afforded perceptual closure and recovery of parts objects. Our results suggest that recovery of parts through collinearity of local contour elements was an important factor for efficient activation of object representations.

Attneave and Arnould (1956; see also Zasne, 1970) stressed the importance of maxima on the outline contour of objects as informative regions for object recognition. Experiment 2 showed that priming was equivalent (at contrast values corresponding to identification threshold and above threshold) for forms whose contour was composed of maxima and for forms in which the maxima were deleted. A larger priming effect appeared for forms with maxima when the two-dimensional corners were replaced by three-dimensional corners (in
Experiment 4). The larger priming effect for 3D local features relative to line segments and flat corners* can occur at two different levels of processing. At sensory levels of processing, 3D features might be detected earlier than line segments because of a higher energy content. The number of pixels was larger in these features than in line segments and flat corners, and energy content affects detection and identification thresholds (Oltzak & Thomas, 1985; Thomas, 1985). An earlier availability of 3D local features relative to the two other types of features would provide them with a temporal advantage for perceptual organization of the form and access to stored object representations. At levels of processing involving object representations, objects might be stored in the form of three-dimensional entities. This type of description is economic for the rapid identification of those objects for which there is constant orientation change as the observer moves in natural environments. Evidence for this type of description is suggested by studies showing that priming is not affected by the viewing angle of objects as long as all the component parts remain visible (Biederman & Gerhardtstein, 1993).

In conclusion, our study shows that local features can be more efficient than global shape information in picture priming. It also shows that, among these local features, physical characteristics providing information about the three-dimensionality of an object particularly facilitate its subsequent identification. Our results do not, however, allow us to distinguish between an advantage for forms containing 3D features due to the activation of stored three-dimensional object representations or due to an earlier availability of 3D features at sensory levels of processing. This point requires further investigation.

REFERENCES


*For all RTs were longer for 3D forms in Experiment 4 than for flat corners in Experiment 2, but one difference in RTs between the two experiments affected the three priming conditions equally.

The increased priming in Experiment 4 is due to the lower stimulus in Experiment 4 rather than to the physical features in themselves.


Manuscript received 1 November 1993
Revised manuscript received 1 December 1994