VISUAL COGNITION, 2001, *8* (6), 801–821

Is global shape sufficient for automatic object identification?

Roberto Dell'Acqua and Remo Job
University of Padova, Italy
Jonathan Grainger
CNRS & Université de Provence, France

The present paper reports two experiments that investigate the critical features of an object shape that automatically elicit recognition. Silhouettes of real objects (targets) and meaningless patterns (fillers) in both canonical and non-canonical formats were presented to subjects, in an attempt to test whether information about the global shape of an object was sufficient for automatic object identification. In Experiment 1, target–filler discriminability was evaluated by means of a reality-decision task. In Experiment 2, subjects had to perform an elongation-decision task, previously shown to be sensitive to the influence of automatically activated object identities (Dell'Acqua & Job, 1998). Contrary to the previous findings, the present study shows that, although silhouettes were identified with surprising good accuracy in the reality-decision task, effects of object identity on the elongation-decision task were negligible.

Figure–ground organization entails segmenting an attended portion of the visual field into regions that are objects and regions that are backgrounds. Contrary to backgrounds, objects seem to possess a specifiable shape and appear to occlude the backgrounds. Since the 1980s, the majority of the models proposed to describe the organization of the mental processes involved in object recognition have relied on the notion that objects must be separated from backgrounds before any operations related to matching objects to their stored mental representations can take place (Biederman, 1987; Marr, 1982; Ullman, 1996; but see

Please address all correspondence to R. Dell'Acqua, Dipartimenti di Psicologia, Via Venezia, 8, 35131 Padova, Italy. Email: dellacqu@psy.unipd.it

This work was supported by a STARTUP Grant (FFMA 1998) from the University of Padova to all authors, and by a grant from COFIN MURST (1997) to Remo Job. Roberto Dell'Acqua is currently supported by a postdoctoral grant from the Italian Ministry of Scientific Research. The authors are indebted to Muriel Boucart, Glyn Humphreys, and Pierre Jolicoeur for useful comments on a previous draft of this manuscript.

http://www.tandf.co.uk/journals/pp/13506285.html DOI:10.1080/13506280042000153

Lowe, 1985, and Peterson & Gibson, 1994). Ancillary to this assumption was the hypothesis that the outer boundary of an object, namely information about its global shape, must play a key role in the process of object recognition (see Quinlan, 1991, for a discussion on this point). Intuition suggests that global shape, in particular conditions, may be solely responsible for recognition. Ordinarily, indeed, people in normal environments do not seem to have trouble in recognizing objects on the basis of global shape information. Entering a moderately lit room does not completely disrupt our ability to reach for an object, especially when the situation implies an active search for that object. What is controversial in linking global shape information to the process of recognition is whether recognition acts as a non-controllable, ballistic process once global shape information is processed by the visual system. The debate surrounding this point has produced several empirical investigations, whose results are apparently in conflict.

On the one hand, the notion that object global shape is a primitive in visual cognition has received empirical support from studies that have shown that objects are automatically identified in conditions in which either attention is directed to the global shape, or stimulus information is impoverished to such an extent that only global shape information can be recovered from an object. One of the most convincing demonstrations on this line has been provided by Boucart and Humphreys (1992), and Boucart, Humphreys, and Lorenceau (1995). In these studies, the authors employed a form-comparison and a colour-comparison task. Subjects were presented with a reference object, followed by the presentation of a pair of objects. A speeded response was required to indicate which of the pair of objects matched the reference according to either global shape features, e.g., spatial orientation and size of the objects (Boucart & Humphreys, 1992, 1994), or the relative density distribution of two colors on the object outer boundary (Boucart et al., 1995). In both paradigms, subjects were explicitly instructed to neglect the identity of the objects, and focus on the task-relevant perceptual dimension. The results showed consistently faster matching responses when the target and the reference were semantically related compared to the condition in which the stimuli were semantically unrelated. These results were not found when subjects were instructed to make their responses according to other perceptual dimensions (e.g., motion or texture colour). Such findings were taken as evidence that the processing of global shape information automatically induced the activation of stored object representations.

On the other hand, experimental studies that have employed visual search paradigms with objects as stimuli have come to rather different conclusions. Wolfe and Bennet (1997), for instance, devised an elegant series of experiments, in which subjects were presented with multiple displays of objects, with each object composed of a fixed number of local attributes (different types of curves and vertices). Target objects shared a certain number of these local attributes with distractors, the only difference between these two categories of stimuli being their global shape. The authors reported seven experiments in which the similarity of the local attributes shared by targets and distractors, the level of search practice, and the familiarity of the shapes of their stimuli were systematically manipulated. It was hypothesized that automatic processing of object global shape would have lead to close-to-flat visual search slopes when the number of distractors was varied. However, in no case did the results conform to this prediction. The rather steep visual search slopes obtained across the seven experiments were taken as evidence that global shape was not processed automatically.

THE PRESENT EXPERIMENTS

In order to provide an empirical investigation on the issue of whether global shape processing leads to automatic activation of the identity of the objects, we adopted a close variant of the paradigm originally proposed by Dell'Acqua and Job (1998), in which subjects were instructed to perform a perceptual task on objects that did not require their identification. In Dell'Acqua and Job's study, subjects were instructed to indicate whether the main axis of elongation of an object was vertical or horizontal. Line drawings of real-world objects were intermixed with non-objects, and copies of these stimuli were submitted to a graphical manipulation so as to reverse the direction of their original elongation. The results revealed a selective cost in responding to the elongation of the real-world objects that were graphically distorted. The effect of the distortion was absent when subjects responded to non-objects. This finding suggested that performance in the elongation-decision task was affected by the automatic activation of the canonical representation of the real-world objects. When the actual elongation on the screen and the 'semantic' elongation of the real-world objects mismatched, competition between opposite response codes resulted in longer reaction times. In the present study, Dell'Acqua and Job's line drawings were replaced with the silhouettes of the same objects, in the attempt to evaluate whether the same pattern of results would be observed under conditions in which the to-be-processed information was limited to the global shape of the objects. In Experiment 1, subjects were required to perform a reality decision test on the silhouettes. Subjects had to categorize each silhouette as that of a real-world object or a meaningless pattern. The aim of Experiment 1 was to provide a direct estimate of the speed and accuracy in accessing stored information about the objects selected for the present experiments. The same paradigm as that devised by Dell'Acqua and Job was used in Experiment 2, in order to test whether global shape is sufficient information to elicit automatic object identification.

EXPERIMENT 1

Method

Subjects. The subjects were 14 undergraduate students (6 females) at the University of Padova who participated for course credit. The age of the subjects ranged from 23 to 29 years. All were naive to the purpose of the experiment and all had normal or corrected-to-normal vision.

Material. The stimuli consisted of the silhouettes of 32 real objects (half animals and half artefacts; hereafter, targets) selected from the Snodgrass and Vanderwart (1980) set of stimuli and from the British Picture Vocabulary Scale. Half of the targets had a vertical axis of canonical elongation (vertical targets; e.g., candle), and half had an horizontal axis of canonical elongation (horizontal targets; e.g., shoe). Vertical and horizontal targets were matched for familiarity (see Dell'Acqua & Job, 1998, for the familiarity values, and for the complete list of stimuli). A corresponding set of 32 fillers was generated by changing the orientation of portions of the outermost boundary of the targets, so that each filler closely matched the corresponding target in the extent of the vertical and horizontal axes of elongation. To create the distorted stimuli, targets and fillers were submitted to an additional graphical manipulation. The original (canonical) format of the stimuli was modified such that the canonical elongation of the stimuli was reversed. A set of 64 distorted stimuli was generated by shortening the vertical axis and by lengthening the horizontal axis of the vertical stimuli, and by applying the complementary manipulation to the horizontal stimuli. As a result of this manipulation, distorted stimuli with a horizontal elongation were obtained from canonical stimuli with a vertical elongation, and distorted stimuli with a vertical elongation were obtained from canonical stimuli with a horizontal elongation (see Figure 1, for an example).

Procedure. The stimuli were displayed on a 15" NEC screen (cathode ray tube, resolution 640 × 480 pixels) controlled by a Pentium 166 MHz CPU and MEL software. At a viewing distance of 90 cm, each stimulus fit in a square portion of the screen with a side of less than 6° of visual angle. The vertical stimuli ranged from 3.3° to 5.3°, and the horizontal stimuli ranged from 3.5° to 5.6°. The stimuli were presented in black (RGB coordinates: 0, 0, 0) on the light gray background of the screen (RGB coordinates: 90, 90, 90).

On each trial, a fixation point (a black dot of 0.20°) was displayed in the centre of the screen for 500 ms. At the offset of the fixation point, a stimulus was displayed for 150 ms. Following the stimulus, a masking pattern was exposed for 100 ms. The masking pattern was composed of a pseudo-random combination of parts taken from different silhouettes. No silhouette could be identified when the mask and the silhouette overlapped on the screen. Subjects were instructed to decide as quickly and accurately as possible whether each

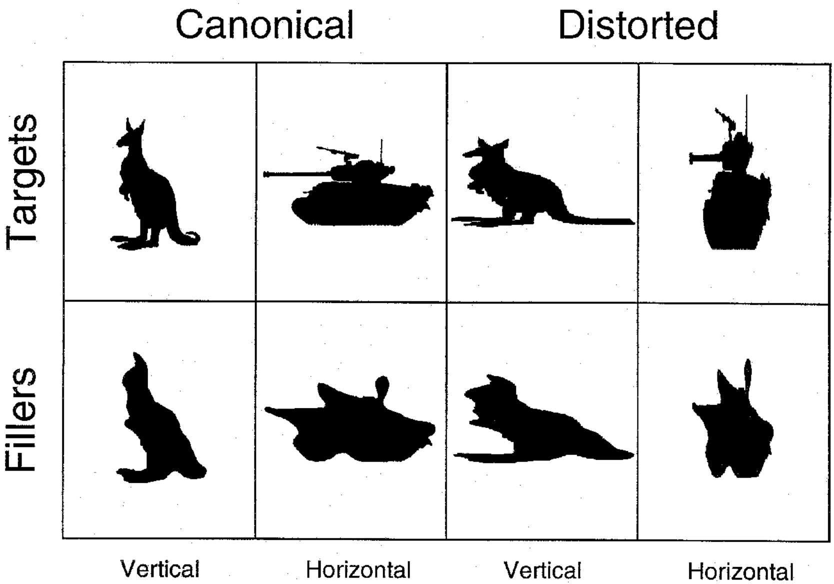

Figure 1. Examples of targets (upper four frames) and fillers (lower four frames), in both canonical and distorted formats.

stimulus was the silhouette of a real object or a meaningless pattern by pressing one of two keys of a response box. Response times (RTs) were measured from stimulus onset until the subject's manual response. After the execution of the manual response, an intertrial interval of 2 s elapsed before the presentation of the fixation point for the following trial. The total set of 128 stimuli was presented in two different blocks of 64 stimuli each, with a short rest between the blocks. For each block, there were four different orders of randomization. The trials were randomly ordered within the constraints of having no more than four repetitions of the same type of stimulus (target vs. filler), or format (canonical vs. distorted), on successive trials. Order of block presentation and hand-response mappings were fully counterbalanced across subjects. Before the beginning of the experimental session, each subject performed 32 practice trials with stimuli that were not presented during the experimental session. The experiment was about 30 min long.

Two factors were independently manipulated within our experimental design. The first factor was "type of stimulus", with two levels, meaningful objects (i.e., targets) vs. meaningless patterns (i.e., fillers). The second factor was "format" of the stimuli, with two levels, "canonical" vs. "distorted" stimuli, that were defined according to whether or not the silhouettes were graphically distorted.

Results

The analyses concentrated on correct RTs and proportion of hits. Correct RTs were first screened for outliers using a procedure based on the computation of the cut-off bounds of each subject's performance. Cut-off bounds were computed by adding and subtracting two standard deviations from each subject's overall mean RT. Each RT not included in the range delimited by the cut-off bounds (less than 1.5%) was replaced by the cut-off bound value. Correct RTs and proportion of hits were then submitted to ANOVA, with both subjects ($F1$) and items ($F2$) as random factors. The factors considered in the ANOVA were type of stimulus (target vs. filler), and format (canonical vs. distorted). All factors were treated as within in the analyses with subjects and items as random factors. The significance level chosen was $p < .05$.

Mean RT analysis. The overall mean RT was 531 ms. A summary of the results is reported in Figure 2a. The main effect of type of stimulus was significant, $F1(1, 13) = 104.1$, $MSe = 668$, $p < .001$ and $F2(1, 31) = 23.7$, $MSe = 5943$, $p < .001$, indicating shorter RTs to targets than to fillers. The main effect of format was significant, $F1(1, 13) = 5.8$, $MSe = 465$, $p < .04$ and $F2(1, 31) = 10.8$, $MSe = 2067$, $p < .003$, indicating shorter RTs to canonical stimuli than to distorted stimuli. The interaction between type of stimulus and format was significant, $F1(1, 13) = 6.9$, $MSe = 263$, $p < .03$ and $F2(1, 31) = 12.4$, $MSe = 1607$, $p <$

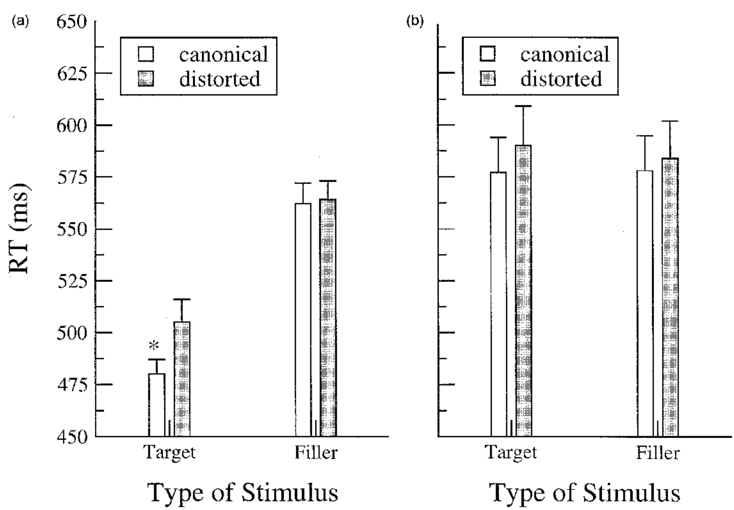

Figure 2. (a) Mean RTs and 95% confidence intervals (Loftus & Masson, 1994) in Experiment 1 (reality-decision task), as a function of type of stimulus and format. The asterisk highlights the crucial significant difference between RT to targets in the canonical format and RT to targets in the distorted format. (b) Mean RTs and 95% confidence intervals in Experiment 2 (elongation-decision task), as a function of type of stimulus and format.

.002. As Figure 2 suggests, the interaction was due to the fact that the effect of format was evident on targets but not on fillers. To support this hypothesis, separate analyses were conducted on targets and fillers data. The effect of format was significant for targets, $F1(1, 13) = 17.3$, $MSe = 257$, $p < .002$ and $F2(1, 31) = 16.0$, $MSe = 2641$, $p < .001$, and was not significant for fillers (all Fs < 1).

Proportion of hits analysis. The mean proportion of hits was .87. The main effect of type of stimulus was not significant (all Fs < 1), indicating an equal proportion of hits for targets and for fillers. The main effect of format was significant, $F1(1, 13) = 20.8$, $MSe = 0.003$, $p < .001$ and $F2(1, 31) = 4.7$, $MSe = 0.026$, $p < .04$, indicating a lower proportion of hits for distorted stimuli than for canonical stimuli (.84 vs. .90, respectively). Type of stimulus interacted with format, $F1(1, 13) = 40.2$, $MSe = 0.001$, $p < .001$ and $F2(1, 31) = 11.4$, $MSe = 0.029$, $p < .002$. The proportion of hits for canonical fillers and distorted fillers was constant (.86 vs. .88, respectively), whereas the proportion of hits was higher for canonical targets than for distorted targets (.96 vs. .80, respectively).

Discussion

The results of Experiment 1 are important in several respects. First, subjects were able to distinguish the silhouettes of real-world objects from the corresponding meaningless counterparts with a surprising high level of accuracy. We take this finding as providing a direct demonstration that the silhouettes of the real-world objects selected for the present study conveyed sufficient information for their identification.

Furthermore, the results provide an absolute temporal estimate of the accessibility of information about the structural representation of the real-world objects (e.g., Humphreys, Lamote, & Lloyd-Jones, 1995; Kroll & Potter, 1984; Lupker, 1988). On average, subjects were almost 60 ms faster in making the correct decision that a silhouette was that of a real-world object than the opposite decision. This suggests that the reality decision task was performed via a comparison that was self-terminating on a match between a real object and its long-term representation, and exhaustive otherwise (Ratcliff, 1978). Alternatively, this result could reflect the influence of a response bias generated by treating one response dimension (i.e., “real object”) as a positive dimension, and the opposite response dimension as its negation (e.g., by treating “meaningless pattern” to the same extent as “non-real object”; e.g., Pike & Ryder, 1973). In either case, however, the results of Experiment 1 provide the indication that, by the time subjects correctly judged a silhouette as that of a meaningless pattern (on average, after 563 ms), any matching operations of the silhouettes of the real-world objects to their stored mental representations were reasonably terminated.

A further central aspect that should be considered is that the results of Experiment 1 replicate Dell'Acqua and Job's (1998) findings using line drawings. In both their study and in the present experiment, performance was disrupted when subjects judged the distorted real-world objects. We argue that this pattern of results may be informative about the type of information processed during the present reality-decision task. On the hypothesis, for instance, that the reality of the silhouettes covaried with some physical characteristic of the stimuli (e.g., different degrees of contour curvature across the stimuli falling in the two response categories), and that subjects categorized the stimuli exclusively on the basis of such physical characteristics, an effect of the distortion confined to real objects would be difficult to explain. Such an effect could only be accounted for by assuming that the graphical distortion had differential effects on the global physical appearance of the real objects and the meaningless patterns. Instead, it is more likely that the distortion effect is due to the fact that subjects were actually relying on their visual knowledge about objects in categorizing the present stimuli. In this view, we assume that the reaction times to the distorted real objects were inflated by the time taken to carry out an additional processing step, possibly involving normalization of the distorted stimuli, before a correct "real" response was emitted.

EXPERIMENT 2

The similarity in the results of Experiment 1 and the results obtained by Dell'Acqua and Job (1998) constitutes converging evidence that the activation of information about canonical elongation from stored representations of objects, whether explicitly required for the present reality decision task or automatically elicited by virtue of the presentation of line drawings in Dell'Acqua and Job's experiment, had a specific influence on performance. On the hypothesis that global shape information is sufficient to automatically elicit object identification, we predict that the distortion effect observed in Experiment 1 should also be observed under conditions in which the task does not rely on object identification. This hypothesis was tested in Experiment 2.

Method

Subjects. Subjects were 28 undergraduate students (12 females) at the University of Padova who participated for course credit. The age of the subjects ranged from 24 to 29 years. All were naive to the purpose of the experiment and all had normal or corrected-to-normal vision. None of the subjects had participated in Experiment 1.

Material. The material consisted of the stimuli used in Experiment 1.

Procedure. The stimuli were displayed on a 14" Macintosh RGB screen (cathode ray tube, resolution 640 × 480 pixels) controlled by an Apple computer (Macintosh Quadra 700). The dimensions of the stimuli, the general experimental settings, and the temporal sequence of events on each trial of the present experiment were the same as those adopted in Experiment 1. Subjects were instructed to decide as quickly and accurately as possible whether the elongation of each stimulus was either more vertical or more horizontal by pressing one of two keys of the computer keyboard. The instructions stressed the importance of judging the elongation of the stimuli regardless of their identity, which was defined as irrelevant to the aim of the experiment. Response times (RTs) were measured from stimulus onset until the subject's manual response. After the execution of the manual response, an intertrial interval of 2 sec. elapsed before the presentation of the fixation point for the following trial.

Three factors were independently manipulated within our experimental design. The first factor was type of stimulus, with two levels, meaningful objects (i.e., targets) vs. meaningless patterns (i.e., fillers). The second factor was elongation of the stimuli, with two levels, "vertical" stimuli (e.g., a candle and its associate filler) vs. "horizontal" stimuli (e.g., a shoe and its associate filler). The third factor was format of the stimuli, with two levels, "canonical" vs. "distorted" stimuli, that were defined by the absence vs. presence of the graphical manipulation that was applied to half of the silhouettes[1].

Results

The analyses concentrated on correct RTs and error rates. Correct RTs were first screened for outliers using the same procedure as that used in Experiment 1. Less than 2% of the RTs were replaced by the relative cut-off bound values. Correct RTs and error rates were then submitted to a series of ANOVAs, with both subjects ($F1$) and items ($F2$) as random factors. The factors considered in all ANOVAs were type of stimulus (target vs. filler), elongation (vertical vs. horizontal), and format (canonical vs. distorted). All factors were treated as within in the analysis with subjects as random factors. Elongation was treated as between in the analysis with items as random factor.

Mean RT analysis. The overall mean RT was 584 ms. The main effect of elongation was significant in the by-subjects analysis only, $F1(1,27) = 4.2$, MSe

[1]In Experiment 2, the two levels of the factor elongation always referred to the original elongation of the stimuli (i.e., to the elongation of the stimuli before the graphical manipulation), and not to the actual elongation of the stimuli when displayed on the screen. To make it clear, while "vertical" and "horizontal" stimuli in the "canonical" condition have, respectively, a vertical and horizontal elongation when displayed on the screen, in the "distorted" condition, "vertical" and "horiziontal" stimuli actually have a horizontal and a vertical elongation on the screen, respectively (see Figure 1).

= 841, $p < .05$, indicating faster responses to vertical stimuli (578 ms) than to horizontal stimuli (586 ms). The effect of elongation was qualified by the significant interaction between elongation and format, $F1(1, 27) = 27.5$, $MSe = 3759$, $p < .001$ and $F2(1, 14) = 8.0$, $MSe = 4542$, $p < .02$. Vertical stimuli were responded to faster in the canonical condition (i.e., when they were vertically elongated) than in the distorted condition (i.e., when they were horizontally elongated), $F1(1, 27) = 26.2$, $MSe = 2976$, $p < .001$, whereas horizontal stimuli were responded to faster in the distorted condition (i.e., when they were vertically elongated) than in the canonical condition (i.e., when they were horizontally elongated), $F1(1, 27) = 13.4$, $MSe = 2296$, $p < .002$. The main effect of Format was marginally significant in the by-subjects analysis only, $F1(1, 27) = 3.6$, $MSe = 1514$, $p < .07$. There was no effect of type of stimulus, nor of the interaction between type of stimulus and format (Fs < 1; see Figure 2).

RT distribution analysis. Before discussing our preferred interpretation of the present finding, a few caveats should be mentioned, all related to the fact that we are in the presence of a null effect of identity activation on the elongation-decision task. As all null effects, the results invite caution. Our effort in the forthcoming sections is to substantiate the null effect while showing that objections based on lack of power in the present experimental design are probably unfounded, by reporting the results of several further analyses carried out on the RT data. In this optic, RTs in the elongation decision task were submitted to a distributional analysis, in order to evaluate the possibility that significant effects of the interaction between type of stimulus and format could be confined to late components of the RT distribution (e.g., Balota & Spieler, 1999). This analysis was motivated by the fact that the generation of a "semantic" response code (i.e., contingent on real-world object canonical elongation) and the generation of a perceptual response code (i.e., contingent on the actual elongation of the object on the screen) have been hypothesized to have different time courses. Widespread evidence, in fact, indicates that the generation of a perceptual code precedes that of a "semantic" code (e.g., Navon, 1977; Zeki, 1993). Consequently, we argue that the probability to detect interference between these codes would be maximized by selectively considering the portion of RTs that were long enough for the two response codes to concurrently affect the overt response to a silhouette. A fitting procedure was used that allowed us to decompose the RT distribution of each subject, in each cell of the type of stimulus by format design, into two major components, an early component characterized by a normal (Gaussian) distribution of processing completion times, and a late component characterized by an exponential distribution of processing completion times. The convolution of these two RT components has a distribution called ex-Gaussian distribution, which has been shown to closely fit observed RT distributions (e.g., Heathcote, Popiel, & Mewhort, 1991; Ratcliff, 1978). The ex-Gaussian distribution has three parameters: mu and sigma (the mean and standard deviation of the underlying Gaussian distribution, respectively), and tau (the mean and standard deviation of the underlying exponential distribution). According to the line of argument proposed earlier, inspection of the late (tau) RT components should be informative about whether null effects of identity activation were indeed reliable or spurious.

To carry out the distributional analysis, the data from each two (randomly chosen) subjects, in each cell of the design, were Vincent averaged after the decomposition of each subject's distribution in 10 quantiles, each corresponding to 10% of the available RTs for that subject. Ex-Gaussian functions were then generated with the use of RTSYS software (Heathcote, 1996). The goodness-of-fit of the theoretical ex-Gaussian function to the empirical data was evaluated for each Vincentized distribution, for each cell of the present experimental design, using a chi-square test, with none of the fits rejected at the .05 level of probability (i.e., all ex-Gaussian fits were good). Probability density functions were estimated via SIMPLEX optimization algorithm (Nelder & Mead, 1965). Figure 3 displays the fit of the ex-Gaussian functions (curved lines) across subjects to the empirical correct RTs (bars), as a function of type of stimulus and format.

The fitting procedure allowed a comparison of the RT distributions across the cells of the type of stimulus by format design based on the parameters mu, sigma, and tau. These parameters, averaged across subjects, are reported in Table 1. Independent within-subjects ANOVAs showed null effects ($F < 1$) of the interaction between type of stimulus and format on both mu, sigma, and tau parameters.

Split-half RT analysis. It must be noted that the null effects of identity activation shown in Experiment 2 can simply be viewed as a drastic reduction in the pattern of interference observed in Experiment 1. In this view, the difference between the results of Experiment 1 and Experiment 2 would reflect a *quantitative* difference in the amount of distortion-mediated interference effect, which was substantial in Experiment 1 and close-to-nil in Experiment 2. If this were true, one ought to be able to find statistical conditions that maximize the

TABLE 1
Distributional parameters (in ms) computed following the fit of ex-Gaussian functions to the RT distribution in each cell of the Type of Stimulus by Format design

	Target		*Filler*	
	Canonical	*Distorted*	*Canonical*	*Distorted*
Mu	294	286	299	301
Sigma	57	52	46	58
Tau	132	149	128	134

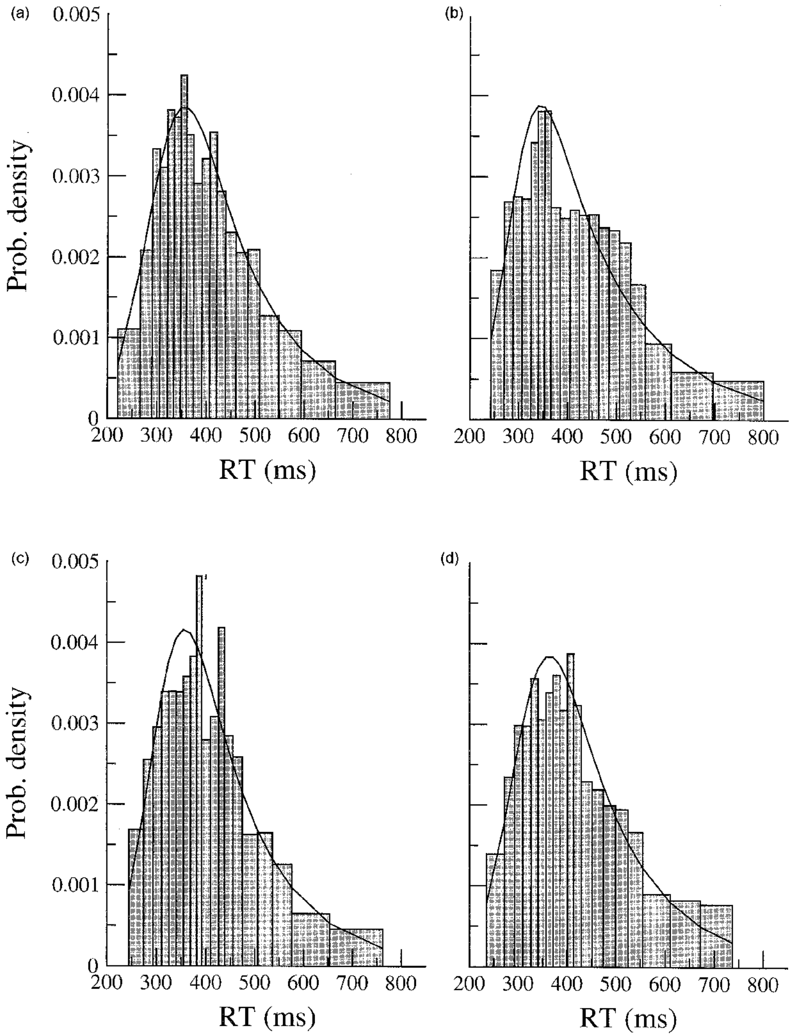

Figure 3. Ex-Gaussian fits (curved lines) plotted against the empirical data (decomposed in 20 quantiles), as a function of type of stimulus and format. (a) Target canonical, (b) target distorted, (c) filler canonical, and (d) filler distorted.

chances to observe a relation between the results of Experiment 1 and Experiment 2, because such a relation should be preserved under the assumption of no *qualitative* differences between the two tasks.

The rationale underlying the analysis was that a significant interaction between type of stimulus and format in Experiment 2 (i.e., the hallmark of automatic identification), if any, could potentially be detected when focusing on the subset of real objects that least contributed to the overall distortion effect observed in Experiment 1, that is, those objects that were identified with similar speed in both their canonical and distorted version. This is because we assume that, for such objects, a "semantic" code was available earlier than the "semantic" code generated following the presentation real objects that, in their distorted version, contributed to a greater extent to the overall distortion effect observed in Experiment 1. To carry out such an analysis, the data from the real objects in Experiment 1 were analysed separately in order to compute the absolute size of the distortion effect associated with each target, by subtracting the mean RT to a given target in the canonical condition from the mean RT to the same target in the distorted condition. These RT differences were submitted to a median-split analysis, on the basis of which two groups of items were individuated: A group of items associated with a below-median contribution to the overall size of the distortion effect, and a group of items associated with above-median contribution to the overall size of the distortion effect. This newly created factor was treated as a by-item factor and included in the design of the analysis carried out on the results of Experiment 2. The results are reported in Figure 4.

There was no apparent modulation of the distortion effect in Experiment 2 as a function of the amount of the distortion effect shown in Experiment 1. A by-item ANOVA corroborated this impression indicating null effects (Fs < 1) of the interaction between type of stimulus and format both for the group of items associated with a greater (i.e., above-median) distortion effect in Experiment 1 and for the group of items associated with a smaller (i.e., below-median) distortion effect in Experiment 1.

Another way to look at these results is to directly correlate the amount of distortion effect in Experiment 1 with the amount of distortion effect in Experiment 2. A scattergram of amount of distortion effect in Experiment 1, plotted against amount of distortion effect in Experiment 2 for each item, is reported in Figure 4c. Also in Figure 4c, we have reported the equation of the linear regression function relating the two measures. A simple correlation analysis, which revealed a non-significant correlation between these two measures, $t(30) = 1.2$, $p > .2$, help to reinforce further the hypothesis of the independence between the results of Experiment 1 and the results of Experiment 2.

Error rate analysis. The overall mean error rate was 8.5%. The main effect of elongation was significant in the by-subjects analysis, $F1(1, 27) = 5.7$, $MSe = .006$, $p < .03$, reflecting a higher error rate for horizontal stimuli (10%) than for canonical stimuli (7%). The effects of format interacted with the elongation of the stimuli, $F1(1, 27) = 37.5$, $MSe = 0.008$, $p < .001$ and $F2(1, 14) = 14.8$, $MSe = 0.006$, $p < .001$. The results indicated a lower error rate for the vertical stimuli in the canonical condition (3%) than in the distorted condition (10.8%), $F1(1, 27) = 27.1$, $MSe = 0.006$, $p < .001$, and a lower error rate for the horizontal stimuli in the distorted condition (6%) than in the canonical condi-

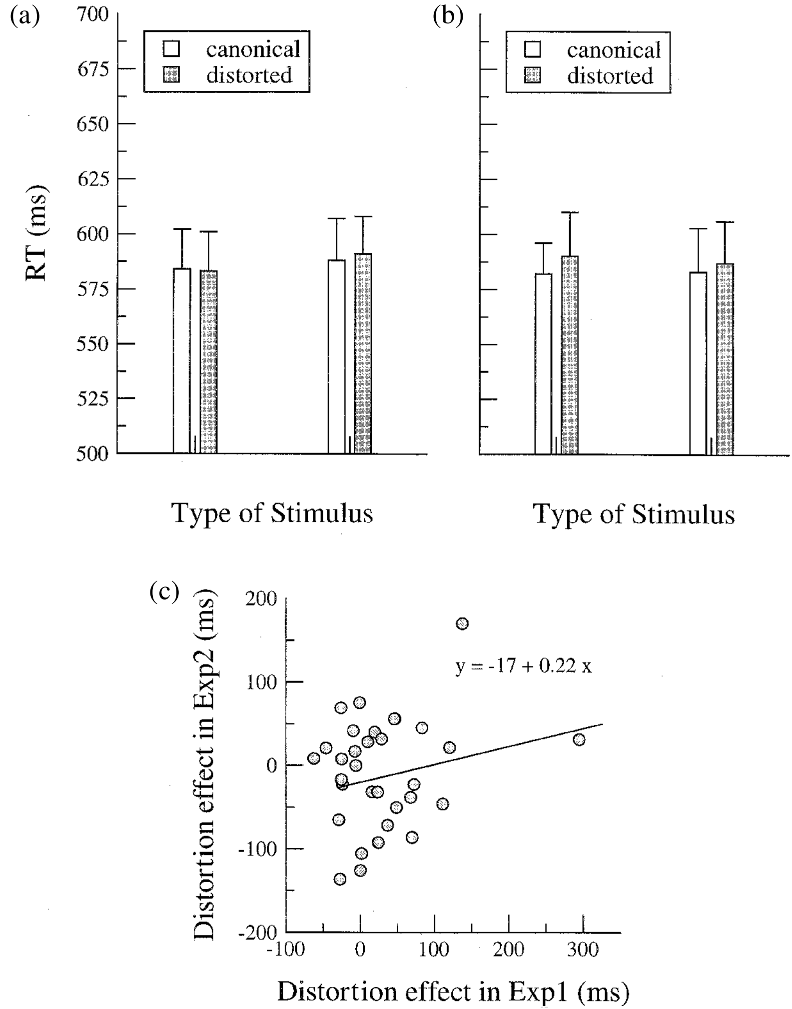

Figure 4. (a) Mean RTs and 95% confidence intervals in Experiment 2 (elongation-decision task) as a function of type of stimulus and format. The graph reports mean RTs to items associated with below-median (smaller) contribution to the distortion effect observed in Experiment 1 (see main text). (b) Mean RTs and 95% confidence intervals in Experiment 2 (elongation-decision task) as a function of type of stimulus and format. The graph reports mean RTs to items associated with above-median (greater) contribution to the distortion effect observed in Experiment 1. (c) Scattergram of the size of distortion effect for individual items in Experiment 1 plotted against the size distortion effect for individual items in Experiment 2. The equation refers to the linear regression function relating the two estimates.

tion (13%), $F1(1, 27) = 27.1$, $MSe = .006$, $p < .001$. Type of stimulus interacted with format in the by-subjects analysis, $F1(1, 27) = 6.8$, $MSe < 0.006$, $p < .02$, reflecting a higher error rate for targets in the distorted condition (10.1%),

$F1(1, 27) = 4.5$, $MSe = 0.006$, $p < .05$, than for targets in the canonical condition, and for fillers in the distorted and canonical condition (7%, 8.1%, 6.9%, respectively). In these latter three cells of the experimental design, error rates did not differ significantly, $F1(1, 27) = 1.9$, $MSe = 0.006$, $p < .15$. No other factor reached the significance level.

Split-half error rate analysis. Although not significant in the by-item analysis carried out on mean error rate in the elongation-decision task, the trend for a higher error rate in responding to the distorted targets is in partial disagreement with the results revealed by the RT analyses. For this reason, error rate was further analysed as a function of the results of Experiment 1, in order to shed light on the possible sources of the significant interaction between type of stimulus and format. Mirroring the argument underpinning the analysis of the RT distribution, one hypothesis might be that the higher error rate for the distorted target in Experiment 2 was generated by a proportion of trials in which distorted targets were actually identified on the part of the subjects involved in the elongation-decision task. Consequently, it might be expected that errors in Experiment 2 were concentrated on those items that were easier to identify (i.e., less prone to identification errors) in Experiment 1. A test for this hypothesis can be provided by conditionalizing the analysis of errors in Experiment 2 on the error rate in identifying the distorted targets in Experiment 1. To this end, the data from distorted targets in Experiment 1 were separately considered, and each distorted target mean error rate was computed. Error rates for distorted targets were submitted to a median-split, following which two groups of distorted targets were generated, one group associated with an above-median error rate, and one group associated with a below-median error rate. This newly created factor was treated as a by-item factor, and included in the design of the analysis carried out on the error rate in Experiment 2. The results of this analysis are reported in Table 2.

There was no evident modulation of the interaction between type of stimulus and format as a function of the accuracy in identifying the distorted targets computed on the basis on the results of Experiment 1. A by-item ANOVA

TABLE 2
Percentages of errors in the elongation decision task (Experiment 2) conditionalized on the error rate (low vs. high) in identifying distorted targets in Experiment 1

	Target		*Filler*	
	Canonical	*Distorted*	*Canonical*	*Distorted*
Low error rate	7.9%	12.4%	12.3%	7.8%
High error rate	6.0%	6.3%	6.2%	7.9%

carried out on error rate showed that the newly created factor did not interact significantly (Fs < 1) with type of stimulus and format.

Discussion

The results of Experiment 2 can be summarized as follows. RTs to vertically elongated stimuli (i.e., vertical canonical and distorted horizontal stimuli) were shorter than RTs to horizontally elongated stimuli (i.e., horizontal canonical and distorted vertical stimuli). This result mirrors a well-established effect in studies of human perception, showing that patterned "verticality" is easier to detect than other spatial perceptual dimensions, such as "horizontality" (e.g., Schwarz & Ischebeck, 1994). This result was also found by Dell'Acqua and Job (1998) in their study using line drawings. Although puzzling, a discussion of this interesting result is beyond the scope of the present work.

Of greatest interest, the RT results of Experiment 2 did not reveal any difference between targets and fillers when the effect of distortion applied to their shapes was considered. This null effect was substantiated by a distributional analysis, and by split-half analyses aimed at evaluating possible dependencies between the results of Experiment 1 and the results of Experiment 2. To summarize, neither early (i.e., mu and sigma) nor late (i.e., tau) components of the ex-Gaussian functions fitted to the observed RT distributions in the cells of the present design revealed an influence of the automatic activation of semantic information on the time to carry out the elongation-decision task. Furthermore, no dependency of the (close-to-nil) amount of distortion effect estimated on the basis of the results of Experiment 2 on the (substantial) amount of distortion effect found in Experiment 1 was observed as a result of a direct correlation carried out on item mean RTs. Taken together, the present results suggest that objects were not automatically identified upon presentation of their silhouettes.

GENERAL DISCUSSION

The results of the present investigation are clear-cut. In Experiment 1, when asked to identify objects on the basis of their silhouettes, responses were strongly modulated by the graphical distortion applied to the stimuli. When subjects responded to the silhouettes of distorted real-world objects, RTs were longer compared to the conditions in which canonical targets, or canonical and distorted fillers, were responded to. In Experiment 2, in contrast, when asked to judge the main axis of elongation of a silhouette, responses were not influenced by object identity.

The finding of null effects of object identity bears a close resemblance with the results obtained by Boucart and Humphreys (1994) in a study in which both line drawings and silhouettes were used as stimuli. The authors compared the performance in two size-matching tasks, one in which two line drawings were to be compared, and one in which two silhouettes were to be compared. Evidence for automatic activation of object identity was found for line drawings (i.e., in the form of better performance with semantically related pairs of stimuli), but not for silhouettes. Boucart and Humphreys hypothesized that whether or not semantic information influenced the size-matching performance depended on the relative speed of processing semantic and physical information. If the identification process takes longer for silhouettes than for line drawings, a viable explanation of the absence of a semantic effect with silhouettes was that the decision based on size was made before the activation of any semantic information. Put differently, the failure to detect the semantic effect with silhouettes was accounted for by hypothesizing that the visual system successfully exploited the earlier availability of perceptual information to produce an overt response, which was thus not affected by more slowly, but nevertheless automatically, activated semantic information. We have one problem with this interpretation, however. This “horse-race” model does not seem to mesh well with the results of Experiment 2, because it predicts that semantic activation should affect the late portion of the RT distribution. The results of the present Experiment 2, in contrast, show an absence of a semantic effect on both early and late components of the RT distribution. The reliability of this null semantic effect is further reinforced by a comparison between the results of Experiment 1 and Experiment 2. A simple visual inspection of Figure 2 yields the straightforward conclusion that, under the present experimental conditions, semantic information could potentially be activated before a decision about stimulus elongation was made. The longest mean RT observed in Experiment 1 was 563 ms, suggesting that 563 ms was sufficient time to identify a real-world object on the basis of its silhouette. Subjects, however, took, on average, 584 ms to carry out the elongation-decision task, making the response times in the two experiments quite comparable.

As the reader may have surmised at this point, we favour a different explanation of the present findings, which calls on the issue of the modulation of putatively automatic effects on performance by task-induced strategies. The present pattern of results indicates that the activation of structural representations of objects exerted an effect on performance under the requirements of the reality-decision test, whereas the same information was “behaviorally ineffective” under the requirements of the elongation-decision test. In one case, the activation of memory representations was required to successfully carry out the reality-decision task. In the other case, the task could be carried out based on information that did not elicit automatic identification. This dissociation suggests that the automaticity of a mental process is tightly connected with the precipitating conditions in which its effects on behavior are detected (e.g., Melara & Mounts, 1993). It is now widely accepted, for instance, that priming effects are dependent on the task performed on prime stimuli. With words as stimuli, semantic priming effects, that are usually observed in conditions of unconstrained processing of the primes, are absent if subjects are required to perform a letter-search on the primes (e.g., Henik, Friedrich, & Kellogg, 1983). In the classical colour-naming condition of a Stroop task, close-to-nil amounts of word-to-colour interference effects are normally observed when colour information is restricted to portions of the word stimuli (Besner & Stolz, 1999; Besner, Stolz, & Boutilier, 1997). As recently suggested, these findings indicate that, under conditions permitted by particular stimulus- or task-manipulations, functional lock-out points (or context-dependent blocks) may be instantiated in the flow of processing required to carry out a cognitive task (Besner & Stolz, 1999; Besner et al., 1997; Chiappe, Smith, & Besner, 1996; Dell'Acqua & Grainger, 1999; Stolz & Besner, 1996). The results of the present study show that, at least in the testing conditions we have devised, subjects were able to exert control on the activation of semantic information. Therefore, the extreme view on automatic object identification put forward by Boucart and colleagues' (Boucart & Humphreys, 1992; Boucart et al., 1995) should be weakened somewhat in light of the present findings.

Are there possible ways to account for the present results without jeopardizing Boucart and colleagues' view on identification automaticity following the processing of global shape information? Although leaning towards a negative answer to this question, we are also clearly prone to acknowledge the possibility that alternative explanations, potentially congruent with Boucart and colleagues' view, may well apply to the present results, insofar as a theory on the type of information forming the global shape of an object is lacking to date. It may be argued, indeed, that "processing of the global shape" in Boucart and colleagues' view meant the active analysis of a global configuration composed of outline contour and parts, while the present study focused on the extreme case in which "global shape" was reduced to only the outline contour. In addition, it may be argued that the metrics we provided for comparing the speed to access semantic and structural information was misleading, to the extent that the speed of semantic access was established using a procedure (i.e., the reality-decision test) that differed from the elongation-decision test used to derive an estimate of the time taken to access global shape information about the sample of objects used in the present experiments. The strongest case we could build, in this optic, would be to use the same form-matching task used by Boucart and colleagues, and show absence of semantic effects when silhouettes were presented, an issue we leave to future research.

The evidence that identity information is automatically activated upon presentation of line drawings of the objects (Dell'Acqua & Job, 1998), but not of their silhouettes, invites the following cautiously advanced speculation. The objective difference between these two types of stimuli is that features indicating object parts and their spatial relations are present in line drawings, and are absent in silhouettes. Parts play an undoubtedly important role in visual cognition (e.g., Biederman, 1987; Biederman & Cooper, 1991; Hummel &

Biederman, 1992). The present findings lend support to the idea that parts, and not the global shape, play a pre-attentive role in object visual cognition, as recently argued by Lamberts and Freeman (1999). Incidentally, this same idea may be used to explain why identities do not "pop out" of multi-element visual displays when the dimension defining targets and fillers is their global shape (Wolfe & Bennet, 1997; see the introduction to this paper). The present results point to processing strategies in which what is assumed to be later processing, that is, semantic processing, is modulated by stimulus information extracted earlier in processing (e.g., Sanocki, 1991, 1993; Sanocki, Bowyer, Heath, & Sarkar, 1998). To reiterate the present case, automatic recognition did not occur when parts were invisible. Although this type of processing contingency has been advanced to account for variations in the computational costs of object recognition, our suggestion is that a formally equivalent dependence of semantic processing on the type of perceptual information extracted from objects may be postulated to account for the present results. In our view, recognition operates as a ballistic process only when particular object constituents are provided to the visual system. The present circumstances suggest that processing of the global shape renders the activation of identity information contextually tunable.

REFERENCES

Balota, D.A., & Spieler, D.H. (1999). Word frequency, repetition, and lexicality effects in word recognition tasks: Beyond measures of central tendency. *Journal of Experimental Psychology: General, 128,* 32–55.

Besner, D., & Stolz, J.A. (1999). Unconsciously controlled processing: The Stroop effect reconsidered. *Psychonomic Bulletin and Review, 6,* 449–455.

Besner, D., Stolz, J.A., & Boutilier, C. (1997). The Stroop effect and the myth of automaticity. *Psychonomic Bulletin and Review, 4,* 221–225.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94,* 115–147.

Biederman, I., & Cooper, E.E. (1991). Priming contour-deleted images: Evidence for intermediate representations in visual object recognition. *Cognitive Psychology, 23,* 393–419.

Boucart, M., & Humphreys, G.W. (1992). Global shape cannot be attended without object identification. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 785–806.

Boucart, M., & Humphreys, G.W. (1994). Attention to orientation, size, luminance and color: Attentional failure within the form domain. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 61–80.

Boucart, M., Humphreys, G.W., & Lorenceau, J. (1995). Automatic access to object identity: Global information, not particular physical dimensions, is important. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 584–601.

Chiappe, P.R., Smith, M.C., & Besner, D. (1996). Semantic priming in visual word recognition: Activation blocking an domains of processing. *Psychonomic Bulletin and Review, 3,* 249–253.

Dell'Acqua, R., & Grainger, J. (1999). Unconscious semantic priming from pictures. *Cognition, 73*, B1–B15.

Dell'Acqua, R., & Job, R. (1998). Is object recognition automatic? *Psychonomic Bulletin and Review, 5,* 496–503.

Heathcote, A. (1996). RTSYS: A DOS application for the analysis of reaction time data. *Behavior Research Methods, Instruments, and Computers, 28,* 427–445.

Heathcote, A., Popiel, S.J., & Mewhort, D.J.K. (1991). Analysis of response time distributions: An example using the Stroop task. *Psychological Bulletin, 109,* 340–347.

Henik, A., Friedrich, F.J., & Kellogg, W. (1983). The dependence of semantic relatedness effects upon prime processing. *Memory and Cognition, 11,* 366–373.

Hummel, J.E., & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review, 99,* 480–517.

Humphreys, G.W., Lamote, C., & Lloyd-Jones, T.J. (1995). An interactive activation approach to object processing: Effects of structural similarity, name frequency and task in normality and pathology. *Memory, 3,* 535–586.

Kroll, J.F., & Potter, M.C. (1984). Recognizing words, pictures, and concepts: A comparison of lexical, object, and reality decision. *Journal of Verbal Learning and Verbal Behavior, 23,* 39–66.

Lamberts, K., & Freeman, R.P.J. (1999). Building object representations from parts: Tests of a stochastic sampling model. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 904–926.

Loftus, G.F., & Masson, M.E J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin and Review, 1,* 476–490.

Lowe, D.G. (1985). *Perceptual organization and visual recognition*. Boston: Kluwer Academic.

Lupker, S.J. (1988). Picture naming: An investigation of the nature of categorical priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 444–455.

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

Melara, R.D., & Mounts, J. R. W. (1993). Selective attention to Stroop dimensions: Effects of baseline discriminability, response mode, and practice. *Memory and Cognition, 21,* 627–645.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology, 9,* 353–383.

Nelder, J.A., & Mead, R. (1965). A simple method for function minimization. *Computer Journal, 7,* 308–313.

Peterson, M.A., & Gibson, B.S. (1994). Must figure-ground organization precede object recognition? An assumption in peril. *Psychological Science, 5,* 253–259.

Pike, R., & Ryder, P. (1973). Response latencies in yes/no detection task: An assessment of two basic models. *Perception and Psychophysics, 13,* 224–232.

Quinlan, P.T. (1991). Differing approaches to two-dimensional shape recognition. *Psychological Bulletin, 109,* 224–241.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85,* 59–108.

Sanocki, T. (1991). Effects of early common features on form perception. *Perception and Psychophysics, 50,* 490–497.

Sanocki, T. (1993). Time course of object identification: Evidence for a global-to-local contingency. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 878–898.

Sanocki, T., Bowyer, K.W., Heath, M.D., & Sarkar, S. (1998). Are edges sufficient for object recognition? *Journal of Experimental Psychology: Human Perception and Performance, 24,* 340–349.

Schwarz, W., & Ischebeck, A. (1994). Coactivation and statistical facilitation in the detection of lines. *Perception, 23,* 157–168.

Snodgrass, J.G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 174–215.

Stolz, J.A., & Besner, D. (1996). The role of set in visual word recognition: Activation and activation blocking as non-automatic processes. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 1166–1177.

Ullman, S. (1996). *High-level vision: Object recognition and visual cognition*. Cambridge, MA: MIT Press.

Wolfe, J.M., & Bennet, S.C. (1997). Preattentive object files: Shapeless bundles of basic features. *Vision Research, 37,* 25–43.

Zeki, S.M. (1993). *A vision of the brain.* Oxford, UK: Blackwell Scientific Publications.

Manuscript received November 1999
Revised manuscript received July 2000