

Is object recognition automatic?

ROBERTO DELL'ACQUA and REMO JOB
University of Padua, Padua, Italy

A new paradigm is proposed that involves a simple judgment on an object's perceptual feature that is independent of object identity. Subjects were required to categorize as vertical or horizontal the main axis of elongation of an object picture. Both fake and real-world objects were presented, and a graphical manipulation was applied to their shapes so that the canonical elongation of the real-world objects was incongruent with the elongation of their shapes after the manipulation. The results showed an influence of the identity of the objects on the perceptual task in the form of a cost in judging the elongation of the incongruent real-world objects. The results are taken as evidence for automatic activation of objects' stored representations. A "horse-race" model of the influence of these representations on the perceptual task is proposed in the final section of the article.

One interesting question concerning the ability to identify an object is whether such a mental skill can be regarded as automatic—namely, whether identification cannot be prevented from running to completion once an object is presented within the visual field. Three different kinds of psychological effects usually have been taken to reflect the automaticity of object identification.

The first effect is the so-called negative priming effect. Tipper (1985) presented subjects with two superimposed pictures of objects (prime display) and asked them to selectively attend to only one of the two stimuli (target) and to ignore the other (distractor). A third picture (probe) was presented following the prime display, and subjects were required to name the probe. Probes and distractors could be identical, semantically related, or unrelated. Results showed that naming responses were slower when probes and distractors were either identical or semantically related than when they were unrelated. Negative priming effects have also been observed in studies in which the symbolic domain of the stimuli was varied—that is, when written words were presented as probes, and pictures as distractors (Tipper & Driver, 1988).

The second effect is the semantic priming effect. In a series of elegant experiments, Carr, McCauley, Sperber, and Parmelee (1982; see also McCauley, Parmelee, Sperber, & Carr, 1980) employed a staircase procedure to establish recognition thresholds for pictures presented under masked exposure and showed substantial facilitation in a naming task when target pictures or target words were preceded by subthreshold related picture primes. Similar re-

sults have also been found by Vanderwart (1984), who showed significant cross-form priming in a lexical decision task when the stimulus onset asynchrony (SOA) between prime pictures and related target words was shorter than 250 msec. This interval had been indicated by Neely (1977; see also Heyer, Briand, & Smith, 1985) as critical for the detection of automatic, as opposed to strategic (conscious), facilitatory effects.

The third type of effect is picture-word interference. Smith and Magee (1980) presented subjects with lists of object names, each superimposed on an object's picture. Although subjects were told that the pictures were of no importance for the task, the results showed that the relationship between names and pictures had a dramatic impact on subjects' performance. Categorizing the words was disrupted when incongruent pairings (names superimposed on semantically unrelated objects) were included in the lists. Glaser and Dungenhoff (1984) and Glaser and Glaser (1989) extended these results by showing longer categorization times when the presentation of a single object was accompanied by the presentation of a semantically related distractor (see MacLeod, 1991, for a detailed review).

In all these studies, the automaticity of object recognition has been conceptually maintained in light of the evidence that although informed about the irrelevance of part of the stimulation presented on each trial, subjects could not refrain from processing the information contained in either prime or distracting objects. Some, however, have questioned this empirical characterization of the automaticity of object recognition. Boucart and Humphreys (1992, 1994; Boucart, Humphreys, & Lorenceau, 1995), for example, have pointed out that identification of the irrelevant information might be contingent on the particular task requirements adopted in these studies. Specifically, the requirement of categorizing or naming target objects could have caused some form of logical recoding of the instructions (Simon & Sudalaimuthu, 1979) and therefore could have induced semantic processing of all the stimuli presented on a given trial. Fol-

This work was supported by a grant from CNR (96.01848.CT11) awarded to R.J. Part of the results reported in the present article was presented at the Seventh Conference of the European Society for Cognitive Psychology, Lisbon, 1994. The authors are grateful to Pierre Jolicoeur, Jonathan Grainger, and one anonymous reviewer for useful comments on a previous draft of this manuscript. Correspondence should be addressed to R. Job, Dipartimento di Psicologia dello Sviluppo e della Socializzazione, via Venezia, 8, 35100 Padova, Italy (e-mail: remojob@psico.unipd.it).

lowing their argument, the critical evidence for automatic object identification would be provided by a task that required perceptual processing of the relevant information. To provide such a test, Boucart and her colleagues adopted a variant of a successive form-comparison task and presented subjects with a single object (reference), followed at a fixed temporal interval by a pair of objects. A speeded response was required to indicate which of the pair of objects matched the reference according to global-shape features such as spatial orientation and size (Boucart & Humphreys, 1992, 1994). Even in this case, although subjects were explicitly instructed to filter out information about the identity of to-be-matched objects, the results showed faster matching responses when the target and the reference were semantically related than in the condition in which they were unrelated.

The putatively more stringent evidence provided by Boucart and Humphreys's (1992, 1994) paradigm, however, may be challenged by the existence of at least two potential confounds. The first confound arises from the long (650-msec) interval separating the presentation of the reference and the target stimulus. One might argue, indeed, that the effect they found was brought about by semantic activation mediated by holding in memory the reference stimulus for that period of time—that is, until the target was presented and an outcome for the matching task was selected. The second possible confound stems from their use of a matching task. The argument in this case would be that a successive comparison could be performed only by engaging particular processing pathways (leading to identification) that could be actively inhibited when performing a different task.

In order to overcome these confounds, as well as those previously discussed, the present investigation was based on a paradigm in which the logic of requiring perceptual, instead of semantic, processing of task-relevant information was implemented. In addition, the confounds linked to the use of a comparison task were ruled out by requiring a two-alternative forced-choice decision about a perceptual feature of a single object per trial. As will be more clear in what follows, the rationale of the experiment rested on the Stroop logic of comparing subjects' performance in two different conditions, one in which task-relevant information about a perceptual feature of an object was congruent with task-irrelevant information about its identity, and one in which these two types of information were incongruent.

In the present study, subjects were required to examine a picture displayed on the screen of a computer and decide as quickly as possible whether its main axis of elongation was aligned more to the vertical or more to the horizontal direction. Half of the stimuli were composed of pictures of real-world objects (hereafter, targets). A graphical manipulation was applied to their shapes in order to generate incongruity between the task-relevant perceptual dimension (actual elongation on the screen) and the task-irrelevant "semantic" dimension (canonical elongation of the target objects). For example, consider

the main axis of elongation of a kangaroo depicted in a "usual" perspective (see top left panel in Figure 1). Relative to the vertical and the horizontal coordinates of the frontal surface of the screen, such an axis would clearly be aligned more with the vertical than with the horizontal direction. In this case, the to-be-judged perceptual dimension—that is, the main axis of elongation of the actual pattern on the screen—would be congruent with what can be confidently assumed to be the main axis of elongation of a canonical representation of a kangaroo. For the case of the incongruent stimuli, the graphical manipulation we applied was to reverse the extent of the perpendicular axes of elongation in such a way that we shortened the vertical axis of elongation and stretched the horizontal axis of elongation of the kangaroo. Thus, incongruity between task-relevant and task-irrelevant information was caused by the fact that the main axis of elongation of the actual pattern on the screen (now horizontal) did not correspond to the main axis of elongation of the canonical representation of a kangaroo. The other half of the stimuli were composed of meaningless fillers generated from the target objects. The fillers were generated by modifying the internal structure of the target objects and by maintaining most of their external structure. Fillers were submitted to the same graphical manipulation as that applied to the target objects. Since the fillers closely matched the target objects with respect to their global shape characteristics, our assumption was that decisions about the main axis of elongation of the filler stimuli could constitute an appropriate baseline against which to estimate the possible effect of task-irrelevant identity information on the perceptual decision (see DiGirolamo & Kanwisher, 1995, for the use of an equivalent manipulation). We introduced this graphical manipulation because subjects have been shown to have a preferred (canonical) perspective from which to both view and imagine objects. One of the key characteristics of this canonical representation is that it maximizes salient features of an object's shape, such as its main axis of elongation in 3-D space (Marr, 1982; Marr & Nishihara, 1978). In addition, processing of an object seen from distorted perspectives is impaired as compared with processing of an object in its canonical perspective (Johnson, Paivio, & Clark, 1996; Palmer, Rosch, & Chase, 1980; see also Ullman, 1996, chap. 7).

In our paradigm, the instructions stressed the importance of focusing on the actual elongation of the stimuli on the screen. Thus, given that information about the targets' identity was to be neglected, the prediction was straightforward. If the presentation of a target object automatically elicited the activation of its canonical representation, we could reasonably predict that such activation should influence the perceptual judgment of its elongation. Although both a pattern of facilitation and a pattern of interference could be predicted, we suspected that the Stroop-like logic implemented by this new paradigm would have brought about the well-known asymmetry between these two effects; that is, if an effect showed

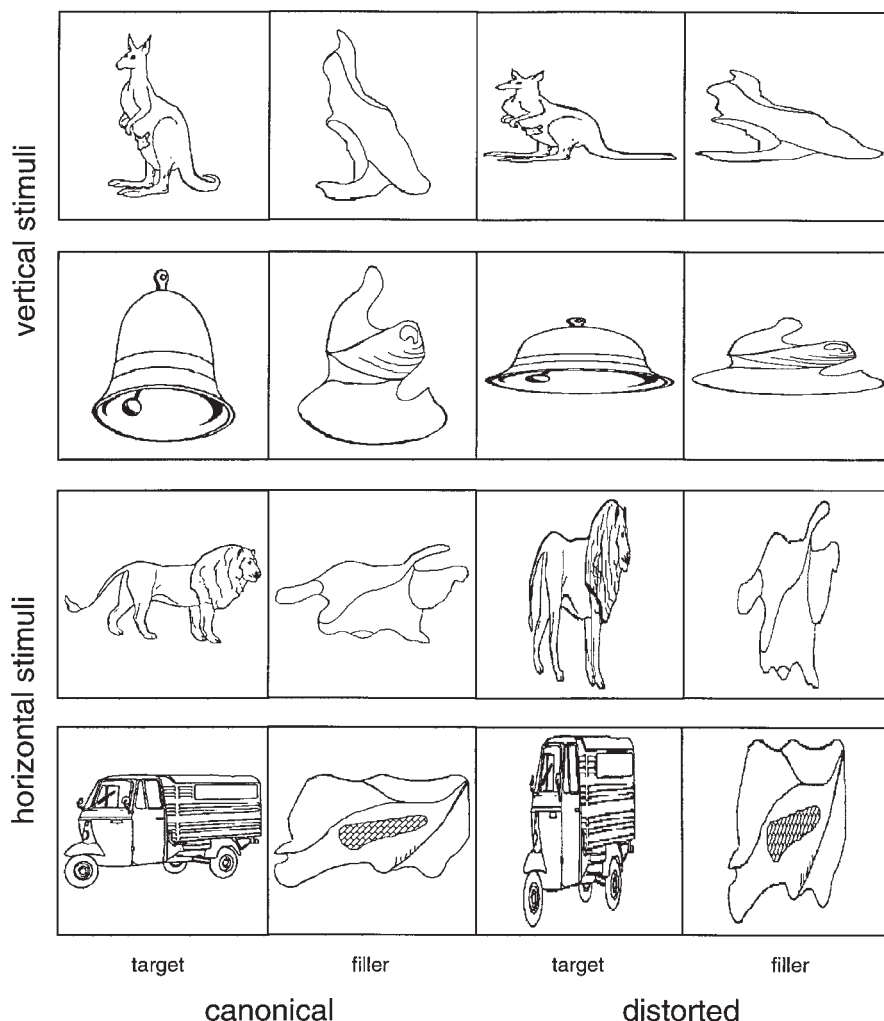


Figure 1. Examples of targets and fillers, in both canonical (left) and distorted (right) format. Stimuli with a vertical elongation are shown in the top two rows; stimuli with a horizontal elongation are shown in the bottom two rows.

up, it would be more likely to have been detected in the condition in which task-relevant and task-irrelevant information were incongruent (interference) than the opposite (facilitation).

METHOD

Subjects

Subjects were 28 undergraduate students (17 females) at the University of Padua who participated for course credit. The age of the subjects ranged from 24 to 31 years. All were naive to the purpose of the experiment and all reported normal or corrected-to-normal vision.

Design and Stimuli

Target stimuli consisted of line drawings of 32 real objects (half animals and half artifacts) selected from the Snodgrass and Vanderwart (1980) set and the British Picture Vocabulary Scale (Dunn, Dunn, Whetton, & Pintilie, 1982; see the Appendix for a list of the material). Half of the targets had a vertical axis of canonical elongation (vertical targets; e.g., ostrich and candle), and half had a horizontal axis of canonical elongation (horizontal targets; e.g., rhinoceros and shoe).¹

The familiarity of the meaningful stimuli was rated by asking a group of 10 subjects to judge on a 7-point scale how familiar the targets were according to their own personal experience. Mean familiarity values are reported in the Appendix. A one-way analysis of variance (ANOVA) was performed on these values, with elongation (vertical vs. horizontal) as a within-subjects factor. The mean rating values did not differ between the two sets of stimuli ($F < 1$). A corresponding set of 32 fillers was generated by altering both the segments of the contour and the vertices inside the contour of each target. Thus, each filler exactly matched the corresponding target in terms of the extent of the vertical and horizontal axes of elongation and roughly matched it in terms of the global shape of the contour (see Figure 1 for an example). An attempt was made to keep constant the orientation of the internal line segments and the visual complexity of the stimuli.

To create the incongruent stimuli, the perceptual layouts of both the targets and the fillers were submitted to a further graphical manipulation. The original (canonical) format of the stimuli was modified so that the canonical elongation of the stimuli was reversed. A set of 64 distorted stimuli was generated by shortening the vertical axis, lengthening the horizontal axis of the vertical stimuli, and applying the opposite 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 an horizontal elongation (see Figure 1 for an example²).

In order to evaluate the effect of the inversion of the canonical elongation on the recognizability of the material used in the present experiment, each stimulus was displayed on the screen of a computer, followed by a mask. The exposure durations of the stimulus and the mask (150 and 100 msec, respectively) and the general experimental settings were the same as in the experiment (see below). The subjects ($n = 6$) were instructed to name the stimulus when it depicted a target object and to say "fake" otherwise. Since the only point of this task was to show that a sufficient amount of information about the identity of the stimuli could be extracted within the masked exposure duration of 150 msec, the response was unspeeded. For targets, both "fake" responses and incorrect names were treated as errors; for fillers, errors were constituted by productions of real concepts' names. The rather negligible error rate indicated that despite the radical modification of their canonical shape, all the targets could be accurately recognized (0.47% of errors) and distinguished from the corresponding fillers (0.28% of errors). Mean error rates were submitted to an ANOVA, with type of stimulus (target vs. filler) and format (canonical vs. distorted) as within-subjects factors. No effect was significant (all F s < 1).

The major/minor axis ratio (reported in the Appendix) of each stimulus was computed as an index of the absolute vertical and horizontal elongation of the stimuli when displayed on the computer screen. An ANOVA was performed on these values, with elongation (vertical vs. horizontal) and format (canonical vs. distorted) as within-subjects factors. The analysis showed no significant effects (all F s < 1).

In summary, three factors were independently manipulated in our experimental design. The first factor was type of stimulus, which had two levels, targets (i.e., meaningful objects) and fillers (i.e., meaningless patterns). The second factor was elongation of the stimuli, with two levels, vertical stimuli (e.g., a giraffe and its associate filler) and horizontal stimuli (e.g., a rhinoceros and its associate filler). The third factor was format, with two levels, canonical and distorted stimuli, which were defined by the absence or presence of the graphical manipulation that characterized half of the total list of stimuli. Note that 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. That is, whereas a vertical and a horizontal stimulus in the canonical condition had, respectively, a vertical and a horizontal elongation on the screen, in the distorted condition they actually had a horizontal and a vertical elongation, respectively (Figure 1).

Procedure

The stimuli were displayed on a 14-in. Macintosh RGB screen (resolution 640×480 pixels) controlled by an Apple computer (Macintosh Quadra 700). At a viewing distance of 90 cm, each stimulus could be inscribed in a square 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 center of the screen for 500 msec. At the offset of the fixation point, a stimulus was displayed for 150 msec. Following the stimulus, a masking pattern was exposed for 100 msec. The masking pattern was generated by filling a circle (7° of diameter) with a blurred background composed of random dots and lines. This particular masking pattern was generated so that no stimulus could be recognized when the stimulus and the mask overlapped on the screen.

Subjects were instructed to decide as fast 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.

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 ran-

domly ordered within the following constraints: No more than four consecutive repetitions of the same response (vertical vs. horizontal), type of stimulus (target vs. filler), and format (canonical vs. distorted) were allowed to occur. Order of block presentation and hand-response pairings (vertical-horizontal) were fully counterbalanced across subjects. Before the beginning of the experimental session, each subject performed 32 practice trials with stimuli that were not included in the experimental session. The experiment was about 45 min long.

RESULTS

The analyses focused on correct RTs and error rates. Correct RTs were first screened for outliers using a procedure based on the computation of the cutoff bounds of each subject's performance. Cutoff bounds were computed by adding (higher bound) and subtracting (lower bound) 2 SD s from each subject's overall mean RT. Each RT not included in the range delimited by the cutoff bounds (less than 1%) was replaced by the value corresponding to the respective cutoff bound. Correct RTs and error rates were then submitted to an ANOVA, with both subjects ($F1$) and items ($F2$) as random factors. The factors involved in the ANOVA 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 factor. Elongation was treated as between in the analysis with items as random factor. The significance level chosen was $p < .05$.

RTs Analysis

The overall mean RT was 615 msec. The main effect of the type of stimulus was significant [$F1(1,27) = 9.3$, $MS_e = 2,721$, $p < .001$, and $F2(1,14) = 7.0$, $MS_e = 1,432$, $p < .02$], indicating faster RTs to fillers (605 msec) than to targets (626 msec). The main effect of format was significant in the analysis with subjects as random factor only, indicating faster RTs to canonical stimuli (605 msec) than to distorted stimuli [$F1(1,27) = 19.9$, $MS_e = 1,280$, $p < .001$, and $F2(1,14) = 1.4$, $MS_e = 5,641$, $p < .27$]. The interaction between type of stimulus and format was significant [$F1(1,27) = 19.2$, $MS_e = 1,788$, $p < .001$, and $F2(1,14) = 9.5$, $MS_e = 5,642$, $p < .01$]. The results are reported in panel A of Figure 2, where RTs are plotted as a function of the type of stimulus and the format of the stimuli. In panels B1 and B2, the results are reported for each of the levels of the factor elongation. As can be seen, RTs to distorted targets (655 msec) were slower than those to canonical targets, canonical fillers, and distorted fillers (605, 608, and 601 msec, respectively), which did not differ from each other [$F(1,27) = .093$, $MS_e = 317$, $p > .76$]. This pattern was evident for both vertical (panel B1) and horizontal (panel B2) stimuli.

The main effect of elongation was significant [$F1(1,27) = 22.3$, $MS_e = 1,955$, $p < .001$, and $F2(1,14) = 7.5$, $MS_e = 1,332$, $p < .02$], reflecting faster RTs to vertical stimuli (601 msec) than to horizontal stimuli (629 msec). The factor elongation interacted with format [$F1(1,27) = 20.1$, $MS_e = 7,186$, $p < .001$, and $F2(1,14) = 9.1$, $MS_e = 5,642$, $p < .01$]. The results indicated that RTs to vertical

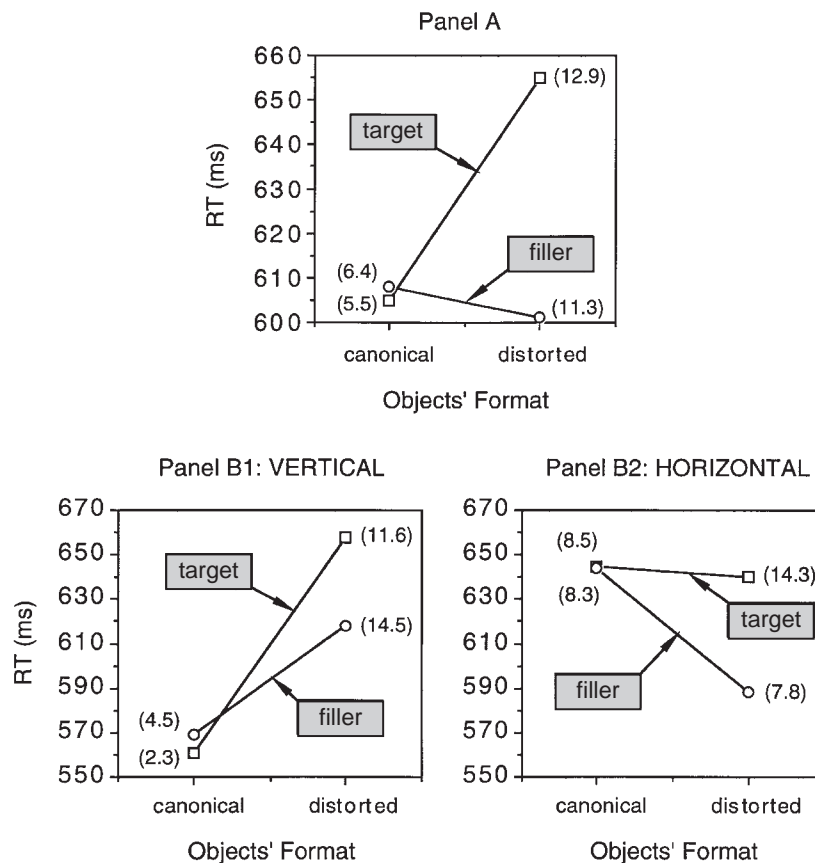


Figure 2. Panel A: Mean response times (RTs) and mean percentages of errors (in parentheses) as a function of type of stimulus and format. Panel B1: Mean RTs and mean percentages of errors (in parentheses) as a function of type of stimulus and format for vertical stimuli. Panel B2: Mean RTs and mean percentages of errors (in parentheses) as a function of type of stimulus and format for horizontal stimuli.

(i.e., when they were vertically elongated) stimuli were faster in the canonical condition (566 msec) than in the distorted (i.e., when they were horizontally elongated) condition [644 msec; $F(1,27) = 4.5$, $MS_e = 1,681$, $p < .001$] and that RTs to horizontal stimuli were faster in the distorted (i.e., when they were vertically elongated) condition (613 msec) than in the canonical (i.e., when they were horizontally elongated) condition [645 msec; $F(1,27) = 6.3$, $MS_e = 4,037$, $p < .02$]. There was no hint of a three-way interaction (all $F_s < 1$).

Error Rate Analysis

The overall mean error rate was 9%. The main effect of format was significant [$F(1,27) = 28.4$, $MS_e = .008$, $p < .001$, and $F(1,14) = 7.7$, $MS_e = .008$, $p < .02$], reflecting a higher error rate for distorted stimuli (12.1%) than for canonical stimuli (6.0%). Format interacted with the elongation of the stimuli in the analysis with subjects as random factor only [$F(1,27) = 7.6$, $MS_e = .004$, $p < .02$, and $F(1,14) = 2.5$, $MS_e = .008$, $p > .1$]. The results indicated a lower error rate for the vertical stimuli in the canonical condition (5.4%) than in the distorted

condition [13.0%; $F(1,27) = 31.9$, $MS_e = .008$, $p < .001$], whereas no difference was found in the error rates for the horizontal stimuli in the canonical and the distorted conditions [9.4% and 11.0%, respectively; $F(1,27) = 2.4$, $MS_e = .008$, $p > .13$]. Type of stimulus interacted with elongation [$F(1,27) = 9.1$, $MS_e = .005$, $p < .01$ and $F(1,14) = 5.2$, $MS_e = .003$, $p < .04$]. The results indicated a higher error rate for the targets in the horizontal condition (7.0%) than in the vertical condition [11.4%; $F(1,27) = 5.0$, $MS_e = .011$, $p < .04$], whereas no such difference was found for vertical and horizontal fillers [9.4% and 8.1%, respectively; $F(1,27) = 1.7$, $MS_e = .004$, $p > .2$].

Summary

RTs were slower and error rates were higher when subjects had to judge the elongation of the distorted targets than when they had to judge the elongation of the corresponding distorted fillers, the canonical targets, or the canonical fillers. This result strongly suggests a semantic influence on the ability to perform a simple task based on a decision that could be made regardless of the

object's identity. As predicted, and mirroring the pattern commonly found in most of the Stroop studies (see MacLeod, 1991), the effect of identity information on the perceptual judgment turned out to be substantial in the incongruent condition, but not in the congruent condition. This lack of a difference between targets and fillers in the congruent (i.e., canonical) condition may seem inconsistent, at first blush, with a predictable object-superiority effect (Weisstein & Harris, 1974) in the results of the present paradigm. In fact, this type of effect is normally found when the perceptual task requires processing of a feature *embedded* in an object, and not when the task requires processing of an object as a whole (Boucart, Delord, & Giersch, 1994). Although several accounts have been proposed to explain the consistent asymmetry between facilitation and interference in Stroop tasks (e.g., Cohen, Dumbar, & McClelland, 1990; Glaser & Glaser, 1989; Lindsay & Jacoby, 1994), the case we build in the forthcoming section will be restricted to the fact that an interaction was found, and that this interaction reflects, at some level, the influence of automatically activated stored representations of real-world objects on the speed of processing one feature of their perceptual appearance.

A second result is that RTs to vertically elongated stimuli (i.e., canonical/vertical and distorted/horizontal stimuli) were faster than those to horizontally elongated stimuli (i.e., canonical/horizontal and distorted/vertical stimuli). This bias toward more prompt responses to vertically oriented objects has also been found by Schwarz and Ischebeck (1994) in the context of line orientation judgments. Although this effect is interesting, a closer investigation is beyond the scope of the present work.

DISCUSSION

In the classical Stroop experiments, the relevant dimension is the color of a word, and the irrelevant dimension is the word itself. When color and word information are incongruent (green letters spelling the word *red*), a cost is usually reported in processing the color. In our experiment, the relevant dimension was the elongation of an object, and the irrelevant dimension was the identity of the object. When actual and canonical elongation of real-world objects did not match, perceptual judgments were hindered. This evidence suggests that at some level during the flow of visual processing, a cross-talk between one or more stages required by the perceptual task and one or more stages involved in (involuntary) processing of an object's identity must occur. Recent accounts of the Stroop effect given in the framework of parallel distributed models (e.g., Cohen et al., 1990), for example, assume that the flow of activation for relevant and irrelevant information follows distinct parallel pathways, and interference arises where the two simultaneously active pathways intersect at the level of the selection of an overt response. With an important difference, such a model would nicely fit the results of the present study: Whereas color and lexical information can conceivably be assumed to undergo computationally separable stages of processing, it is difficult to adopt an equivalent perspective dealing with shape and identity information. More feasibly, one has to assume that a high degree of functional dependency characterizes these two types of information (Pinker, 1984; Ullman, 1984, 1989). From a logical point of view, stored representations of objects can be activated only after a certain amount of perceptual information available in the stimulus has been processed (see, e.g., Hummel & Biederman, 1992). Furthermore, from an empirical point of view, both global-on-local precedence ef-

fects (Navon, 1977; Podrouzek, Modigliani, & Di Lollo, 1992; Pomeranz, 1983) and naming errors due to global visual similarity among objects (Vitkovitch, Humphreys, & Lloyd-Jones, 1993) suggest that the perceptual information computed before complete access to a stored representation concerns global properties of an object's shape (see also Marr, 1982).

An interesting prediction emerges from this argument that is based on the absolute speed in performing the perceptual task of the present study. For factors generally catalogued under the rubric of preparatory strategy (De Jong & Sweet, 1994; Pashler, 1994), a fast and correct performance in a task is usually taken to reflect the outcome of a system that has been set to efficiently cope with prespecified environmental demands. In an experimental situation, subjects are instructed to process some kind of stimulation and execute a response according to specific, task-relevant attributes of it. Despite these instructions, however, a certain amount of trial-by-trial variability in setting the correct preparation for the task is to be expected. Indeed, when a speeded response is required, fast RTs are taken to reflect a particularly successful attempt to set the internal system to cope with the instructions, whereas slow RTs reflect a looser attempt to deploy advanced preparation for the most efficient performance. In our experiment, subjects were explicitly instructed to respond as fast as possible to specific information about the global shape of the objects (i.e., elongation) and to filter out information about their identity. As argued, fast RTs probably reflect a more efficient preparation for the task. The prediction is as follows: Given the prior availability of the perceptual information with respect to the activation of objects' stored representations, fast RTs in judging the elongation of the objects should reflect a performance less affected by the semantic factor manipulated in the present paradigm.

To test this hypothesis, a further analysis was carried out on the present data set. For each subject, the total number of trials was divided into the categories type of stimulus (target vs. filler) and format (canonical vs. distorted). RTs in each of the four cells were further divided into tertiles, with this three-level variable treated as an additional within-subjects variable. Mean RTs in the resulting $2 \times 2 \times 3$ design were then submitted to an ANOVA, with all the factors treated as within-subjects factors. As predicted, the results for the RTs yielded a significant three-way interaction [$F(2,54) = 8.9$, $MS_e = 1,029$, $p < .001$], which lends statistical support to the striking result plotted in Figure 3. As predicted, the cost of judging the elongation of the distorted targets increased as RTs became longer.³

The results of our study deserve a further comment. As recently pointed out by Melara and Mounts (1993), definitions of mental processes in terms of the dichotomy automatic versus voluntary are fraught with problems. On the one hand, one is led to infer from a consistent body of evidence that a particular effect is automatic in all the circumstances tested to date (i.e., in an absolute sense). On the other hand, "there is no guarantee that this effect will withstand all forthcoming strategies or manipulations" (Melara & Mounts, 1993, p. 643). This issue is not new, and calls into play recent findings concerning the elimination of normally robust automatic effects as a consequence of a number of task and stimulus manipulations. Specifically, striking demonstrations of the influence of the "set" adopted by subjects in performing a task on putative automatic processes are accruing in the field of studies on automaticity. For example, it is well known that semantic priming effects disappear if subjects perform a letter search on the prime (Henik, Friedrich, & Kellogg, 1983; Henik, Friedrich, Tzelgov, & Tramer, 1994; Smith, Theodor, & Franklin, 1983). Furthermore, even the received view of the Stroop interference as caused by mandatory activation of stored lexical representations (Cohen et al., 1990; see also MacLeod, 1991) has been repeatedly challenged. An elegant demonstration in this context has been provided by Besner, Stolz, and Boutilier (1997). In a classical color categorization task, subjects had to process the color of printed color words. On half of the trials, fully colored words were presented, and on the other half of the trials, only one letter of the words was colored. The results were clear-cut. A smaller Stroop effect was found for words with only one letter colored than for fully colored words. This finding seems to fly in the face of models that rest on the assumption that the presentation of a word "automatically" elicits processing of its meaning. Instead, the authors suggested that this "less extensive processing" of the meaning of the words reflects a context-dependent block that constrains the flow of activation between serially

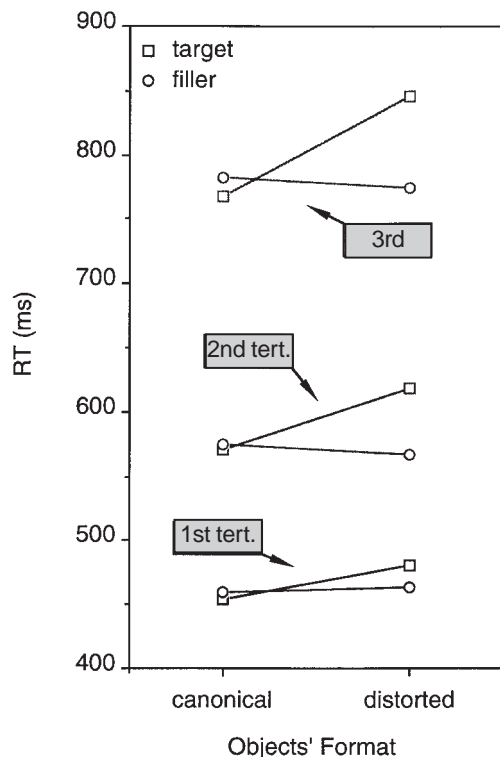


Figure 3. Mean response times (RTs) as a function of type of stimulus, format, and RT tertile.

organized lexical and semantic stages of processing (see also Chiappe, Smith, & Besner, 1996; Henderson, 1987; Smith, Besner, & Miyoshi, 1994; Stolz & Besner, 1996). Given this bulk of evidence, we would like to conclude provocatively by stressing that the effect we found was restricted to the particular settings and task requirements we have described in the Method section of this paper. We found what Melara and Mounts (1993) called a "precipitating" condition, in which object identification is unavoidable, and we leave the interesting question of which are the conditions in which identification can be strategically controlled to future investigations.

REFERENCES

- BESNER, D., STOLZ, J. A., & BOULIER, C. (1997). The Stroop effect and the myth of automaticity. *Psychonomic Bulletin & Review*, *4*, 221-225.
- BOUCART, M., DELORD, S., & GIERSCH, A. (1994). The computation of contour information in complex objects. *Perception*, *23*, 399-409.
- BOUCART, M., & HUMPHREYS, G. W. (1992). Global shape cannot be attended without object identification. *Journal of Experimental Psychology: Human Perception & 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 & 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 & Performance*, *21*, 584-601.
- CARR, T. H., MCCAULEY, C., SPERBER, R. D., & PARMELEE, C. M. (1982). Words, pictures, and priming: On semantic activation, conscious identification, and the automaticity of information processing. *Journal of Experimental Psychology: Human Perception & Performance*, *8*, 757-777.
- CHIAPPE, P. R., SMITH, M. C., & BESNER, D. (1996). Semantic priming in visual word recognition: Activation blocking and domains of processing. *Psychonomic Bulletin & Review*, *3*, 249-253.
- COHEN, J. D., DUMBAR, K., & MCCLELLAND, J. L. (1990). On the control of automatic processes: A parallel distributed account of the Stroop effect. *Psychological Review*, *97*, 332-361.
- DE JONG, R., & SWEET, J. B. (1994). Preparatory strategies in overlapping-task performance. *Perception & Psychophysics*, *55*, 142-151.
- DIGIROLAMO, G. J., & KANWISHER, N. (1995, November). *Accessing stored representations begins within 155 msec in object recognition*. Paper presented at the annual meeting of the Psychonomic Society, Los Angeles.
- DUNN, L. M., DUNN, L. M., WHETTON, C., & PINTILIE, D. (1982). *British Picture Vocabulary Scale*. Windsor, U.K.: NFER-NELSON.
- GLASER, W. R., & DUNGELHOFF, F.-J. (1984). The time course of picture-word interference. *Journal of Experimental Psychology: Human Perception & Performance*, *10*, 640-654.
- GLASER, W. R., & GLASER, M. O. (1989). Context effects in Stroop-like word and picture processing. *Journal of Experimental Psychology: General*, *118*, 13-42.
- HENDERSON, L. (1987). Word recognition: A tutorial review. In M. Coltheart (Ed.), *Attention & Performance XII* (pp. 171-200). Hillsdale, NJ: Erlbaum.
- HENIK, A., FRIEDRICH, F. J., & KELLOGG, W. (1983). The dependence of semantic relatedness effects upon prime processing. *Memory & Cognition*, *11*, 366-373.
- HENIK, A., FRIEDRICH, F. J., TZELGOV, J., & TRAMER, S. (1994). Capacity demands of automatic processes in semantic priming. *Memory & Cognition*, *22*, 157-168.
- HEYER, K. DEN, BRIAND, K., & SMITH, L. (1985). Automatic and strategic effects in semantic priming: An examination of Becker's verification model. *Memory & Cognition*, *13*, 228-232.
- HUMMEL, J. E., & BIEDERMAN, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review*, *99*, 480-517.
- JOHNSON, C. J., PAIVIO, A., & CLARK, J. M. (1996). Cognitive components of picture naming. *Psychological Bulletin*, *120*, 113-139.
- LINDSAY, D. S., & JACOBY, L. L. (1994). Stroop process dissociation: The relationship between facilitation and interference. *Journal of Experimental Psychology: Human Perception & Performance*, *20*, 219-234.
- MACLEOD, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*, 163-203.
- MARR, D. (1982). *Vision*. San Francisco: Freeman.
- MARR, D., & NISHIHARA, H. K. (1978). Representation and recognition of the spatial organisation of three-dimensional shapes. *Proceedings of the Royal Society of London: Series B*, *200*, 269-294.
- MCCAULEY, T. H., PARMELEE, C. M., SPERBER, R. D., & CARR, T. H. (1980). Early extraction of meaning from pictures and its relation to conscious identification. *Journal of Experimental Psychology: Human Perception & Performance*, *6*, 265-276.
- MELARA, R. D., & MOUNTS, J. R. W. (1993). Selective attention to Stroop dimensions: Effects of baseline discriminability, response mode, and practice. *Memory & Cognition*, *21*, 627-645.
- NAVON, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, *9*, 353-383.
- NEELY, J. H. (1977). Semantic priming and retrieval from lexical memory: Role of inhibitionless spreading of activation and limited-capacity attention. *Journal of Experimental Psychology: General*, *106*, 226-254.
- PALMER, S., ROSCH, E., & CHASE, P. (1980). Canonical perspective and the perception of objects. In J. Long & A. D. Baddeley (Eds.), *Attention & Performance IX* (pp. 135-151). Hillsdale, NJ: Erlbaum.
- PASHLER, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*, *2*, 220-244.
- PINKER, S. (1984). Visual cognition: An introduction. *Cognition*, *18*, 1-63.
- PODROUZEK, K. W., MODIGLIANI, V., & DI LOLLO, V. (1992). Lateral masking as a determinant of global dominance. *Perception*, *21*, 705-716.
- POMERANZ, J. R. (1983). Global and local precedence: Selective attention in form and motion perception. *Journal of Experimental Psychology: General*, *112*, 516-540.

- SCHWARZ, W., & ISCHEBECK, A. (1994). Coactivation and statistical facilitation in the detection of lines. *Perception*, **23**, 157-168.
- SIMON, J. R., & SUDALAIMUTHU, P. (1979). Effects of S-R mapping and the response modality on performance in a Stroop task. *Journal of Experimental Psychology: Human Perception & Performance*, **5**, 176-187.
- SMITH, M. C., BESNER, D., & MIYOSHI, H. (1994). New limits to automaticity: Context modulates semantic priming. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 104-115.
- SMITH, M. C., & MAGEE, L. E. (1980). Tracing the time course of picture-word processing. *Journal of Experimental Psychology: General*, **109**, 373-392.
- SMITH, M. C., THEODOR, L., & FRANKLIN, P. E. (1983). The relationship between contextual facilitation and depth of processing. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **9**, 697-712.
- 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 & 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 & Performance*, **22**, 1166-1177.
- TIPPER, S. P. (1985). The negative priming effect: Inhibitory priming by ignored objects. *Quarterly Journal of Experimental Psychology*, **37A**, 571-590.
- TIPPER, S. P., & DRIVER, J. (1988). Negative priming between pictures and words in a selective attention task: Evidence for semantic processing of ignored stimuli. *Memory & Cognition*, **16**, 64-70.
- ULLMAN, S. (1984). Visual routines. *Cognition*, **18**, 97-159.
- ULLMAN, S. (1989). Aligning pictorial description: An approach to object recognition. *Cognition*, **32**, 193-254.
- ULLMAN, S. (1996). *High-level vision: Object recognition and visual cognition*. Cambridge, MA: MIT Press.
- VANDERWART, M. (1984). Priming by pictures in lexical decision. *Journal of Verbal Learning & Verbal Behavior*, **23**, 67-83.
- VITKOVITCH, M., HUMPHREYS, G. W., & LLOYD-JONES, T. J. (1993). On naming a giraffe a zebra: Picture naming errors across different categories. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 243-259.
- WEISSTEIN, N., & HARRIS, C. S. (1974). Visual detection of line segments: An object-superiority effect. *Science*, **186**, 752-754.

NOTES

1. Given the lack of an agreed-upon metric of the perceptual attributes of the objects selected for our experiment, we had to rest on some assumptions concerning the issue of their canonicalness. One of these assumptions is that the population of the stimuli from which we drew the sample used in our study was composed of objects depicted in their canonical perspective. The assumption might be deemed as directly supported by the rating values available, regretfully, only for the Snodgrass and Vanderwart (1980) set of pictures (see image-agreement norms). Our view is that this assumption should reasonably hold also for the objects extracted from the British Picture Vocabulary Scale.

2. Note that the aim of the graphical distortion was not to completely reverse the extension of the major and minor axes of the objects' elongation. Had we kept constant the major/minor axis ratios between canonical and distorted objects, most of the meaningful stimuli used in the experiment would have resulted in nonidentifiable objects.

3. An additional analysis carried out on the distribution of the errors across the three levels of the RT tertile factor indicated that the mean percentage of errors increased from 4% to 6% to 21% as RTs became longer [$F(1(2,54)) = 68.4$, $MS_e = .004$, $p < .001$]. Although obviously inconsistent with a speed-accuracy tradeoff interpretation of the finding plotted in Figure 3, this result is to be taken with some caution since the three cells of the design for this additional analysis could be generated only by including incorrect responses.

APPENDIX

Item	Rc	Rd	Familiarity
Canonical Vertical Elongation			
KANGAROO	1.50	2.05	3.38
SEA HORSE	2.07	1.81	2.95
FLAMINGO	1.97	1.37	3.75
GIRAFFE	1.48	1.71	3.91
GORILLA	1.66	2.00	4.05
AWL	2.00	1.47	4.75
PENGUIN	2.22	1.62	4.65
OSTRICH	1.23	1.88	2.85
BELL	1.22	2.46	3.90
CANDLESTICK	2.46	1.62	3.55
CARAFE	1.42	1.90	4.90
LAMP	2.06	1.76	2.10
PINT MUG	1.28	2.07	5.85
SALOPETTE	2.72	2.29	5.35
LADDER	1.59	1.66	5.55
TRAFFIC LIGHT	1.20	2.42	6.90
Canonical Horizontal Elongation			
DUCK	1.05	2.28	6.20
WHALE	2.68	1.82	3.40
CAMEL	1.45	1.79	4.00
SWAN	1.40	1.61	5.30
LION	1.95	1.89	4.75
PORCUPINE	1.57	1.54	3.50
RHINOCEROS	1.74	1.71	3.90
TURTLE	2.00	1.72	5.70
CANNON	2.07	1.43	1.85
WHEEL BARROW	1.74	1.44	3.80
BED	2.30	1.85	6.95
VAN	1.70	1.60	4.05
SHOE	1.55	2.35	6.75
SAW	2.75	1.57	3.95
TRICYCLE	1.04	2.26	3.90
CASE	1.21	2.38	6.50

Note—Rc, major/minor axis ratio for canonical targets ($M = 1.76$, $SD = .484$). Rd, major/minor axis ratio for distorted targets ($M = 1.85$, $SD = .317$).

(Manuscript received May 9, 1997;
revision accepted for publication December 22, 1997.)