Efficient visual search by category: Specifying the features that mark the difference between artifacts and animals in preattentive vision

DANIEL T. LEVIN and YUKARI TAKARAE
Kent State University, Kent, Ohio

ANDREW G. MINER
University of Illinois at Urbana-Champaign, Urbana, Illinois

and

FRANK KEIL
Yale University, New Haven, Connecticut

In this report, we explored the features that support visual search for broadly inclusive natural categories. We used a paradigm in which subjects searched for a randomly selected target from one category (e.g., one of 32 line drawings of artifacts or animals in displays ranging from three to nine items) among a mixed set of distractors from the other. We found that search was surprisingly fast. Target-present slopes for animal targets among artifacts ranged from 10.8 to 16.0 msec/item, and slopes for artifact targets ranged from 5.5 to 6.2 msec/item. Experiments 2–5 tested factors that affect both the speed of the search and the search asymmetry favoring detection of artifacts among animals. They converge on the conclusion that target–distractor differences in global contour shape (e.g., rectilinearity/curvilinearity) and visual typicality of parts and form facilitate search by category. We argue that existing theories are helpful in understanding these findings but that they need to be supplemented to account for the specific features that specify categories and to account for subjects' ability to quickly locate targets representing heterogeneous and formally complex categories.

In most visual search experiments, the subject attempts to detect a target that is characterized on every trial by the presence or absence of some single feature or conjunction of features. For example, the task might be to detect a red X among a set of black Xs and red Os. In general, targets are almost always single objects or small sets of objects that share some basic characteristic. Far fewer studies have explored the complex features that characterize natural object categories. In the experiments reported here, subjects searched for targets specified only by their membership in a broad natural category and hidden among a mixed set of distractors that also had in common only membership in a different broad category. Specifically, in one block of trials, subjects searched for 1 of 32 animals in a mixed set of artifacts, and in the other they searched for 1 of 32 artifacts among a mixed set of animals. It is important to note that, on any given trial, the subjects did not know the specific target—they only knew that it would be some animal or some artifact. Thus, no

single instance of a given feature defined either the targets or the distractors. Instead, whatever the visual features that distinguish these categories, they were variously instantiated and not necessarily present in all the display items.

We are interested in two issues. First, given evidence that distinctions between artifacts and natural kinds lay the foundation for learning more specific categories (Keil, 1989), it is plausible that visual search findings may reveal an efficient process for distinguishing these categories perceptually. Although the parameters of the task, considered in the context of current theories of visual search, might lead one to predict a slow and difficult search, the foundational nature of these well-learned categories may allow a much faster search than otherwise might be expected. Second, we will attempt to specify at least some of the features that affect search slopes in this task. Although no simple feature or property will consistently distinguish targets from distractors, there are a number of global perceptual differences between artifacts and animals that could facilitate search if present in a given category exemplar. This is particularly interesting in the context of an efficient search. In such a case, the features driving the search should be available preattentively,[1] and by understanding them we can gain insight into the degree of complexity available in preattentive vision. Although a long tradition of research has asked this

question for simple abstract stimuli, no findings we know of have tested what features inherent to natural object categories affect search slopes. The feature we explored first was the relatively rectilinear shape of artifact contours, as compared with the relatively curvilinear shape of animal contours. In addition, a number of other factors were tested for their ability to predict search slopes including complexity, ambiguity, visual typicality, and similarity between the target object and the distractor categories.

Visual Search by Category

Research using a wide variety of paradigms has shown that humans are capable of efficiently perceiving and remembering natural objects. Early experiments on natural scene perception concentrated on the degree to which a large number of individual scenes could be remembered. Most of these studies emphasized our apparently limitless ability to discriminate familiar scenes from novel ones (Nickerson, 1968; Shepard, 1967; Standing, Conezio, & Haber, 1970), whereas later work focused on the amount of time necessary to encode a scene, for both on-line identification and later recognition. Several studies found that brief 100- to 200-msec exposures were sufficient to identify objects and scenes, but not to remember them (Intraub, 1981; Potter, 1976; Potter & Faulconer, 1975).

Other research using a visual search paradigm has shown that a specific object can be located relatively quickly. For example, Biederman, Blickle, Teitelbaum, and Klatsky (1988) found that subjects performed above chance at detecting a named target in arrays presented for 100 msec. However, they also found that increases in display size reduced accuracy, prompting them to suggest that object search was not capacity free. More recently, Zelinsky, Rao, Hayhoe, and Ballard (1997) had subjects locate a color image of a target object in a pseudorealistic setting (e.g., a stuffed animal among other toys in a crib). On each trial, the target image was shown in isolation; then the subjects indicated whether it was present on the table. An analysis of eye movements suggested that the first saccade after fixation was directed toward the middle of the display (which contained no objects, because the stimuli were arranged in a semicircle). The second saccade then proceeded in the general direction of the target, followed by further saccades that were directed with increasing precision at the target. Therefore, although search reaction times (RTs) indicated a 28-msec/item search rate, the subjects' eye movements revealed an ability to determine which part of the screen included the target without the benefit of a serial scan of each object in turn.

Although this rapid target identification suggests some type of preattentive coding, the data reviewed above are based on identifying particular targets and therefore could plausibly be driven by simple perceptual cues. For example, the rapid detection in Zelinsky et al.'s (1997) research could be based on strategic selection of distinctive perceptual features, such as object color and basic aspects of shape. Subjects asked to detect a screwdriver

might search for the particular shape and orientation of its shaft, thus doing the search on the basis of the kind of simple perceptual feature that is traditionally thought to drive parallel search. The same could be said about the scene classification research, which typically uses basic-level target descriptions, if not actual images of the target to be detected. However, in Intraub's (1981) rapid identification task, subjects successfully responded on the basis of the nonpresence of a category of object (e.g., they might be instructed to "look for a picture that is not of house furnishings and decorations," p. 608). This does suggest a more sophisticated process, because no particular target was selected. Indeed, this sophistication does find support in a few studies that have assessed the degree to which subjects can detect targets solely on the basis of their membership in a category that presumably is marked by no particular perceptual feature.

The best-known example of this kind of visual search is Egeth, Jonides, and Wall (1972), who showed that subjects could locate a number among letters, or could do the reverse, in parallel. Jonides and Gleitman (1972) further showed that an ambiguous "O" stimulus (which could be the number zero or the letter "o") could be located in parallel among both numbers and letters. Thus, they argued that their findings reflected a truly categorical visual search that could not be explained solely on the basis of the direct association between a given shape and a given response. However, a number of attempts to replicate this effect have failed (e.g., Duncan, 1983; Krueger, 1984), so it must be considered to be tentative at best.

Although the status of efficient alphanumeric category search is questionable, more recent data from experiments using broad natural categories reinforce the possibility of efficient categorical visual search. Thorpe, Fize, and Marlot (1996) tested subjects' ability to locate an animal in a series of natural scenes and found that RTs were as short as 400 msec, implying that the animals had actually been perceived in less than 200 msec. Evoked potentials confirmed this by showing a divergence in brain activity for scenes with and without animals as soon as 150 msec after stimulus onset (Thorpe et al., 1996). If we assume that their images contained a sufficient variety of animals so that they were not all marked by the presence of a particular part (such as legs), these data suggest that complex natural scenes can be rapidly parsed by category.

This finding is particularly interesting because it may reflect the foundational nature of Thorpe et al.'s (1996) stimulus categories. The contrast between artifacts and animals may, in fact, represent one of the first ways that children divide the world and may serve throughout adulthood as a conceptual framework for more specific learning (Keil, 1989). For example, Gutheil, Vera, and Keil (1998) taught young children a new property for a given animal and then asked them whether it would be present in other objects. The children readily attributed the new properties to a wide variety of living things, ranging from dogs to worms, yet refused to generalize it to nonliving objects. In addition, young children understand a num-

ber of entailments on the basis of the animal/artifact distinction. They know that living things grow as they age, whereas artifacts do not (Rosengren, Gelman, Kalish, & McCormick, 1991), and that animals can move by themselves, whereas artifacts are less likely to do so (Massey & Gelman, 1988). Furthermore, precursors of a distinct understanding of the living-kind/artifact difference can be seen in infancy. For example, Woodward (1998) found that 6-month-old infants distinguish between the goal-directed actions of a hand and the more automatic actions of a stick, and Bertenthal (1993) showed that infants distinguish between biological and nonbiological motion.

The assumption that the artifact/animal distinction serves a cognitive foundation is also reinforced by neuroanatomical findings. It appears that a subset of brain injuries can selectively impair recognition of one category or another. For example, Farah, McMullen, and Meyer (1991) described two visual agnosics who had difficulty recognizing living things, as compared with artifacts, a deficit that could not be explained on the basis of simple differences between the categories, such as image complexity, familiarity, and so forth. Sacchett and Humphreys (1992) reported the opposite deficit—a selective inability to name artifacts. Although the specific interpretation of this deficit has been controversial (Funnell & Sheridan, 1992; Gaffan & Heywood, 1993; Stewart, Parkin, & Hunkin, 1992), most authors agree that these data reflect some kind of specialization by category or by basic between-category processing differences, and more recently, Caramazza and Shelton (1998) argued that they may reflect a set of distinct visual subsystems that correspond closely to cognitive between-category distinctions. If this is true, efficient visual search for these categories may involve detection of activity in one of these systems.

On the whole, then, the research converges to suggest that scenes can be rapidly parsed by category, even by broadly inclusive categories, such as artifacts and animals. Both visual search and rapid serial visual presentation tasks reveal efficient categorical parsing of scenes at a number of levels. This efficiency may not be limited to well-defined or narrowly defined categories and may include broad, foundational categories, such as artifacts and natural kinds. If an efficient search by category is possible, it is important to understand what specific features drive the search. Understanding the features that drive an efficient search will provide information about the representations available in preattentive vision. Previous research testing object recognition and identification in normal and brain-injured subjects suggests that a wide range of potential features could distinguish these categories.

What Features Distinguish
Artifacts and Animals?

In the experiments reported here, a visual search task was based on natural object categories that are associated with a rich set of concepts and different inferential patterns. The contrast between the categories is also as-

sociated with a variety of systematic perceptual differences, ranging from differences in the importance of part configuration to differences in basic contour shapes. A number of authors have suggested that living things are processed by a system that makes more extensive use of part configuration as a differentiating cue. Farah (1995) argued that face-specific deficits are the most dramatic examples of this kind of configural processing deficit and that those deficits for recognizing living things that often accompany face-specific deficits are similar in this respect. Artifacts and such objects as alphanumeric characters are processed by using more information about parts in isolation from their configuration.

Another possibility is that visual information is globally more central to identifying living things (Warrington & Shallice, 1984). Warrington and her colleagues proposed that living things are mainly defined by their sensory features and that artifacts are mainly defined by their functional features. This proposal implies that perceptual differences between animals and artifacts lie in how we interact with them. A number of studies support this view. First, findings in perception suggest that artifacts are coded explicitly in terms of body-scaled action affordances. For example, Warren (1984) found that an estimate of the degree to which a step is climbable is closely tuned to optimal energy efficiency for the subject's particular size. Second, recent neuroimaging data have suggested that images of artifacts simultaneously activate primary visual areas and an area associated with generating action terms, whereas images of animals do not activate this second area (Martin, Haxby, Lalonde, Wiggs, & Ungerleider, 1995; Martin, Wiggs, Ungerleider, & Haxby, 1996).

A more low-level perceptual distinction between artifacts and natural kinds is that they have different contour shapes. For example, Zusne (1975) showed subjects random shapes that were either polygons with straight lines or similar closed forms with curved lines. He found that the subjects generated more associations with living things for the curved shapes. More recently, Kurbat (1997) measured the average local curvature on the image contours of artifact and natural kinds and found that artifacts were, on average, characterized by straight edges and natural kinds by more curved edges. He also found that curvilinearity was a significant predictor of whether an agnosic patient would recognize a given object. Other research has suggested that infants use contour shapes to differentiate categories. Van De Walle and Hoerger (1996) tested this hypothesis in infants, using a stimulus set of toys that had global part arrangements characteristic of one category, but with contours characteristic of another category. Thus, one stimulus might be a dog with parts suitably arranged to represent a head, a body, and four legs. However, each of these parts was made with straight edges and corners, giving the toy contours characteristic of an artifact. Infants were shown a series of toys representing one category, habituated to them, and then were given a test toy. As in previous research (Mandler & McDonough, 1993), the infants dishabituated to the new toy if it rep-

resented a different category from the toys on the habituation series. In addition, however, Van De Walle and Hoerger found that the infants dishabituated to a test toy from the same category as the habituated toys so long as it had contours characteristic of the nonhabituated category.

In summary, it is possible that the visual system uses specialized processes for dealing with artifacts and natural kinds. The question is, what features does it use to distinguish these categories during the first stages of processing? Different proposals range from action affordances to contour shape. Here, we used a visual search paradigm to test feature contrasts at two basic levels: a complex level that required relatively complete structural descriptions of form versus a more simple perceptual level based on contour shape (e.g., curvilinearity/ rectilinearity). In addition, we tested the degree to which the visual typicality and similarity between the target object and the distracting category affected search slopes. It is important to point out that we are not committed to a top-down model of object perception or search, whereby semantic-level categorizations affect search slopes. Instead, the effectiveness of a categorical contrast between targets and distractors is assumed to lie in the way that long-term learning (or genetics) has caused early and mid-level vision to reflect natural breakpoints in the structure of the visual world.

The Present Study

With the experiments reported in this paper, we both assessed the overall speed of visual search by category and, in addition, attempted to understand the features that distinguish artifacts from natural kinds. Although these two goals are somewhat different, they converge in that a fast search would suggest a feature contrast that is deeply embedded in preattentive vision. Experiment 1 was a basic visual search experiment in which subjects searched for 1 of 32 targets of one category among a mixed set of distractors chosen from a set of 32 objects of the other category. Experiment 2 tested the degree to which complete structural descriptions drive the search by having subjects search for jumbled objects of one category among jumbled objects of the other. In Experiments 3 and 4, a possible low-level perceptual feature was tested. Objects of both categories were assessed for the degree to which their contours were rectilinear. Search slopes for each object were then collected and correlated with the rectilinearity measure, with the prediction that highly rectilinear objects would be more quickly detected. Finally, in Experiment 5, we completed a series of regression analyses predicting the slopes observed in Experiments 3 and 4, using a variety of variables, including complexity, ambiguity, contour shape, and typicality.

EXPERIMENT 1

Experiment 1 was a basic visual search experiment in which the search target was defined by category. Subjects therefore searched for 1 of 32 different artifacts among a mixed set of animal distractors, or did the reverse, searching for 1 of 32 animals among a mixed set of artifacts. The basic issue of interest here was the efficiency of visual search by category, as reflected by search slopes.

Before proceeding, a note on the stimulus sets used in this paper is in order. We chose to test the contrast between animals and artifacts as the search categories despite the fact that a more broad classification scheme might subsume the animal category within all living things, or even all natural kinds. Despite the psychological salience of all of these levels of classification, we chose to use animals because it simplified development of the stimuli (because it allowed us to avoid a number of complexities, such as the question of natural kinds that are, in some ways, similar to artifacts, such as fruits and vegetables). In doing so, we follow Caramazza and Shelton (1998), Martin et al. (1996), and Thorpe et al. (1996) in assuming that animals represent a psychologically salient and potentially dissociable category.

Method

Subjects. Eight Cornell University undergraduates participated in the present experiment in exchange for course credit.

Stimuli. The stimulus set consisted of 64 realistic black-and-white line drawings of artifacts and animals (32 of each; Figures 1A–1B). The drawings were shaded internally with a uniform light gray. Some of the stimuli were from the original Snodgrass and Vanderwart (1980) set, whereas others were culled from various collections of clip art. They were selected from a larger set pseudorandomly, with the constraint that they had to be reasonably unambiguous (this was most important for the artifact pictures). In addition, some animals without legs protruding from the bottom were purposely selected (e.g., the butterfly, the fish, and the seahorse) to minimize the degree to which that particular feature was characteristic of the whole set. Once selected, the two sets of 32 images were equalized for average pixel area to prevent size from becoming a cue to reinforce the search. The size of the stimuli varied somewhat but averaged approximately 100 (h) $\times$ 80 (v) pixels (approximately 5 $\times$ 4 cm).

Apparatus. Search trials were presented in 256-level gray scale on Macintosh LC computers with 12-in. monitors. The subjects sat in a comfortable unconstrained position, with an average viewing distance of approximately 60 cm, and completed the experiment in small groups ranging in size from 1 to 5, using different computers in the same room.

Procedure. The experiment consisted of two blocks of 144 trials, each preceded by 12 practice trials. In one block the subjects searched for an animal among artifacts, and in the other they searched for an artifact among animals. Presentation order of blocks was counterbalanced across subjects. The subjects were instructed that they would be seeing a series of displays containing a number of objects and that their job was to hit the "1" key if an object from the target category was present in the display and to hit the "2" key if none was present. Thus, for each trial, no specific object was named as a target. Instead, the subjects had to detect any of the animals among artifacts, or do the reverse. The subjects were instructed to use two fingers on their dominant hand to respond and to respond as quickly as possible without sacrificing accuracy. The subjects were also informed that they would receive feedback after each trial in the form of a "+" for correct responses and a "−" for incorrect responses.

On each trial, the subjects saw a display containing three, six, or nine objects. The objects were presented in randomly selected locations within a 3 $\times$ 3 matrix on the screen. The approximate visual angle subtended by each object was 4.7º $\times$ 3.8º, and the approximate angle of the entire display was 15.5º $\times$ 12.5º. On half of the trials an object from the target category was present, and on half it
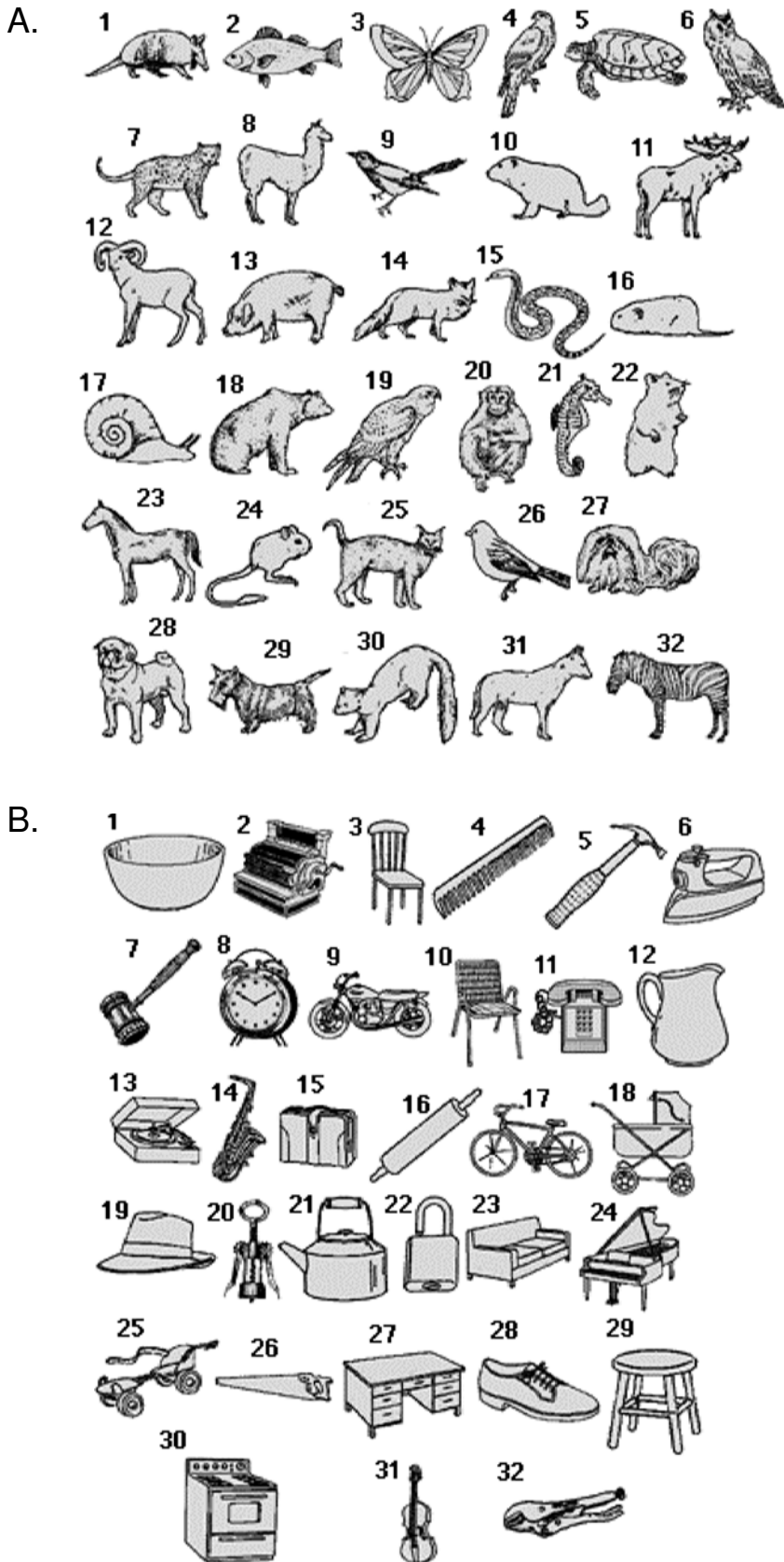
Figure 1. (A) Animal stimulus set. (B) Artifact stimulus set.

was absent. There were 24 trials in each cell of the design, for a total of 288 trials (24 trials × 2 tasks [animal target /artifact target] × 3 display sizes [3,6,9] × 2 target conditions [target-present /target-absent]). Each trial was response terminated and was immediately followed by feedback, which remained on the screen for 350 msec. Targets for each trial were chosen randomly from among the 32 members of each category, and distractors were chosen randomly with replacement.[2]

Results

All the analyses reported in this paper have focused on target-present data, although target-absent data are reported and graphed for the sake of completeness. Although target-absent data are sometimes informative, they tend to reflect decision stage processes and search termination criteria (see Chun & Wolfe, 1996) that are not of central interest for the purposes of this report.

Errors. The average error rate was 1.95%. The mean error rate for target-present trials was 2.86%, and the mean error rate for target-absent trials was 1.04%. The error rates for target-present trials were entered into a two-factor (task × display size) within-subjects analysis of variance (ANOVA). Neither the effects of these factors nor the interaction between them was significant, although the error rate for artifact target trials was 3.65% and the error rate for animal target trials was 2.08% [$F(1,7) = 4.200$, $MS_e = 6.975$, $p = .0796$]. The error rates across display sizes were 4.17%, 3.12%, and 3.65% for three-, six-, and nine-item artifact target displays and 1.04%, 1.04%, and 4.17% for three-, six-, and nine-item animal target displays, respectively.

Search reaction times. Prior to analysis, the RT data from error trials were eliminated, here and throughout this report. The RT data were entered into a two-factor (task × display size) within-subjects ANOVA. In addition, average search slopes were computed for each condition.

The analysis revealed two main findings. First, search was highly efficient. The average target-present slope for detecting an artifact among animals was 5.5 msec/item (target-absent slope = 27 msec/item), whereas the target-present slope for animals among artifacts was 16 msec/item (target-absent slope = 41.5 msec/item). Second, a near-significant search asymmetry was revealed in the task × display size interaction [$F(2,14) = 2.58$, $MS_e = 1,776$, $p = .1114$], in which the search slope for artifacts was shallower than that for animals (see Figure 2).

One additional aspect of the data worth mentioning is that there is a crossover in RTs for target-present three-item displays such that RTs were significantly faster when detecting an animal among artifacts (578 msec) than when detecting an artifact among animals (624 msec, $p < .01$, Duncan test).

Discussion

On the whole, search slopes for target-present trials were surprisingly shallow, and the artifact target search fell within the traditional cutoff of 10 msec/item for a putatively parallel search (Theeuwes, 1993). A number of factors would lead to the prediction of a slow, ineffi-
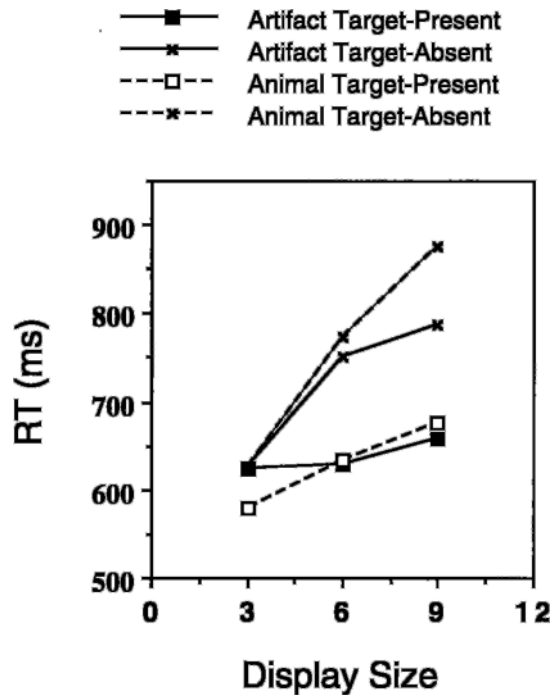

Figure 2. Visual search reaction times (RTs) for animal targets and artifact targets.

cient search. First, no specific perceptual primitive appears to have been sufficient to mark the target. According to the original feature integration theory (FIT; Treisman & Gelade, 1980), parallel search is possible only when a single, discriminable perceptual primitive distinguishes targets from distractors. Although some perceptual feature might at least partially have distinguished the categories, it would have been nothing like the primitives previously studied, and in addition, it would have been unlikely to have reliably separated all the members of the target category from the nontarget category, because of the heterogeneity of our stimuli. Other variations and revisions of FIT also seem unable to explain why the search was so fast. For example, Wolfe's (Wolfe, 1994; Wolfe, Cave, & Franzel, 1989) guided search model is more flexible in that it allows parallel searches for targets defined by a conjunction of features, but again it is difficult to point to a pair or a triplet of features that would have reliably distinguished targets from distractors. Furthermore, even if some conjunction of primitives were to have distinguished targets from distractors, it would have been unlikely that the low-noise conditions specified in guided search would have been met here (e.g., that activations on each feature map strongly distinguished target locations from distractor locations).

The Duncan and Humphreys (1989) model of visual search also appears to predict a very slow search. In particular, one of the major factors affecting search speed, according to Duncan and Humphreys, is the degree of heterogeneity among distractors. Displays with highly

varied distractors will be difficult to search, whereas displays with very similar distractors will be easy to search. The distractors used in the present study were highly heterogeneous, yet the targets were located rapidly. These sets of objects shared no diagnostic color, orientation, part, or part structure. Although the distractors were probably similar in some respects, these respects were isolated among a large number of perceptually available features. Thus, there was no a priori means of selecting a set of features on which to base an estimate of homogeneity. In this context, the notion of distractor homogeneity becomes insufficient to explain why visual search might have been fast or slow in this case. (However, the model may provide some explanation for the direction of asymmetry. Animal distractors appear more homogeneous than artifacts because of common parts, such as legs, which might facilitate search for artifact targets. This possibility will be examined further in Experiments 3 and 4.) It is therefore critical to specify the features that distinguish the particular categories, if we are to understand search behavior. Experiments 2 and 3 were designed to test whether search is based on higher level features and/or more simple features, such as individual parts or curvilinear/rectilinear contour shape. To examine the effect of higher order information, the targets and distractors used in Experiment 2 were jumbled. Jumbling was intended to disrupt structural coding and therefore interfere with semantic processing.

Before going on to Experiment 2, a note about target-absent search slopes is in order. In Experiment 1, the ratio of target-present to target-absent slopes was 1:4.91 for the artifact targets and 1:2.59 for the animal targets. Thus, for the artifact targets, target-absent search was quite slow relative to target-present search. It seems likely that this occurred because artifact search is easier than subjects think it should be, making them more hesitant to terminate the search than they should be. Certainly, this may represent an interesting effect, perhaps indicating that this search task is considerably easier than subjects expect, but in the interest of focus we will defer exploration of these differences for a later date and will continue to analyze only target-present slopes in the remainder of this paper.

EXPERIMENT 2

If efficient visual search is being driven by a complete structural code, jumbling the features of the objects should markedly slow the search. Experiment 2 was therefore identical to Experiment 1, with the exception that all the stimuli were jumbled by trading locations and orientations of most of the major parts within each object. The subjects searched for a jumbled artifact among jumbled animals or did the reverse, searching for a jumbled animal among jumbled artifacts.

Method

Subjects. Fourteen Cornell University undergraduates completed the present experiment in exchange for course credit. Of these, 1 was dropped from the analysis, owing to a high error rate (16%).

Stimuli. The stimulus set consisted of jumbled versions of the 64 images used in Experiment 1. Images were jumbled by using a standard image-processing program (Adobe Photoshop) to cut, rotate, and paste sections of each image. In general, most object parts were cut and pasted as wholes to avoid creating unnatural line ter-
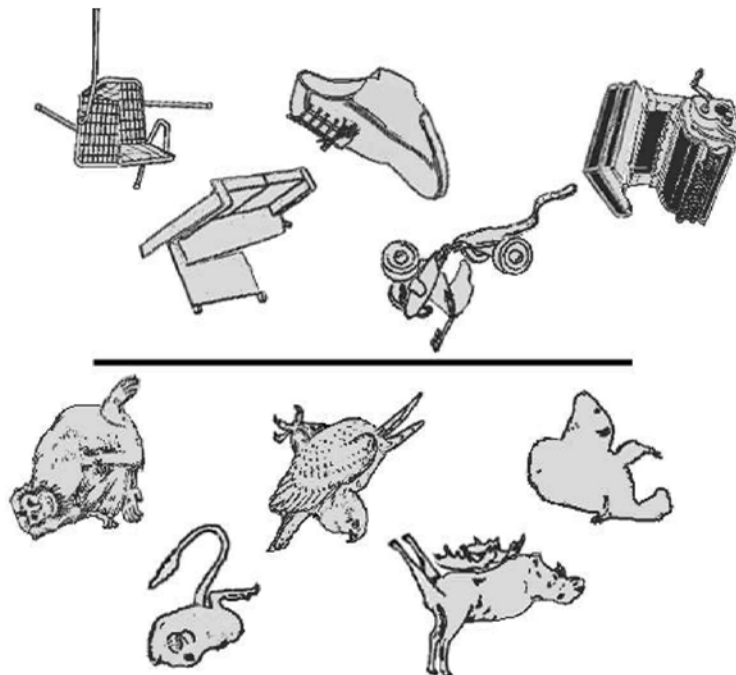

Figure 3. Sample jumbled objects.

minations or radical changes in contour direction (see Figure 3). However, this constraint was not possible to satisfy in all cases.

Apparatus and Procedure. The apparatus and procedure were identical to those in Experiment 1.

Results

Errors. The average error rate was 3.15%. The mean error rate for target-present trials was 5.12%, and the mean error rate for target-absent trials was 1.18%. The error rates for target-present trials were entered into a two-factor (task × display size) within-subjects ANOVA. The task effect and the task × display size interactions were nonsignificant ($Fs < 0.1$), whereas the display size effect was significant [$F(2,24) = 4.500, MS_e = 40.509, p = .0219$]. The error rates across display sizes were 4.17%, 2.56%, and 6.73% for three-, six-, and nine-item artifact target displays and 3.52%, 4.17%, and 9.62% for three-, six-, and nine-item animal displays, respectively.

Search reaction times. The data were entered into a within-subjects two-factor task (artifact target or animal target) × display size (three, six, or nine) ANOVA. An additional ANOVA included a between-subjects configuration (normal/jumbled) factor that allowed comparison between Experiments 1 and 2.

The primary finding in this experiment is plainly visible in Figure 4. Jumbling features had little effect on search slopes, which remained quite shallow. The mean slope for artifact target trials was 12.3 msec/item, as com-

pared with 5.5 msec/item in Experiment 1 (66.8 msec/item on target-absent trials), and the slope for animal target trials was 17.1 msec/item, as compared with 16.0 msec/item in Experiment 1 (75.1 msec/item on target-absent trials). The search asymmetry was not significant ($F < 1$). Although this represents a slowing of approximately 7 msec/item in the artifact target trials, as compared with Experiment 1, the between-experiments configuration × display size interaction did not approach significance, even when only RTs for artifact target trials were included in the analysis ($Fs < 1$).

Although search slopes were similar between Experiments 1 and 2, average RTs increased by 108 msec [$F(1,19) = 10.934, MS_e = 31,399, p = .0037$].

Finally, mean target-present RTs in three-item displays once again crossed over, with animals (691 msec) being detected more quickly than artifacts (702 msec), although the comparison was nonsignificant.

Discussion

The major finding in this experiment was that jumbling features slowed the search only minimally. Although the slowdown in the search slopes for artifacts suggests some interference with the search-relevant features, the small size of the effect points away from the possibility that the search is based on a completed structural code. This is, of course, compatible with many theories of visual search and preattentive vision. In particular, recent work has suggested that efficient search cannot be based on a completed structural description. Wolfe and Bennett (1997) used a variety of configural features (i.e., features defined by the relationship between at least two parts) and showed that, in all cases, search was relatively slow. As such, Wolfe and Bennett argued that objects in preattentive vision are "shapeless bundles" of features. Similarly, Rensink (2000) has argued that preattentive representations of objects are loosely organized and transitory. It is important to note, however, that these representations are unlikely to be limited to unprocessed primitives, such as line segments and colors. As has been noted by both Rensink (2000) and Wolfe and Bennett (1997), parallel search is possible on the basis of a number of relatively complex features, including shape from shading (Ramachandran, 1988) and three-dimensional (3-D) orientation (Enns & Rensink, 1991). In addition, Donnelly, Humphreys, and Riddoch (1991) found that contour closure could support efficient search, implying that some limited processing of form occurs prior to the arrival of focal attention. He and Nakayama (1992) went farther and suggested not only that preattentive vision accounts for depth relations, but that visual search cannot access the prior feature-based representations at all.

Clearly, therefore, Experiment 2 does not completely eliminate the possibility that complex perceptual features drive visual search by category. For example, it is possible that semantic information can be accessed on the basis of individual parts. It is important to note that even if the features that distinguish these categories are exclusively perceptual, the idea that they instantiate a
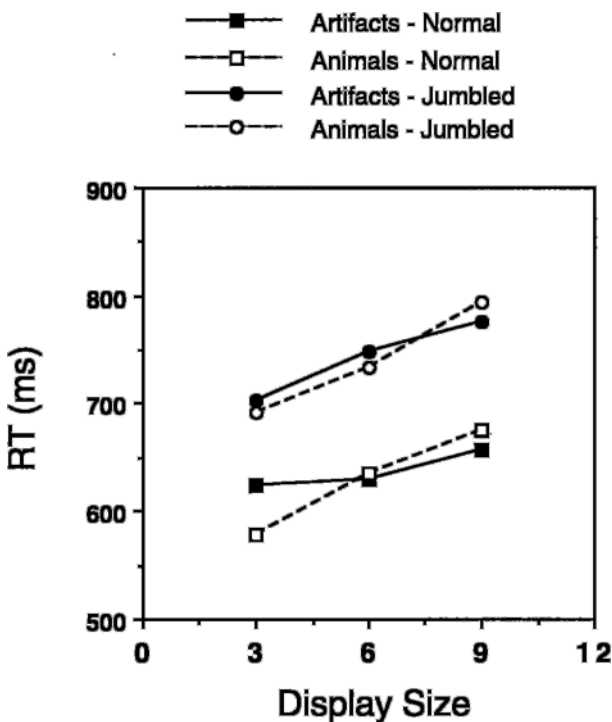

Figure 4. Combined results from target-present trials in Experiments 1 and 2. Boxes represent Experiment 1 (visual search for nonjumbled artifacts and animals), and circles represent Experiment 2 (visual search for jumbled artifacts and animals).

substantive category remains plausible, given findings of category-specific visual impairments reviewed above. However, given that Experiment 2 generally pointed away from completed structural descriptions, Experiments 3 and 4 investigated more simple features by testing for correlations between rectilinearity and visual search slopes, whereas Experiment 5 tested the predictive power of other factors that might be available on the basis of visual information less complex than structural descriptions, but more complex than primitives.

EXPERIMENT 3

To test the hypothesis that artifact rectilinearity facilitates visual search by category, it was necessary to devise some way of measuring the global shape of the object contours. We therefore measured rectilinearity by using a standard edge detection algorithm that allows detection of pixel locations characterized by an image edge and measures the orientation of that edge (Sobel method; see Jain, Kasturi, & Schunck, 1995). Given edge orientations at each edge location, it was possible to obtain a measure of rectilinearity by simply computing the degree to which edges were collinear with each other. This measure was taken on each of the 64 artifact and animal images. Individual rectilinearity measures were then correlated with behavioral data from a new experiment in which subjects completed a sufficient number of trials to allow an estimate of search slopes for each stimulus. The search task was identical to that in Experiment 1.

Method

Subjects. Six Cornell undergraduate and graduate students completed the present experiment. These included two of the authors, D.T.L. and A.G.M.

Apparatus. The experiment was run on computers with 14-in. (image size, 26.9 cm horizontal × 20.2 cm vertical) 256-level grayscale displays set at a resolution of 640 × 480. The stimuli were therefore somewhat larger than those in Experiments 1 and 2 and subtended a maximum of 5.2° × 4.3°, whereas the entire display subtended approximately 20.1° × 15.8°.

Procedure. The subjects completed a search experiment similar to that in Experiment 1. The major difference was that each stimulus used in Experiment 1 served as a target a total of 16 times, 8 times in a three-object display and 8 times in a nine-object display. The subjects therefore completed a total of 2,048 trials (64 stimuli × 2 display sizes × 8 trials per display size × 2 target conditions). To avoid fatigue, the experiment was broken up into four sessions, each of which lasted approximately 20 min. The 512 trials in each session included two blocks of 256 trials, one for the animal target condition, and one for the artifact target condition. In addition, 8 practice trials preceded the 256 experimental trials in each block. The subjects were given instruction sheets and were allowed to complete the four sessions at their convenience by going to the laboratory unsupervised over a period of 2–3 days. Many of the subjects completed two sessions in a single sitting, but all were advised not to attempt more.

Contour analysis. Several measures of edge relatedness were computed to measure differences in contour shape between pictures of artifacts and animals. All of them use an edge detection algorithm that assesses the degree to which each pixel's surround suggests an edge. In the present experiment, this was done by measuring the slope of change in luminance values over 5 × 5 pixel regions of space. The orientation of the edge was determined by taking derivatives in two orthogonal directions (different by 90°) and combining them by taking the arcsine of their ratio (Sobel method; see Jain et al., 1995). Once all the pixels in the image were assigned an *edge value* (high if the region contained information suggesting an edge) and an orientation, a threshold was applied to the edge values, which selected a subset of pixels to be classified as edges. Orientations of all pairs of neighboring edges were then subtracted to produce a measure of the deviation between edge orientations. Throughout this report, this basic measure will be referred to as *edge pair deviation* (EPD). The basic assumption is that pictures with a relatively high percentage of edge pairs that are either collinear or at right angles are rectilinear, whereas pictures with few such pairs are curvilinear.

A key issue is whether to base computations on the entire edge map or only on the map of external edges. Following Kurbat (1997) and Tranel, Logan, Frank, and Damasio (1997), we used only the external edges. The primary benefit of using the whole edge map is that it includes more information and edges that are likely to be used by the subjects to identify the object. The major disadvantage of using the whole map is that it includes internal edges that are parts of textures and small surface markings in addition to major contours (a problem that is especially acute in the line drawings). Although texture and surface markings could plausibly be important factors in driving this search task, it is extremely difficult to reliably distinguish these from contours that might match those used by the visual system (Sanocki, Bowyer, Heath, & Sarkar, 1996). External contours have no special status as markers of part boundaries, but they are at least continuous and also likely to be a fair match to a subset of contours that are actually used by the visual system to parse scenes and identify objects.

Figure 5 summarizes the basic output of this analysis for a typical animal and a typical artifact. These figures show distributions of EPD values across a range of linear distances from 2 to 40 pixels in steps of two. The horizontal rows in the graph represent different distances, whereas the columns represent different EPD values (with bins ranging from 0°–5° to 85°–90°). The size of the square in each cell of the chart represents the relative number of edge pairs within a given range of EPD values and distances. Cell values were computed by taking the ratio of the number of actual edge pairs to the expected value based on the total number of edge pairs in the image divided by the number of cells in the chart. Thus, observations on each cell are normalized for the number of edge pairs in the image, but not for variance across rows or columns.

As can be seen in Figure 5, the animal image is characterized by a gradual increase in the range of EPD values as edge-to-edge distance increases, showing that increased distance is associated with decreased predictability of local contour orientation, a finding that is similar to the analysis reported by Olshausen and Field (1996). However, the artifact image is quite different in that it appears to violate this relationship. Here, increasing distance is associated with no regular increase in EPD values, which reflects the preponderance of straight lines in the image. Also, the artifact in Figure 5 is characterized by a relative increase in the number of edge pairs at a deviation of 85°–90°, which is caused by the presence of right angles in the image. As edge pairs become more distant, they are more likely to include one edge on each side of a corner.

On the basis of the different distributions for animals and artifacts, a direct measure of rectilinearity was computed for each stimulus, based on the percentage of EPD values in the 0°–5°, 5°–10°, and 85°–90° columns. This particular selection was guided by the assumption that rectilinearity is reflected in the preponderance both of aligned edges and of edges at right angles.

Results

Errors. The average error rate was 2.95%. The mean error rate for target-present trials was 4.34%, and the
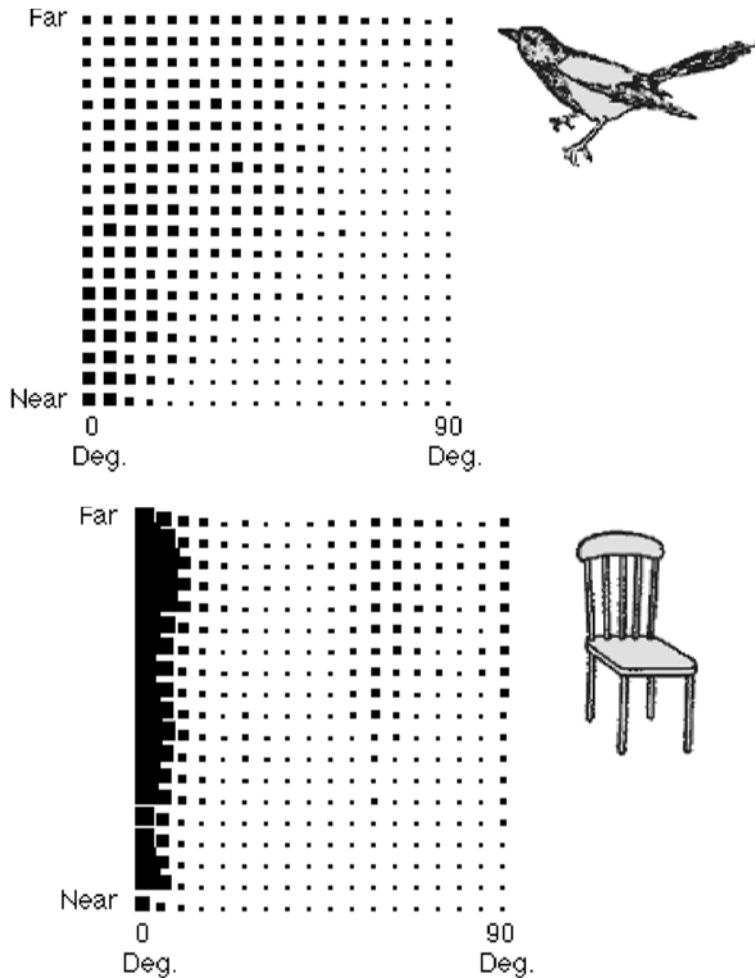
Figure 5. Typical edge pair distributions for one animal and one artifact.

mean error rate for target-absent trials was 1.56%. The error rates for target-present trials were entered into a two-factor (task $\times$ display size) within-subjects ANOVA. The task and display size effects were nearly significant [task effect, $F(1,5) = 5.254$, $MS_e = 1.811$, $p = .0705$; display size effect, $F(1,5) = 6.167$, $MS_e = 6.064$, $p = .0556$], whereas the task $\times$ display size interaction was nonsignificant [$F(1,5) = 1.176$]. The error rates across display sizes were 3.97% and 5.97% for three- and nine-item artifact target displays and 2.21%, and 5.21% for three- and nine-item animal displays, respectively.

Reaction times. Target-present search slopes, averaged over all stimuli, were first entered into a one-factor within-subjects ANOVA, with task (artifact target/animal target) as the single factor. This analysis confirmed that the overall results from the search task replicated those from Experiment 1. The average search slope in the artifact target condition was 5.6 msec/item (target-absent slope = 25.75 msec/item), whereas the average in the animal target condition was 13.7 msec/item (target-absent slope = 31.0 msec/item). Target-present slopes were sig-

nificantly different, replicating the search asymmetry observed in Experiment 1 [$F(1,5) = 70.185$, $MS_e = 2.757$, $p = .0004$]. This asymmetry was also significant in an item-based analysis [$F(1,62) = 35.337$, $MS_e = 18.316$, $p < .0001$]. As in Experiment 1, target-present RTs for three-item displays were faster for animal targets (531 msec) than for artifacts (549 msec), although this effect was nonsignificant [$F(1,5) = 2.697$, $MS_e = 338.5$, $p = .161$]. Target-present RTs for nine-item displays were significantly faster for artifact targets than for animal targets [$F(1,5) = 7.939$, $MS_e = 357.1$, $p = .037$].

An additional analysis assessed learning effects across the four blocks of 512 trials. Mean RTs were entered into a three-factor task (artifact target or animal target) $\times$ display size (three or nine) $\times$ block (one, two, three, or four) within-subjects ANOVA. RTs decreased across blocks (Block 1, 630 msec; Block 2, 569 msec; Block 3, 545 msec; Block 4, 534 msec; $F(1,5) = 15.87$, $MS_e = 2,784$, $p < .0001$]. However, the block $\times$ display size interaction did not approach significance ($F < 1$), indicating no change in search slopes across blocks. The task $\times$ block and task
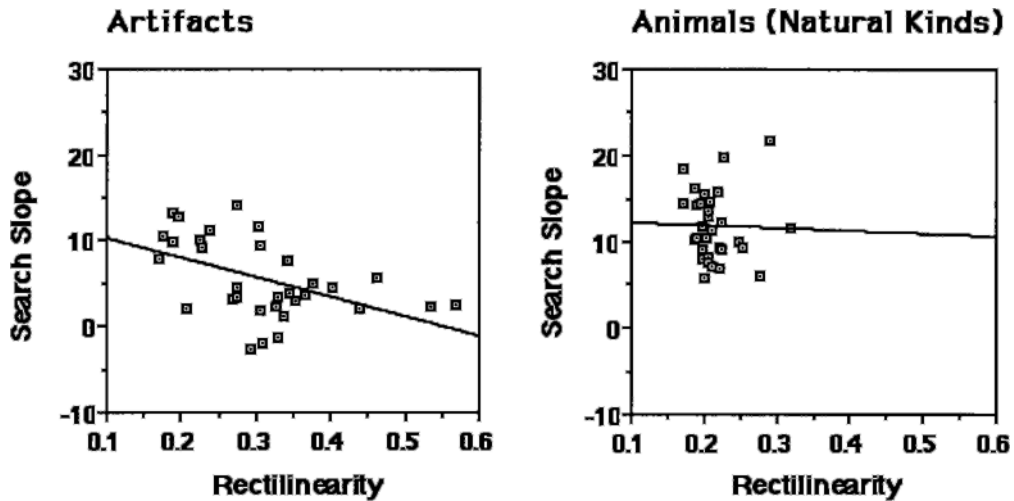
Artifacts

Animals (Natural Kinds)

Figure 6. Correlation between search slopes and rectilinearity in Experiment 3.

× display size × block interactions were nonsignificant (*F*s < 1).

In the second major analysis, search slopes were computed for each stimulus in the experiment and were correlated with rectilinearity. This relationship was significant for artifact slopes ($r = -.48$, $p < .01$), but not for animals ($r = -.026$; Figure 6). As can be seen in Figure 6, one reason for the negative finding in animals appears to be range restriction, in that rectilinearity values varied only slightly in this stimulus set.

Discussion

Experiment 3 makes several things clear. First, it closely replicates Experiment 1, both in terms of the search asymmetry favoring artifacts and in terms of actual search slopes, which were 5.5 and 13.7 msec/item for artifact target and animal target searches, respectively, as compared with 5.6 and 16.0 msec/item in Experiment 1. Second, the experiment reveals a significant correlation between search slopes for individual artifacts and their objectively measured rectilinearity. This relationship is absent for animals.

The most important finding was the correlation between search slopes and rectilinearity for artifact targets. This suggests that rectilinearity could well serve as the feature driving the search asymmetry favoring artifacts. In addition, the pattern of means in Figure 6 suggests that objects with rectilinearity values of approximately .25 are searched for at a normative rate of 12 msec/item, irrespective of category. Artifacts that deviate from this normative contour shape are the ones that drive the asymmetry, because they are easier to locate, although it remains to be seen whether animals with rectilinear contours will also be more easily located among curvilinear distractors. Furthermore, it is possible that different contour populations are represented by early vision in a spatiotopic map of their own. This kind of explanation fits well with re-

search on the connectivity patterns observed in area V1 (Gilbert, 1993).

Although rectilinearity has plausible physiological roots, it might simply be interpreted as an autocorrelation function that is high when edges in one location predict the orientation of edges in another and low when orientations are less correlated. This alternative is, however, very similar to the rectilinearity concept. The major circumstance in which it is different (given that we have isolated a single external contour) is one in which there is a large number of accidentally collinear edges that do not happen to belong to a contour that continues uninterrupted in one direction. A sawtooth edge is one example of this, because it is not straight but distant local edge orientations are likely to be similar nonetheless. At a minimum, visual inspection of the stimuli reveals no such edges in the artifacts, which on the whole do not have external contours that repeat any particular local variation. (Even the handsaw's teeth are too small to be detected by the 5-pixel filters we used.) Thus, a preliminary analysis suggests that rectilinearity is a reasonable way of describing the measure, although other, similar concepts, such as autocorrelation, might be adequate as well.

Although rectilinearity predicts search slopes, it plainly does not explain all of the variance. It is possible that some different measure of rectilinearity would be more valid and successful, but other factors may influence slopes as well. For example, each category has characteristic parts. This is particularly true of animals, many of which have a basically horizontal body and vertical legs. As an internal analysis, we therefore selected 12 animals that seemed to violate the basic horizontal body + vertical legs formula (Animals 2, 3, 4, 6, 15, 17, 19, 20, 21, 22, 26, and 27) to test for differences in search slopes and intercepts. Both search slopes and intercepts for the violators were significantly slower (slope, 14.0 msec; intercept, 511 msec) than those for the other animal stimuli [slope, 10.6 msec,

intercept, 488 msec; slope, $F(1,30) = 6.74, MS_e = 13.313$, $p = .0145$; intercept, $F(1,30) = 4.371, MS_e = 898, p = .0450$]. Therefore, a typical set of parts also seems to facilitate search, at least in animals. This suggests that a collection of features is used to discriminate categories. Since there is no single feature present consistently in animal or artifact categories and animals and artifacts are possibly differentiated along many dimensions, it would be a poor strategy to try to locate a target on the basis of a single feature. The implication that multiple diagnostic features have for theories of visual search will be further examined in Experiment 5 and in the General Discussion section.

However, before concluding that visual search by category can be efficient and that rectilinearity facilitates visual search, we looked more closely at the possibility that stimulus factors biased these experiments. The stimuli used in Experiments 1–3 were simple line drawings of artifacts and animals. Although the drawings were realistic, they still emphasized contours and were mediated by an artist's hand. At a minimum, these factors may introduce unnatural contours that unnaturally facilitate the search. This is particularly relevant for the external contours, which are drawn with heavy lines that may overemphasize form primitives that are not present in natural objects. In addition, it is impossible to know what distortions might be introduced in the process of reproducing the images. For example, it is possible that artists tend to overemphasize the straightness of artifact contours and the curvedness of animal contours, which may accentuate a difference that only minimally marks the difference between actual categories. Another potential interpretive issue is the lack of color or texture information in the drawings, especially given that color and texture information might be more central to the classification of animals, as compared with artifacts (Keil, 1994). For these reasons and more generally to replicate the basic findings of Experiment 3, we developed a set of full-color stimuli of photographic origin and again tested for efficient search and a correlation between contour shape and search slopes.

EXPERIMENT 4

In Experiment 4, we created a set of 64 color photographic stimuli (32 animals and 32 artifacts) on the basis of photo collections and scanned images from books, magazines, and catalogs. As in Experiment 3, 6 subjects searched for each of the stimuli in heterogeneous collections of distractors representing the other category.

Method

Subjects. Six participants completed the present experiment, including D.T.L. and 5 graduate and undergraduate students. One of the graduate students had also completed Experiment 3.

Stimuli. A new set of stimuli was created by assembling a set of 59 artifact images and 72 animals from a variety of sources. The stimuli were chosen from stock photo collections and were scanned in from books, magazines, and catalogs. During the initial stimulus collection, an effort was made to locate images that represented a wide variety of shapes and more specific categories. For the animals,

a special effort was made to reduce the prevalence of canonical body+legs-at-bottom images. Otherwise, the initial stimulus collection process was based on the availability of suitable images that had fully visible external contours and avoided highly unusual views. Once the initial set of stimuli was collected, the final set of 32 for each category was randomly selected, with the constraint that a wide variety of categories be represented. Therefore, when random selection caused overrepresentation of one very narrow category or the removal of too many distinctive shapes, the object was replaced, and another object was chosen. For these reasons, random choices were overruled for four animals. In addition, one artifact was replaced because it had a picture of an animal on it. In several instances, two members of the same basic-level category were allowed into the stimulus set. The Appendix includes the names of all the stimuli in the set.

Once selected, the images were separated from their backgrounds. Some had been photographed on neutral backgrounds, so this process could be done automatically. Others, however, had to be separated manually, using Photoshop clipping paths. Although this does represent some human intrusion in determining external contour shape, only pictures with easily visible figure/ground contours were selected for inclusion, and manual error was most likely 1 pixel or less, given that clipping paths were created using high-resolution versions of the stimuli, which were reduced to approximately one fourth of their original size for presentation. The animals from one of the photo collections were provided with predefined clipping paths.

Apparatus. The apparatus was similar to that used in Experiment 3, with the exception that all the monitors displayed stimuli in 256-level color and with a screen resolution of 600×800 pixels on 15-in. monitors. Individual stimuli subtended approximately 4.9° $\times$ 4.6°, and the entire array subtended approximately 18.4° $\times$ 14.9°.

Procedure. The procedure was identical to that used in Experiment 3. Again, the subjects completed the experiment in four sessions at their convenience. The contour analysis procedure was identical to that used in Experiment 3.

Results

Errors. The average error rate was 3.4%. The mean error rate for target-present trials was 4.55%, and the mean error rate for target-absent trials was 2.36%. The error rates for target-present trials were entered into a two-factor (task $\times$ display size) within-subjects ANOVA. The task and display size effects were significant [task effect, $F(1,5) = 11.279, MS_e = 4.771, p = .0201$; display size effect, $F(1,5) = 40.690, MS_e = 7.921, p = .0374$], whereas the task $\times$ display size interaction was nearly significant [$F(1,5) = 3.525, MS_e = 1.414, p = .1193$]. The error rates across display sizes were 4.30% and 7.81% for three- and nine-item artifact target displays and 2.21% and 3.91% for three- and nine-item animal displays, respectively.

Reaction times. As in previous analyses, only target-present trials were analyzed. Mean search slopes and intercepts for each condition closely replicated those observed in Experiments 1 and 3 (see Figure 7). Search for artifact targets was significantly faster (6.2 msec/item, as compared with 5.6 msec/item in Experiment 3) than search for animals [10.8 msec/item, as compared with 13.7 msec/item; $F(1,5) = 68.544, MS_e = 0.898, p = .0004$]. An item-based analysis comparing the 32 artifact target slopes with the 32 animal target slopes was similarly significant [$F(1,62) = 9.191, MS_e = 34.24, p = .0035$]. Target-absent slopes were 16.8 msec/item in the artifact
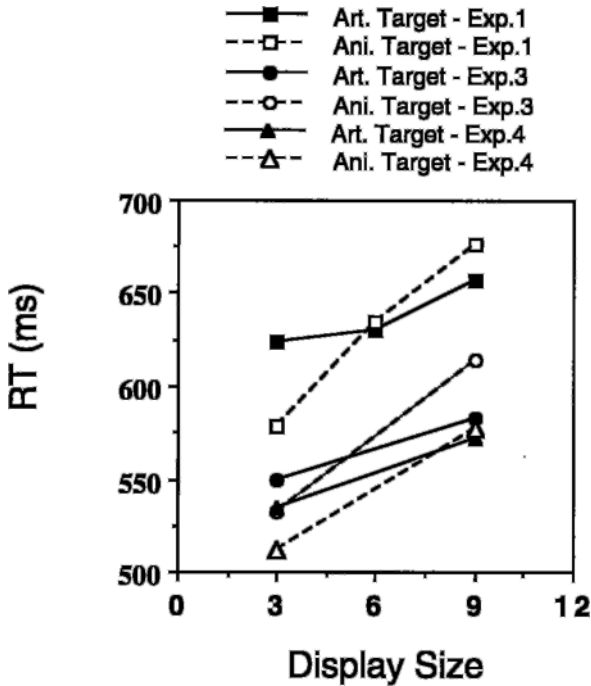
Figure 7. Summary of the search data for Experiments 1, 3, and 4.

target condition and 22.1 msec/item in the animal target condition. Mean RTs for three-item displays in the target-present trials were significantly faster for animal targets (512 msec) than for artifact targets [534 msec; $F(1,5) = 7.890, p = .038$]. There was no significant difference between RTs in large displays (572 msec for animal targets and 577 msec for artifact targets; $F < 1$).

The analysis of RTs over the four blocks of 512 trials revealed no learning effects. RTs did not decrease over blocks, nor was there any interaction between display size and block.

As is evident from Figure 8, rectilinearity was correlated with search slopes ($r = -.38, p < .05$) and intercepts ($r = -.55, p < .002$) for artifacts. For animals, the rectilinearity–slope and rectilinearity–intercept correlations were nonsignificant ($r = .28, p = .11$ and $r = .16$, $p = .38$ respectively).

Discussion

Experiment 4 closely replicates Experiments 1 and 3 by revealing highly efficient search by category in color images of photographic origin. Thus, the present results generalize well across stimulus sets that are perceptually quite different. In addition, the same search asymmetry, nearly the same slopes, and the same rectilinearity–search correlations were observed. The latter finding is particularly important because it confirms that global contour shape affects search by category in stimuli with minimal mediation via an artist's hand.

Before proceeding to Experiment 5, it is important to discuss the difference in search slopes between the tasks. In Experiments 1, 3, and 4, a search asymmetry was observed in which artifacts were located more quickly than animals. In Experiment 4, although the correlation is nonsignificant, the direction of the relationship suggests that relatively rectilinear animals are more difficult to detect among artifacts, the reverse of the relationship between rectilinearity and artifact slopes. This suggests that differences in contour shape between targets and distractors generally facilitate search. Curvilinear animals are relatively easy to detect among rectilinear artifacts, and rectilinear artifacts are relatively easy to detect among curvilinear animals. This, combined with the fact that animal distractors are generally more homogeneous in contour shape, might be sufficient to explain the search asymmetry without appeal to asymmetrically defined category-specifying features. In both of the stimulus sets, the standard deviation of rectilinearity values for animals was less than that for artifacts (Experiment 3, animal $SD =$
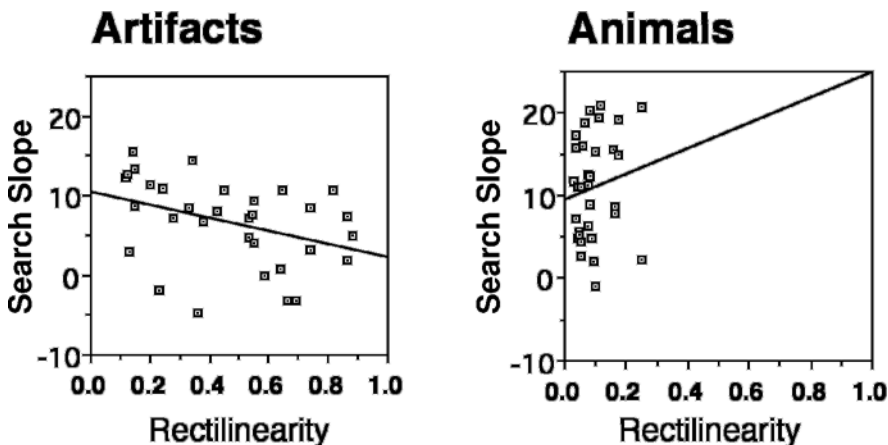

Figure 8. Correlation between search slopes and rectilinearity in Experiment 4. Note that the line of best fit for animal targets is nonsignificant even though it departs from zero, probably owing to range restriction in the rectilinearity measurements.

.032, artifact $SD = .097$; Experiment 4: animal $SD = .059$, artifact $SD = .245$). Thus, the animals had to be located among the heterogeneous artifacts, whereas artifacts could be located among the relatively homogeneous animals. According to Duncan and Humphreys (1989), heterogeneous distractors make targets difficult to detect. Thus, differences in the variability of contour shape may have generally facilitated detection of artifacts. If this is added to the assumption that target–distractor similarity with respect to contour shape also affects search slopes, we can explain both the asymmetry and the responsiveness of both tasks to the rectilinearity of targets.

EXPERIMENT 5

Although it seems that artifacts and animals differ in terms of their external contours, there are other between-category differences that might account for the results of Experiments 3 and 4. First, at a more general level, the rectilinearity measure could reflect general visual complexity. If an object's shape is simple, perhaps its contours do not change direction often, which would result in a high rectilinearity measure. Thus, facilitated detection of artifacts would be based on their simplicity. Second, each object has parts that might be characteristic of either its own category or the other. For example, parts such as handles and heads are characteristic of artifacts and animals, respectively. In addition, some objects have parts characteristic of the contrasting category. For example, the jaws of the pliers (Object 32 in Figure 1B) are similar in appearance to the jaws on an animal, and the antennae on the butterfly (Object 3 in Figure 1A) might be similar to the antenna on a TV. As is suggested by the internal analysis of Experiment 3, the typicality of the object parts may contribute to search efficiency, in addition to the rectilinearity of their contours.

For Experiment 5, we completed a series of regression analyses predicting search slopes for all the stimuli used in Experiments 3 and 4. In addition to rectilinearity, predictor variables included ratings of complexity, ambiguity, global typicality, and a set of five more specific and focused ratings of typicality. Our goal was twofold. First, we wanted to determine whether computed rectilinearity would make a unique contribution to slope predictions independent of other commonly measured variables. Second, we wanted to determine whether other variables would predict the slopes in a given task. If rectilinearity and only rectilinearity predicts search slopes, perhaps the task is more similar to simple feature searches than we had originally thought.

Method

Subjects. The analysis included three sets of ratings completed by different groups of subjects. The first two sets of ratings were done on the black-and-white drawings in Experiment 3 and the color images in Experiment 4, by 11 and 14 subjects, respectively, and were completed as part of the stimulus development process. These subjects were undergraduate and graduate psychology students at Cornell University. For the additional focused ratings of typicality, 7 subjects (graduate students in psychology at Kent State University) rated all 128 objects from Experiments 3 and 4. In all cases, the subjects who completed the ratings were different from those who completed the search experiments.

Materials. The first two sets of ratings used in this experiment were originally collected during the process of generating the original stimulus sets for Experiments 3 and 4. These included ratings of visual complexity, typicality, and ambiguity. The subjects completed the ratings by using paper packets depicting the stimuli (printed in color for the color set). They were instructed to rate complexity and ambiguity on the basis of the visual appearance of the objects as pictured but were told to rate typicality more broadly and to include visual and nonvisual characteristics. All ratings used a 7-point scale, with a rating of 7 being given to *highly typical*, *complex*, and *ambiguous* objects, respectively.

Another set of subjects completed five additional focused ratings of typicality for each of the 128 objects in Experiments 3 and 4. For the first rating (*part-typicality*), the subjects indicated the degree to which each object had parts that were typical of its category, and the instructions emphasized the need to rate typicality relative to "the appearance of the object." The second rating (*configuration typicality*) was similar, except that it emphasized the configuration of parts instead of the parts themselves, asking the subjects, "How well do the relative locations of the parts match your understanding of the typical arrangement for each kind?" Ratings 1 and 2 used scales ranging from *very typical* (rating of 1) to *very atypical* (rating of 7). The third rating (*view typicality*) asked subjects, "How well does the specific view in this specific photograph reveal the parts and configuration of parts as you know them for this object category," and used a scale ranging from *very characteristic* (1) to *not very characteristic* (7). The fourth and fifth ratings were similar to the first and second, except that they focused on the similarity of the parts and configuration of each object to those from the other category. For the fourth rating (*other* part) the subjects were asked, "Does this object have parts that look like they might come from the other kind," and for the fifth (*other whole*) the subjects were asked, "Does this object as a whole look like an object from the other kind?" Anchor points for both scales were *characteristic of own kind* (rating of 1) and *characteristic of other kind* (rating of 7). Although it would have been possible to make the fifth rating parallel the second more closely (by referring directly to the relative locations of parts), we chose to refer to the "whole" because we considered it unlikely that the subjects would rate many objects as specifically sharing a part configuration with the other category.

The subjects were given rating packets that included a numbered printout of the stimuli, an instruction sheet containing the rating scales, and a set of scantron sheets on which to record their responses. All of the subjects were given a verbal summary of the scales, then completed them on their own. The entire task took approximately 1.5 h.

Results

Item reliability. Alpha scores were computed for each item across subjects. For the ambiguity, complexity, and typicality ratings, reliabilities were .88, .95, and .76, respectively, for the drawings and .79, .88, and .66, respectively, for the color images. Reliabilities for the focused typicality ratings were .52, .40, .71, .66, and .73 for part-typicality, configuration-typicality, view-typicality, other-part, and other-whole ratings, respectively. A closer look at the less reliable part- and configuration-typicality ratings shows that artifacts were particularly difficult to rate; reliabilities for artifact part- and configuration-typicality ratings were .39 and .17, respectively. Individual typicality ratings were therefore not included in the artifact analysis. Animal ratings were better (reliabilities

of .58 and .52) but still not very reliable, so individual raters' responses were factor analyzed to determine whether a subset of raters with higher agreement could be isolated. Separate analyses of the animal part-typicality and configuration-typicality ratings each yielded two primary factors (collectively accounting for 60.5% of variance in part-typicality ratings and 52% of variance in configuration-typicality ratings). The first (and strongest) factor in both cases included the same group of 3 raters, so new reliabilities were computed using only these raters. Reliability for the first-factor part-typicality ratings was .86, and the reliability of the configuration-typicality rating was .79. The mean from these raters will constitute the typicality ratings for the remainder of the analysis. (However, significant typicality effects were rerun with the mean of all raters and were not qualitatively different.[3])

Because the part-typicality and configuration-typicality ratings were associated with relatively low reliabilities, and owing to the high part-typicality/configuration-typicality and other-part/other-whole correlations (.743 and .732, respectively), a summary measure was computed that included the mean response to all four ratings. This measure was created by averaging the mean of all four ratings for the 3 first-factor subjects discussed above, with the mean of the more reliable other-part and other-whole ratings for all 7 raters. (Another set of regressions was run that included all 7 raters' data on all four scales, and this produced almost identical results.) The alpha reliability for this composite rating was .81, and it will be referred to as the *focused typicality* rating. This score is high for objects that are atypical of their own category or similar to the other category.

General strategy for regression analysis. Analyses predicting search slopes were completed, using two blocks of variables. For all regressions, search intercepts were entered in an initial block to examine the effect of search termination independent of other variables. In an initial regression, all the remaining variables were entered in the second block. This will be referred to as the *full model.* A second regression was run that focused on maximizing multiple r values with the most parsimonious model while keeping colinearity to a minimum (*basic model*). This was done by avoiding regressions that contained predictors with correlations greater than .60. A correlation matrix including all seven predictors was created for each category. For artifacts one correlation (21 possible) was above .60 (the correlation between the basic typicality rating and complexity was .619), and for animals none was. Finally, these analyses were confirmed with stepwise regressions using all variables, done after a first block forcing intercepts into the equation (stepwise model).

In addition, departures from normality were corrected transformations that were applied to all variables except the focused typicality rating and search slopes. Cube root transformations were used in all cases except for search intercepts, which were adjusted with a reciprocal transformation.

Animal slope predictions. The full animal model accounted for 46.7% of the variance in slopes (see Table 1A).

Table 1A
Regressions Predicting Search Slopes for Animal Targets

Variable	B	SE B	β	T	Sig T
Full Model (Multiple $R = .684$; $R^2 = .467$)					
INT.T	21.660640	5.098476	.504383	4.248	.0001
AMB.T	5.665797	4.036385	.165408	1.404	.1659
CMPL.T	10.510682	4.287268	.267231	2.452	.0174
RLIN.T	19.390005	5.364974	.384114	3.614	.0006
TYP.T	−8.546618	5.125400	−.187687	−1.668	.1010
FocTyp	3.798596	.767565	.552792	4.949	.0000
ViewTyp.T	1.348397	4.526186	.033590	.298	.7669
(Constant)	−66.726922	16.274271		−4.100	.0001
Basic Model (Multiple $R = .655$; $R^2 = .429$)					
INT.T	23.073850	4.871097	.537290	4.737	.0000
CMPL.T	7.566668	4.039317	.192380	1.873	.0660
RLIN.T	18.751544	5.379376	.371466	3.486	.0009
FocTyp	3.926384	.752464	.571389	5.218	.0000
(Constant)	−67.857371	14.063832		−4.825	.0000
Stepwise Model (Multiple $R = .628$; $R^2 = .395$)					
INT.T	22.018171	4.938517	.512708	4.458	.0000
RLIN.T	17.931437	5.472488	.355220	3.277	.0017
FocTyp	4.205522	.752826	.612010	5.586	.0000
(Constant)	−54.961835	12.517628		−4.391	.0000

Note—INT.T, transformed search intercept; AMB.T, transformed ambiguity rating; CMPL.T, transformed complexity rating; RLIN.T, transformed rectilinearity measure; TYP.T, transformed global typicality rating; FocTyp, focused typicality rating; ViewTyp.T, transformed view typicality rating.

Intercepts, when entered alone in an initial block, were nonsignificant predictors of slopes ($\beta = .19$, $p = .13$), but were significant when combined with the other variables in the full regression ($\beta = .50$, $p = .0001$). Other significant predictors were complexity, rectilinearity, and the focused typicality rating. The ambiguity and global typicality measures were marginally nonsignificant, and the view typicality measure did not approach significance (see Table 1A). A more parsimonious model that did not include the marginal predictors was also run. It accounted for 42.9% of the variance in search slopes ($R = .655$) and included intercepts, complexity, rectilinearity, and focused typicality as predictors. In addition, a stepwise regression was run to verify the predictiveness of variables in the basic model. In this regression, intercepts were again entered in a first block; then the other six variables were entered stepwise in the second block. The criteria were $p = .05$ for inclusion and $p = .10$ for exclusion. This analysis largely confirmed the initial analysis by producing a model that included intercepts, rectilinearity, and the focused typicality measure (see Table 1A).

Variations of the basic model were tested by substituting the individual typicality ratings for the summary rating. These showed that the other-whole/configuration-typicality ratings were stronger predictors than the other part/part-typicality ratings and that the part-typicality/configuration-typicality ratings were stronger predictors than the other-part/other-whole ratings. In particular, the model with the configuration-typicality and other-whole ratings explained 56% of the variance in slopes, whereas the model with the part-typicality and other-part ratings explained 38% of the variance in slopes. In the part-

Table 1B
Regressions Predicting Search Slopes for Artifact Targets

Variable	B	SE B	β	T	Sig T
Full Model (Multiple R = .573; R² = .328)					
INT.T	11.910143	5.061855	.335799	2.353	.0222
AMB.T	7.429470	4.454770	.243584	1.668	.1009
CMPL.T	3.827664	3.558746	.158524	1.076	.2867
RLIN.T	−18.209846	5.804449	−.449252	−3.137	.0027
TYP.T	−1.957705	7.059773	−.046435	−.277	.7826
FocTyp	2.512276	1.102144	.305845	2.279	.0265
ViewTyp.T	−3.257734	5.198351	−.080823	−.627	.5334
(Constant)	−18.312236	12.601577		−1.453	.1518
Basic Model (Multiple R = .558; R² = .312)					
INT.T	11.947608	4.974639	.336855	2.402	.0195
AMB.T	6.778213	3.575112	.222232	1.896	.0629
RLIN.T	−19.069971	5.551178	−.470472	−3.435	.0011
FocTyp	2.627636	1.075148	.319888	2.444	.0175
(Constant)	−18.814889	10.151024		−1.853	.0688
Stepwise Model (Multiple R = .520; R² = .270)					
INT.T	13.936630	4.966788	.392934	2.806	.0068
RLIN.T	−16.647235	5.517665	−.410702	−3.017	.0037
FocTyp	2.973238	1.082251	.361962	2.747	.0079
(Constant)	−16.332719	10.281591		−1.589	.1174

Note—INT.T, transformed search intercept; AMB.T, transformed ambiguity rating; CMPL.T, transformed complexity rating; RLIN.T, transformed rectilinearity measure; TYP.T, transformed global typicality rating; FocTyp, focused typicality rating; ViewTyp.T, transformed view typicality rating.

typicality/other-part regression, the part-typicality rating was associated with a β of .587 (p = .0001), whereas the other-part rating was associated with a β of −.10 (p = .431). In the configuration-typicality/other-whole regression, the typicality rating was associated with a β of .572 (p < .0001), whereas the other-whole rating was associated with a β of .18 (p = .118).

Artifact slope predictions. In the full model, intercepts, rectilinearity, and the focused typicality ratings emerged as significant predictors (see Table 1B) in a model that explained 32.8% of slope variance (R = .573). Again, intercepts were nonsignificant predictors when tested alone in an initial block (β = −.013). In the more parsimonious basic model (Table 1B), artifact search slopes were significantly predicted by intercepts, rectilinearity, and the focused typicality ratings and were marginally predicted by ambiguity ratings. Collectively, these variables predicted 31.2% of the variance (R = .558). A stepwise analysis, the same as that run for animal targets, again confirmed the predictiveness of variables in the basic model. It included intercepts, rectilinearity, and the focused typicality measure (see Table 1B).

Regressions substituting individual focused typicality ratings for the summary rating were run for artifacts, using only the other-part and other-whole ratings because the reliability coefficients for the part- and configuration-typicality ratings were very low for artifacts. Two variants of the basic model were run, one using intercepts, rectilinearity, ambiguity, and the other-part rating, and one that substituted the other-whole rating for the other-part rating. The two regressions were very similar, with the

whole rating producing a slightly better fit. The model with the other-whole rating explained 30% of the variance in slopes, whereas the model with the other-part rating explained 28% of the variance in slopes. The other-part predictor narrowly missed conventional levels of significance (β = .231, p = .0740), and the other-whole rating was significant (β = .289, p = .0281).

Discussion

Experiment 5 reveals two primary findings. First, rectilinearity predicts search slopes for both artifacts and animals. Second, in addition to rectilinearity, the focused typicality ratings predict search slopes for both categories. The primary differences between the regressions for each category are that the signs on the rectilinearity coefficients are opposite and that the relative effectiveness of the two focused typicality ratings switch. The former difference reflects the finding that rectilinear artifacts are easy to locate among curvilinear animals and, conversely, curvilinear animals are easy to locate among rectilinear artifacts.

At this point it is interesting to consider the relationship between rectilinearity's successful prediction of search slopes and previous findings that curvature can act as a "basic" or primitive feature in visual search (Treisman & Gormican, 1988; Wolfe, Yee, & Friedman-Hill, 1992). In these previous findings, curved line segments were easily located among straight line segments, whereas straight lines were more difficult to locate among curved segments. Thus, curvature would appear to be a feature, whereas straightness is not, because it is easier to locate a feature-positive item among feature-negative items than the reverse (Treisman & Gormican, 1988). This seems to be in contrast with our finding, in which curvilinearity and rectilinearity appear to constitute different feature values on a single dimension. As was discussed above, the search asymmetry we observed can be explained by assuming that artifact slopes are generally shallower because animal distractors are more homogeneous in contour shape (and probably parts as well). Also, although rectilinear artifacts are more easily located among curvilinear animals, the reverse is also true: Curvilinear animals are more easily located among the rectilinear artifacts. Therefore, there is no real feature coding asymmetry in the present case, whereas there is one favoring curvilinear line segments. This finding is reminiscent of early work on serial visual search in which subjects searched down columns of letters and were able to scan more quickly when looking for curved letters (such as Q) than for rectilinear letters (Z) among rectilinear distractors, whereas the reverse was true when distractors were curvilinear (Neisser, 1963), although it is not clear whether these stimuli would produce a parallel search in a simultaneously scanned object array. One possibility is that curvature is a primitive in line segments, whereas the entire continuum from rectilinearity to curvilinearity has equal status in objects. This qualification points out the difficulty of extrapolating from the primitives defined

by simple stimuli, such as line segments, to more complex objects and reinforces the suggestion that visual search often operates more on completed forms or surfaces than on primitive features (e.g., He & Nakayama, 1992; Suzuki & Cavanagh, 1995).

The switch in effectiveness of the part-typicality rating and the other-part rating reflects the difficulty in rating the visual typicality of artifacts. Both the part-typicality and the configuration-typicality ratings were unreliable for artifacts, suggesting that the subjects were unable to conceive of what the typical artifact ought to look like. For animals, on the other hand, these ratings were much more reliable and proved to be successful in predicting search slopes. As was mentioned above, a number of authors have suggested that distinguishing among animals requires more extensive visual processing, because they tend to share one of a limited number of parts and configurations (e.g., Farah, 1995; Warrington & Shallice, 1984). For this reason, it was probably considerably easier for judges to rate each animal with reference to one or more representations of their typical appearance. Conversely, it was also possible for judges to reliably rate the degree to which artifacts were visually similar to animals. Therefore, it remains an open question as to whether a different typicality rating would have been more successful for artifacts. For example, judges might rate how typical each artifact was with reference to its basic-level category. It is also important to note that the focused typicality ratings were successful where the more global rating was not. Most likely, this was due to the visual emphasis inherent in the former. The global typicality ratings did not specify visual typicality per se and may have been influenced by nonvisual factors, such as membership in a typical species or knowledge of typical behaviors. Generally, however, one finding is clear from Experiment 5: Both rectilinearity and some measure of visual typicality or similarity to the contrasting category predicted search slopes for artifacts and animals.

Another finding should be noted because it conflicts to a degree with the findings of Experiment 2, in which jumbling slowed search only minimally. If the noneffect of jumbling means that configural and form information are completely ineffective in facilitating search, why did the configuration-typicality and other-whole ratings predict search slopes so well? Although it is not possible to be certain, at least two factors may be relevant for producing this finding. First, the part-typicality and configuration-typicality ratings were correlated ($r = .641$ for artifacts and $r = .862$ for animals), as were the other-part and other-whole ratings ($r = .806$ for artifacts and $r = .588$ for animals). Therefore, both ratings may have tapped a general impression of visual typicality that depended on parts and configuration. It is also possible that the raters took the configuration/whole ratings to mean that they should rate the entire object, inclusive of all parts, and perhaps focused on only a single part for the part ratings. Finally, the stimuli in Experiment 2 may have disrupted the relations among some parts, while leaving other aspects of

the object's general form intact. For example, in Figure 3, the basic horizontal body shape is still present in some of the animals, and in addition, some part/whole relationships remain intact (e.g., legs and heads still protrude from the body). Given the research cited earlier that shows that some basic aspects of form may be available preattentively, it is possible that the global shape affected both the ratings and the search slopes. If we assume that animals are more likely to be associated with a specific global shape, this explanation fits well with the fact that whole ratings improved animal slope predictions the most.

We interpret these findings to suggest that visual search by category is a complex process that depends on multiple levels of simultaneous analysis. Consistent with current theory, the faster search (that for the artifact target) is relatively more dependent on contour rectilinearity, which is a plausible perceptual primitive, and in the slower animal search is more dependent on the typicality measure (although this search is also fast, with many targets being located in less than 10 msec/item). However, in contrast to established theory, we find no sharp distinction between tasks that depend on a primitive versus a more conceptual variable, in that both tasks depend to a degree on both kinds of property. Indeed, a recent review by Wolfe (1998b) confirms blurring of the distinction between parallel primitive-based search and a serial search based on focal attention and goes on to suggest that this finding implies the need for a revision in our conceptualization of the visual search task as a whole. In the General Discussion section, we summarize our findings and argue that they suggest the need for an added layer of explanation to account for the efficiency with which people search for members of broad natural object categories.

GENERAL DISCUSSION

The results presented here provide evidence for an efficient process that can distinguish categories and points to specific features that underlie this process. Experiments 1, 3, and 4 converge to show that search for an unspecified artifact in a heterogeneous field of animals proceeds at 5.5–6.2 msec/item. These experiments also reveal a search asymmetry; search slopes for animal targets among artifacts were 10.8–16.0 msec/item. In both cases, shallow slopes were accompanied by relatively low intercepts (512–624 msec), suggesting that the subjects were not withholding responses to three-item displays (and thereby producing artificially low slopes). Experiment 2 tested the hypothesis that the efficient search observed in Experiment 1 was based on high-level features and nonperceptual information by jumbling objects in order to disrupt completion of structural descriptions. Search slopes were not significantly slower for part-jumbled targets, although intercepts were approximately 100 msec longer. Experiments 3 and 4 tested the possibility that between-category differences in contour shape facilitated the search. We found that target–distractor differences in rectilinearity, combined with homogeneity

of contour statistics, facilitate search, and in the case of rectilinear artifacts among relatively homogeneous animals, the slope approaches 0 msec/item. Thus, visual search by category is efficient and can be facilitated by the presence of natural nondefining features. Finally, Experiment 5 tested for other properties of targets that might explain search slopes. In a multiple regression analysis, not only did rectilinearity predict search slopes, but visual typicality (for animals) and similarity to the distracting category (for artifacts) did so as well.

Efficient Search by Category and Theories of Visual Search

The distinction between serial and parallel search is central to many models of visual search. For example, FIT suggests that parallel searches are observed only where a single feature distinguishes targets from distractors. In addition, that feature must be one among a small set of perceptual primitives coded early in the visual system. A target that is a different color from distractors or one that has a fundamentally different shape from distractors will "pop out" and be located without the necessity of a serial search. Other theories, such as guided search (Wolfe et al., 1989), are similar in that they emphasize the role of unique features or unique conjunctions of features in allowing a parallel search. Duncan and Humphreys (1989) described a more general model relying on the similarity between targets and distractors and variations in distractor similarity to explain visual search. All of these theories were developed and tested using artificial stimuli, carefully controlled to reflect simple feature differences that are almost entirely diagnostic of the target–distractor difference. The question is, can these theories serve as useful frameworks to understand visual search for natural categories as well?

We think existing models are helpful in understanding the speed of the search and the search asymmetry. It appears as though parallel search was possible, at least partly, on the basis of features very similar to the perceptual primitives discussed by Treisman and Wolfe. Rectilinearity might be seen as a form primitive that is, indeed, based on early visual codes. As was mentioned previously, rectilinear and curvilinear contours may be represented in functionally different early visual maps. In addition, the variable nature of the rectilinearity feature is clearly not problematic, in that both FIT and guided search include variability in the amount of signal each feature represents, as measured against the noise represented by the distractors. Somewhat rectilinear features induce a smaller amount of signal in the relevant feature map, and highly rectilinear targets induce a stronger signal. As was suggested in the discussion of Experiment 4, the Duncan and Humphreys (1989) model fits our data as well. Similarity between targets and distractors and similarity among distractors both appear to affect search in ways that this model can explain.

Although current models seem to provide a helpful point of departure for understanding, they do not appear to offer a complete explanation. In particular, it appears that no specific primitive feature or conjunction of features can distinguish all targets from distractors. As was mentioned above, rectilinearity might be a primitive, but by itself it is not sufficient for the task, and the predictive power of visual typicality for animal targets and similarity to the distractor category for artifacts points away from an exclusively primitive-based explanation. In addition, artifacts that were not different from animals in rectilinearity were still located quickly (10–12 msec/item). Of course, one could argue that we simply have not found the right primitive and that, if we were to search further, perhaps some simple feature could be found that distinguishes all of our targets from all of our distractors and that, in addition, explains the variance captured by both rectilinearity and typicality. Although a careful look through our stimuli by ourselves and others has yielded no workable candidates, we admit that we cannot prove that none exists. However, proving this kind of negative hypothesis is inherently impossible.

Instead, we should consider more carefully the fundamental advantage of referring to feature primitives in the first place. The benefit of primitives to any theory of vision is to propose a limited set of features that serve as building blocks for more complex representations. Therefore, these features should be few in number, simple, and combine productively to produce a wide variety of representations. However, as Nakayama and Joseph (1997) point out, we might need to question current models of preattentive vision in part because these requirements are apparently not being met. For example, they argue that quite a number of primitives are being cataloged for their ability to drive efficient search and, in addition, that these primitives are not particularly simple and often depend on completed descriptions of surfaces, depth relations, and 3-D structure. Given this critique, primitive-based theories will have difficulty predicting a priori when a search will be efficient on the basis of known stimulus properties. In a sense, Wolfe (1998a) points out this difficulty by suggesting that the list of primitives will be different when measured in different ways. Here, it appears that the list will also be different for different kinds of stimuli within the search task. This problem can be most clearly seen in that curvature and rectilinearity have equal status as feature values on a given dimension, whereas previous research suggested that curvature is a feature against the norm of straightness. As was mentioned above, the difference between the present case and previous findings is that curvature is a characteristic of closed object contours here, whereas in previous research curvature has been a property of line segments. Therefore, it appears that combining segment primitives into contour primitives is not straightforward. This problem may be particularly acute in complex natural stimuli because, in these situations, it seems necessary to posit more and more complex primitives. Therefore, we should worry about being stuck inventing new primitives for each category contrast we test. If one allows for complex primi-

tives that change in different sets of stimuli, especially ones that closely conform to the perceptual boundaries between specific natural categories, they cease to be primitives that can be combined productively. Instead, they might be better described as the perceptual description of the categories themselves.

If the features that contrast specific categories can guide efficient visual search, we should consider the need to explain how these features combine and guide a category search on any given trial. One possibility is that a known category is described by a search template (e.g. Duncan & Humphreys, 1989) that specifies features at a variety of levels of complexity, ranging from simple aspects of form to a list of typical parts. In the case of animals, this might involve a fairly straightforward analysis of the degree to which potential targets have typical animal parts, such as eyes, a head, legs, and a body. As was mentioned above, artifacts are more difficult to associate with a globally typical set of parts or a configuration, and our ratings clearly reflect this. Therefore, the predictiveness of the similar-to-other ratings in artifacts may reflect a between-category difference in the usability of typical visual information, or it may reflect a difference in the level of generality in the analysis between the categories. In either case, detection based on both a simple perceptual feature and the relationship between each object and its category also suggests that processing prior to focal attention proceeds at multiple levels of analysis simultaneously. In part, this analysis might reflect segregated activations in parts of the visual system specialized for processing each category, especially if one assumes that visually typical objects produce more activation in their respective parts of the visual system.

At this point, it is helpful to mention an important limitation to the data reported here. We have tested only one broad categorical distinction and have assumed that the efficiency associated with it is caused by its familiarity or foundational nature. However, although we consider it unlikely, it remains possible that other, similar nonfoundational contrasts would also be efficiently processed, even in cases in which no single feature distinguishes the categories. Therefore, it might be worthwhile to attempt experiments in which the same heterogeneous stimulus set was divided into both foundational categories and more arbitrary (but equally heterogeneous) nonfoundational categories. For example, it might be possible to have subjects search a carefully chosen set of animals and artifacts, using either the animal/artifact distinction or another distinction, such as that between objects typically found in America versus those found in Africa or that between objects larger than a breadbox versus those smaller than a breadbox. Although it might be difficult to argue that inefficiently distinguished nonfoundational contrasts are as separable in form as the efficiently distinguished foundational contrast, the experiment seems worth attempting, because it would further support the hypoth-

esis that real object categories can escape the effects of their formal complexity.

According to the above argument, understanding visual search in natural categories does not require abandoning current theory. Rather, it requires an added level of explanation of how features cohere into known bundles or categories. This explanation will probably include both a general description of the limits of feature combination (a good example is Wolfe & Bennett, 1997) and a description of the specific categories that search stimuli represent. The concept literature provides a good starting point in suggesting that knowledge is organized around a few foundational categories (Carey, 1985; Keil, 1989) that may have similar organizing effects in the visual system. Therefore, category-specific explanations need not proliferate uncontrollably and might limit themselves to broadly inclusive categories. One area of research that seems to confirm the tractability of this problem is neuropsychological findings that brain injuries can selectively reduce visual performance for one broadly inclusive category or another.

Neuropsychological Findings

Recent neuropsychological evidence suggests that both semantic and object recognition systems may be organized by categories and levels of abstraction of features that are different in each stage. Caramazza and Shelton (1998) described a patient, E.W., with a specific deficit in recognizing animals, but not other living or nonliving things. E.W. performed poorly in naming and real/unreal object judgment tasks for animals. Overall, the pattern of deficit suggested that her impairment was specific to the animal kind, rather than to a particular feature (e.g., sensory or functional features). Caramazza and Shelton suggested that the semantic system is organized by category and that the neural substrate for recognizing animals and plants may be distinct from that for other objects because recognizing animals and plants accurately gives a large evolutionary advantage.

If, in fact, semantic storage is organized by category, it is possible that perceptual processing is similarly organized. In fact, a number of patients show a category-specific impairment in the real/unreal object-decision task, which indicates a category-specific impairment in the structural analysis of objects (Mauri, Daum, Sartori, Riesch, & Birbaumer, 1994; Sartori & Job, 1988; Sartori, Job, & Colheart, 1992). Furthermore, a category-specific impairment in the structural analysis stage can be independent of semantic deficits. Sheriden and Humphreys (1993) described a patient with difficulty in recognizing animals and food who performed adequately in real/nonreal object judgment for animals and food. Some evidence suggests that a selective impairment in processing a lower level feature can also cause an impairment in recognizing a particular class of objects. Kosslyn, Hamilton, and Bernstein (1995) reported a study of a prosopag-

nosia patient whose difficulty in recognizing faces seemed to derive from his difficulty in processing curvature.

The neuropsychological data therefore suggest that processing systems may be segregated by category in late vision and, furthermore, that different early visual systems may connect differentially to these later areas. Given that the search asymmetry in the present experiment could be explained by differences in the heterogeneity of the two stimulus categories, it is possible that curvilinear and rectilinear contours are processed by systems that are independent at some point. We might, therefore, expect at some point to observe a patient who has difficulty integrating rectilinear contours to contrast with patients who seem to have a curvilinearity deficit.

Summary and Conclusions

The experiments reported here suggest that an efficient visual search for targets in a broad category depends on both basic perceptual features, such as rectilinearity, and a more sophisticated analysis of part typicality. Thus, the surprisingly efficient visual search finding seems to represent a case in which early visual processes, prior to the arrival of focal attention, reflect the natural categorical organization of the visual world. Although this finding would not necessarily have been predicted on the basis of current theories of visual search, it does not invalidate them. Rather, it suggests that added levels of explanation are necessary to model real-world visual search successfully. In particular, it seems necessary to explain how different perceptual features cohere to form descriptions of targets and distractors that can be used as the basis of efficient visual search.

REFERENCES

BERTENTHAL, B. I. (1993). Infants' perception of biomechanical motions: Intrinsic image and knowledge based constraints. In C. Granrud (Ed.), *Visual perception and cognition in infancy* (pp. 175-214). Hillsdale, NJ: Erlbaum.

BIEDERMAN, I., BLICKLE, T. W., TEITELBAUM, R. C., & KLATSKY, G. J. (1988). Object search in nonscene displays. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **14**, 456-467.

CARAMAZZA, A., & SHELTON, J. R. (1998). Domain-specific knowledge systems in the brain: The animate–inanimate distinction. *Journal of Cognitive Neuroscience*, **10**, 1-34.

CAREY, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.

CHUN, M. M., & WOLFE, J. M. (1996). Just say no: How are visual searches terminated when there is no target present? *Cognitive Psychology*, **30**, 39-78.

DONNELLY, D., HUMPHREYS, G. W., & RIDDOCH M. J. (1991). Parallel computation of primitive shape descriptions. *Journal of Experimental Psychology: Human Perception & Performance*, **17**, 561-570.

DUNCAN, J. (1983). Category effects in visual search: A failure to replicate the "oh–zero" phenomenon. *Perception & Psychophysics*, **34**, 221-232.

DUNCAN, J., & HUMPHREYS, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, **96**, 433-458.

EGETH, H., JONIDES, J., & WALL, S. (1972). Parallel processing in multielement displays. *Cognitive Psychology*, **3**, 674-698.

ENNS, J., & RENSINK, R. A. (1991). Preattentive recovery of three-dimensional orientation from line drawings. *Psychological Review*, **98**, 101-118.

FARAH, M. J. (1995). Dissociable systems for recognition: A cognitive neuropsychology approach. In S. M. Kosslyn & D. N. Osherson (Eds.), *Visual cognition: An invitation to cognitive science* (pp. 101-120). Cambridge, MA: MIT Press.

FARAH, M. J., MCMULLEN, P. A., & MEYER, M. M. (1991). Can recognition of living things be selectively impaired? *Neuropsychologia*, **29**, 185-193.

FUNNELL, E., & SHERIDAN, J. (1992). Categories of knowledge? Unfamiliar aspects of living and nonliving things. *Cognitive Neuropsychology*, **9**, 135-153.

GAFFAN, D., & HEYWOOD, C. A. (1993). A spurious category-specific visual agnosia for living things in normal human and nonhuman primates. *Journal of Cognitive Neuroscience*, **5**, 118-128.

GILBERT, C. D. (1993). Circuitry, architecture, and functional dynamics of visual cortex. *Cerebral Cortex*, **3**, 373-386.

GUTHEIL, G., VERA, A., & KEIL, F. C. (1998). Do houseflies think? Patterns of induction and biological beliefs in development. *Cognition*, **66**, 33-49.

HE, Z. J., & NAKAYAMA, K. (1992). Surface versus features in visual search. *Nature*, **359**, 231-233.

INTRAUB, H. (1981). Rapid conceptual identification of sequentially presented pictures. *Journal of Experimental Psychology: Human Perception & Performance*, **7**, 604-610.

JAIN, R., KASTURI, R., & SCHUNCK, B. G. (1995). *Machine vision*. New York: McGraw-Hill.

JONIDES, J., & GLEITMAN, H. (1972). A conceptual category effect in visual search: O as letter or as digit. *Perception & Psychophysics*, **12**, 457-460.

KEIL, F. C. (1989). *Concepts, kinds, and cognitive development*. Cambridge, MA: MIT Press.

KEIL, F. C. (1994). Explanation, association, and the acquisition of word meaning. *Lingua*, **92**, 169-196.

KOSSLYN, S. M., HAMILTON, S. E., & BERNSTEIN, J. H. (1995). The perception of curvature can be selectively disrupted in prosopagnosia. *Brain & Cognition*, **27**, 36-58.

KRUEGER, L. E. (1984). The category effect in visual search depends on physical rather than conceptual differences. *Perception & Psychophysics*, **35**, 558-564.

KURBAT, M. A. (1997). Can the recognition of living things really be selectively impaired? *Neuropsychologia*, **35**, 813-827.

MACK, A., & ROCK, I. (1998). *Inattentional blindness*. Cambridge, MA: MIT Press.

MANDLER, J. M., & MCDONOUGH, L. (1993). Concept formation in infancy. *Cognitive Development*, **8**, 291-318.

MARTIN, A., HAXBY, J., LALONDE, F. M., WIGGS, C. L., & UNGERLEIDER, L. G. (1995). Discrete cortical regions associated with knowledge of color and knowledge of action. *Science*, **270**, 102-105.

MARTIN, A., WIGGS, C. L., UNGERLEIDER, L. G., & HAXBY, J. V. (1996). Neural correlates of category-specific knowledge. *Nature*, **379**, 649-652.

MASSEY, C. M., & GELMAN, R. (1988). Preschoolers' ability to decide whether a photographed unfamiliar object can move itself. *Developmental Psychology*, **24**, 307-317.

MAURI, A., DAUM, I., SARTORI, G., RIESCH, G., & BIRBAUMER, N. (1994). Category-specific semantic impairment in Alzheimer's disease and temporal lobe dysfunction: A comparative study. *Journal of Clinical & Experimental Psychology*, **16**, 689-701.

NAKAYAMA, K., & JOSEPH, J. (1997). Attention, pattern recognition, and popout in visual search. In R. Parasuraman (Ed.), *The attentive brain* (pp. 279-298). Cambridge, MA: MIT Press.

NEISSER, U. (1963). Decision-time without reaction-time: Experiments in visual scanning. *American Journal of Psychology*, **76**, 376-385.

NICKERSON, R. S. (1968). A note on long-term recognition memory for pictorial material. *Psychonomic Science*, **11**, 58.

OLSHAUSEN, B. A., & FIELD, D. J. (1996). Relations between the "association field" and the statistics of natural scenes [Abstract]. *Investigative Ophthalmology & Visual Science*, **37**, 2370.

POTTER, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Language & Memory*, **2**, 509-522.

POTTER, M. C., & FAULCONER, B. A. (1975). Time to understand pictures and words. *Nature*, **253**, 437-438.

RAMACHANDRAN, V. S. (1988, Month). Perceiving shape from shading. *Scientific American*, **259**, 76-83.

RENSINK, R. (2000). The dynamic representation of scenes. *Visual Cognition*, **7**, 17-42.

ROSENGREN, K. S., GELMAN, S. A., KALISH, C. W., & MCCORMICK, M. (1991). As time goes by: Children's early understanding of growth in animals. *Child Development*, **62**, 1302-1320.

SACCHETT, C., & HUMPHREYS, G. W. (1992). Calling a squirrel a squirrel but a canoe a wigwam: A category-specific deficit for artifactual objects and body parts. *Cognitive Neuropsychology*, **9**, 73-86.

SANOCKI, T., BOWYER, K., HEATH, M., & SARKAR, S. (1996, May). *Are edges sufficient for object recognition?* Paper presented at the Object Perception and Memory Conference, Chicago.

SARTORI, G., & JOB, R. (1988). The oyster with four legs: A neuropsychological study on the interaction of visual and semantic information. *Cognitive Neuropsychology*, **5**, 105-132.

SARTORI, G., JOB, R., & COLHEART, M. (1992). The organization of object knowledge: Evidence from neuropsychology. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp. 451-465). Cambridge, MA: MIT Press.

SHEPARD, R. N. (1967). Recognition memory for words, sentences and pictures. *Journal of Verbal Learning & Verbal Behavior*, **6**, 156-163.

SHERIDAN, J., & HUMPHREYS, G. W. (1993). A verbal-semantic category-specific recognition impairment. *Cognitive Neuropsychology*, **10**, 143-184.

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.

STANDING, L., CONEZIO, J., & HABER, R. N. (1970). Perception and memory for pictures: Single-trial learning of 2500 visual stimuli. *Psychonomic Science*, **19**, 73-74.

STEWART, F., PARKIN, A. J., & HUNKIN, N. M. (1992). Naming impairments following recovery from herpes simplex encephalitis: Category-specific? *Quarterly Journal of Experimental Psychology*, **44A**, 261-284.

SUZUKI, S., & CAVANAGH, P. (1995). Facial organization blocks access to low-level features: An object-inferiority effect. *Journal of Experimental Psychology: Human Perception & Performance*, **21**, 901-913.

THEEUWES, J. (1993). Visual selective attention: A theoretical analysis. *Acta Psychologica*, **83**, 93-154.

THORPE, S., FIZE, D., & MARLOT, C. (1996). Speed of processing in the human visual system. *Nature*, **381**, 520-522.

TRANEL, D., LOGAN, C. G., FRANK, R. J., & DAMASIO, A. R. (1997). Explaining category-related effects in the retrieval of conceptual and lexical knowledge for concrete entities: Operationalization and analysis of factors. *Neuropsychologia*, **35**, 1329-1339.

TREISMAN, A., & GELADE, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, **12**, 97-136.

TREISMAN, A., & GORMICAN, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, **95**, 15-48.

VAN DE WALLE, G. A., & HOERGER, M. (1996). The perceptual foundations of categorization in infancy [Abstract]. *Infant Behavior & Development*, **19**, 794.

WARREN, W. H. (1984). Perceiving affordances: Visual guidance of stair climbing. *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 683-703.

WARRINGTON, E. K., & SHALLICE, T. (1984). Category specific semantic impairments. *Brain*, **107**, 829-854.

WOLFE, J. M. (1994). Visual search in continuous naturalistic stimuli. *Vision Research*, **34**, 1187-1195.

WOLFE, J. M. (1998a). Visual search. In H. Pashler (Ed.), *Attention* (pp. 13-73). Hove, U.K.: Psychology Press.

WOLFE, J. M. (1998b). What can 1 million trials tell us about visual search? *Psychological Science*, **9**, 33-39.

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

WOLFE, J. M., CAVE, K. R., & FRANZEL, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 419-433.

WOLFE, J. M., YEE, A., & FRIEDMAN-HILL, S. R. (1992). Curvature is a basic feature for visual search tasks. *Perception*, **21**, 465-480.

WOODWARD, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, **69**, 1-34.

ZELINSKY, G. J., RAO, R. P. N., HAYHOE, M. M., & BALLARD, D. H. (1997). Eye movements reveal the spatiotemporal dynamics of visual search. *Psychological Science*, **8**, 448-453.

ZUSNE, L. (1975). Curved contours and the associative response. *Perceptual & Motor Skills*, **40**, 203-208.

NOTES

1. The concept of preattentive vision has come under attack recently, given findings that little information seems to be registered under conditions of inattention (e.g., Mack & Rock, 1998). Thus, we refer to preattentive vision in the less literal sense of visual processing that occurs prior to the arrival of focal attention on a specific object or part of an object, while leaving open the possibility that prior to the arrival of focal attention, the scene as a whole is processed by a more broad application of attention.

2. Choosing distractors with replacement eliminates item-to-item dependencies that may allow subjects to focus on a subset of features once a given distractor has been rejected. For example, if one particular distractor is known to be difficult to reject, subjects may be able to reject it once, then assume it will not recur and focus on "easier" features. Choosing distractors with replacement does, of course, have the drawback of causing some displays to have multiple copies of an object among the distractors and may be better suited to a slower search than ours, in which a serial rejection of distractors (or groups of distractors) more strongly affects search slopes. In the present case, it is possible that slopes were artificially reduced if one assumes that duplications among distractors make search easier because those duplications are more likely in more numerous displays (e.g., 1/32 three-item target-present displays, 10/32 six-item target-present displays, and 28/32 nine-item target-present displays will have at least one pair of identical distractors). However, if this were the case, one would expect steeper slopes when comparing three-item displays with six-item displays than when comparing six-item displays with nine-item displays, because the likelihood of duplicates increases more between the six- and the nine-item displays. A look at the data from Experiments 1 and 2 do not support this hypothesis, and the average slopes based on the three- and six-item displays (13.1 msec/item, based on the combined results of all target-present data in Experiments 1 and 2) are almost identical to those computed for the six- and nine-item displays (13.5 msec/item).

3. Regressions with significant part- and whole-typicality ratings were rerun with the mean rating from all 7 raters. For the part-typicality/other-part model predicting artifact slopes, the colinearity measure and the part-typicality rating were again significant (part typicality, $\beta = .436$, $p = .0018$; colinearity, $\beta = .330$, $p = .0065$), whereas the complexity and other-part ratings were nonsignificant (complexity, $\beta = .209$, $p = .0712$; other part, $\beta = -.011$, $p = .931$). The multiple R for the regression as a whole was .568. For the whole-typicality/other-whole model predicting animal slopes, the colinearity measure and the whole-typicality rating were again significant (whole typicality, $\beta = .474$, $p = .0003$; colinearity, $\beta = .429$, $p = .0001$) and the complexity and other-part ratings were also significant (complexity, $\beta = .212$, $p = .0293$; other part, $\beta = .249$, $p = .0384$). The multiple R for the regression as a whole was .715.

APPENDIX
Target Names for the Color Image Set

Animals	Artifacts
badger	ashtray
bear	book
bobcat	hairbrush
lion	calculator
cheetah	calculator
chipmunk	video camera
crab	chair
dolphin	clamp
tropical fish	clock
tropical fish	disk
eagle	CD ROM drive
kangaroo	dolly
male lion	headphones
lizard	helicopter
orangutan	allen wrench set
ostrich	shoe
porcupine	ice scraper
panther	circuit board
frog	binder
rabbit	stereo amplifier
mouse	pen
seal	needle nose pliers
shark	pliers
snake	podium
snake	car
spider	scale
tarantula	screwdriver
stingray	flour sifter
goat	car
tiger	staple remover
zebra	briefcase
otter	sunglasses