

Attending to the parts of a single object: Part-based selection limitations

SHAUN P. VECERA

University of Iowa, Iowa City, Iowa

MARLENE BEHRMANN

Carnegie Mellon University, Pittsburgh, Pennsylvania

and

JOSEPH C. FILAPEK

University of Iowa, Iowa City, Iowa

Studies of object-based attention have demonstrated poorer performance in dividing attention between two objects in a scene than in focusing attention on a single object. However, objects often are composed of several parts, and parts are central to theories of object recognition. Are parts also important for visual attention? That is, can attention be limited in the number of parts processed simultaneously? We addressed this question in four experiments. In Experiments 1 and 2, participants reported two attributes that appeared on the same part or on different parts of a single multipart object. Participants were more accurate in reporting the attributes on the same part than attributes on different parts. This part-based effect was not influenced by the spatial distance between the parts, ruling out a simple spatial attention interpretation of our results. A control study demonstrated that our spatial manipulation was sufficient to observe shifts of spatial attention. This study revealed an effect of spatial distance, indicating that our spatial manipulation was adequate for observing spatial attention. The absence of a distance effect in Experiments 1 and 2 suggests that part-based attention may not rely entirely on simple shifts of spatial attention. Finally, in Experiment 4 we found evidence for part-based attention, using stimuli controlled for the distance between the parts of an object. The results of these experiments indicate that visual attention can selectively process the parts of an object. We discuss the relationship between parts and objects and the locus of part-based attentional selection.

Although any visual scene contains many objects, our visual processing of scenes has a limited capacity. We process only a subset of the objects present in a scene. Attentional mechanisms allow us to restrict processing to a subset of stimuli, to prevent our visual systems from becoming overloaded. Investigations of object-based visual attention recently have been at the forefront of research on attentional processes, demonstrating that, in addition to selecting stimuli on the basis of spatial position (e.g., Eriksen & Hoffman, 1973; Posner, 1980), humans also can select objects, suggesting that there is a limit to the number of objects that are processed simultaneously (e.g., Baylis & Driver, 1992, 1993; Behrmann, Zemel, & Mozer, 1998; Duncan, 1984, 1993a, 1993b; Egly, Driver, & Rafal, 1994; Kramer & Jacobson, 1991; Moore, Yantis, & Vaughan, 1998; Vecera & Farah, 1994).

In the object-based attention literature, the question of how entire objects are selected has been examined. But objects typically are composed of parts, an observation that has not been addressed by theories of object-based attention. Can attention select the individual parts of an object? Object-based attention research suggests, for example, that when attending to a cat, the entire cat is attended and selected; a spatially adjacent ball of yarn is not selected. However, object-based attention research does not inform us about the fate of the object's parts. Can attention be restricted to a part of the cat, such as the cat's tail?

Although parts have been overlooked by object-based attention research, parts have been integral to theories of object recognition. Several theories of object recognition hypothesize that recognition involves first determining the parts of an object (i.e., decomposing an object into its parts) and then representing the object in terms of the parts and the spatial relations among those parts. These part-decomposition models of recognition postulate that part recognition precedes object recognition (e.g., Biederman, 1987; Farah, 1991; Hoffman & Richards, 1984; Hoffman & Singh, 1997; Marr, 1982), and a substantial body of empirical evidence supports part decomposition in human visual processing (Braunstein, Hoffman, & Saidpour, 1989; Driver & Baylis, 1995; Siddiqi, Tresness, & Kimia, 1996).

The authors thank Steve Luck, John Henderson, Art Kramer, Maureen Marron, and three anonymous reviewers for constructive comments on this research. Natalie Lonsdale, Julia Hartwig, John McGoldrick, Valerie Seeley, and Laurel Trautwein assisted with data collection and analysis. Correspondence can be addressed to S. P. Vecera, Department of Psychology, E11 Seashore Hall, University of Iowa, Iowa City, IA 52242-1407 (e-mail: shaun-vecera@uiowa.edu).

Computational algorithms for object identification also have stressed the importance of parts for recognition (e.g., Brooks, 1981; Pentland, 1986). Given the importance of parts for recognition, one may ask if parts are similarly important for guiding visual attention within an individual object. Do the mechanisms that underlie object-based attention also operate at the level of parts? Answers to this question can reveal whether parts are attended and, more theoretically important, whether or not there are similarities between object recognition and object-based attention processes.

Before discussing part-based attention, we first must address how *part* is to be defined. Following Palmer (1977), we acknowledge that there is no fundamental difference between parts and whole objects, except for the level of analysis—parts require a finer level of analysis than the whole object. Objects are composed of parts, and parts themselves may be composed of parts; each of these levels can be referred to as a structural unit (Palmer, 1977). We restrict our definition of *part* to the first structural unit below the structural unit of the object. Furthermore, we consider a part to be a structural unit that contributes to or helps define the overall shape of an object (see Hoffman & Singh, 1997). With this definition, a part becomes a set of elements (e.g., edges) that is a subset of an object's elements that shapes the object (see Feldman, 1999, for a mathematically formal statement of this idea). The elements that define parts can be based on (1) stimulus properties such as proximity, closure, or connectedness (Palmer, 1977), (2) stimulus properties such as inflection points that create minima of curvature in an object's outline (Hoffman & Richards, 1984; Hoffman & Singh, 1997), or (3) function or category cues (e.g., Schyns & Murphy, 1994). In the present experiments, we restrict our conceptualization of a part to those subelements of an object defined on the basis of connectedness and minima of curvature.

There are two interrelated provisions regarding our definition of a part. First, surface features, such as color, brightness, or texture, are not considered parts under our definition, because they are not critical for defining an object's overall shape (also see Wolfe & Bennett, 1996, who use *form* to refer to local attributes that define an object's overall shape; our definition of *part* may be comparable with Wolfe & Bennett's definition of *form*). Second, although parts and objects differ in their hierarchical levels, not all subcomponents of a global shape should be considered *parts*. Following the object recognition literature, our definition of a part requires that the part be involved in forming or determining the shape of the whole object (see Hoffman & Singh, 1997, for a review).

We now turn to results from the visual attention literature that are relevant to the existence of part-based attention. Perhaps the most obvious result from the attention literature is the finding that attention can be allocated to different spatial scales of a stimulus (e.g., Kimchi, 1992; Navon, 1977). This *global-local* research has demonstrated that the local subelements of a display can be at-

tended and selected within a hierarchical stimulus in which a large letter is made up of smaller letters. Although this research has demonstrated that global stimuli are processed faster than local stimuli and that the global level influences processing of the local level, the local elements nonetheless can be attended and recognized (e.g., Kimchi, 1998; Navon, 1977), a result that appears to demonstrate part-based selection. However, results from the global-local paradigm do not demonstrate attention to parts according to our definition of part: In a standard global-local experiment, all of the parts at the local level are arbitrary objects that do not correspond to the elements that compose the shape of the global letter. For example, using minima of curvature and connectedness cues, many people would decompose a capital letter H into three parts—a vertical line segment on the left, a vertical line segment on the right, and a horizontal segment that bisects and connects the two vertical segments. The local elements (i.e., the small letters in a hierarchical display) do not correspond to the parts of the global object that form the shape of that global object.

Another line of research that appears relevant for studies of part-based selection comes from visual search tasks conducted by Wolfe and colleagues (e.g., Bilsky & Wolfe, 1995; Wolfe, Friedman-Hill, & Bilsky, 1994), who were interested in whether visual search performance was different when participants searched for two parts, rather than searching for a part and a *whole*. Participants exhibited a more efficient visual search when looking for a part-whole relation than when looking for a part-part relation. For example, it was easier to search for a red house with yellow windows (whole-part) than to search for a red and yellow house (part-part). As with the global-local results, however, the whole and part features examined in these studies tended to be surface features (e.g., color or texture) that were not critical for establishing the shape of the object. Wolfe's own interpretation of these results does not appeal to part-based attention, but to preattentive grouping processes. Participants may find it easier to group *wholes* together and process them as a single group, as in processing all of the red houses in a display simultaneously (Wolfe et al., 1994). Other results from Wolfe's laboratory (e.g., Bilsky & Wolfe, 1995) may be more applicable to part-based attention, however, and we will return to these results in the General Discussion section.

Finally, a recent study by Watson and Kramer (1999) provides what is perhaps the most relevant evidence for part-based attentional selection. Watson and Kramer were not attempting to demonstrate part-based selection; instead, they were interested in the bottom-up cues to which object-based attention is sensitive. In their third experiment, Watson and Kramer included a manipulation of the parts of an object. Participants viewed objects (wrenches) that had well-defined parts and wrenches with poorly defined parts; part goodness was defined by minima of curvature. The results from this experiment indicated that the goodness of the parts influenced the allocation of object-

based visual attention: Objects with well-defined parts showed a smaller object-based attention effect than did objects with poorly defined parts. A straightforward interpretation of this result is that when participants attended an object with well-defined parts, attention was restricted to one of the two parts, making it difficult to report the two relevant attributes of the object. When participants attended an object that contained poorly defined parts, attention was not restricted to the parts, making it easier for object-based attention to be spread across the entire object.

Despite the importance of demonstrating that object-based attention is sensitive to bottom-up cues such as minima of curvature, Watson and Kramer's (1999) results do not provide a conclusive demonstration of part-based attention (although Watson and Kramer should not be faulted for this, because their focus was on object-based attention). Their findings may not have been due to part-based attention but, instead, to grouping on the basis of common region (Palmer & Rock, 1994): The object that had poorly defined parts occupied a larger region. Participants could have been encouraged to group the two attributes together on the basis of this larger common region. Because two physically different displays (objects with well-defined or poorly defined parts) were compared, the results could be due to any stimulus-level differences between the displays; no mention need be made of *parts* to interpret these findings. Second, in this task there is no comparison between dividing attention across two parts and focusing attention within a single part and, therefore, no measure of part-based attention. Watson and Kramer's results suggest that part goodness may influence object-based attention, but their results do not speak to part-based attention.

In the present research, we apply our definition of a part to a focused- versus divided-attention task. We ask whether visual attention is limited in the number of parts that can be processed from a single object, just as attention is limited to the number of objects that can be processed at any one time. We explored part-based attention with displays containing single, multipart objects, shown in Figure 1A. To determine whether attention could be selectively restricted to a single part of an object, participants reported two attributes that were either on the same part or on different parts. Thus, we can examine the attentional costs associated with dividing attention across the parts of an object. Our procedure also permits us to compare physically similar displays in the *same-part* and *different-parts* conditions to prevent stimulus differences from influencing attentional allocation.

Our procedure was similar to that developed by Duncan (1984) for studying object-based attention. Our stimulus objects consisted of three parts—a main body and two crossbars. We verified that participants viewed these objects as having a main body and two crossbars by asking 12 naive participants to decompose the objects into their “most naturally occurring parts” (see Palmer, 1977, and Siddiqi et al., 1996, for this procedure). All of the par-

ticipants decomposed the stimuli into the main body and the two separate horizontal crossbars, consistent with parsing the parts of an object at minima of curvature (see Vecera, Behrmann, & McGoldrick, 2000).

Each of the two crossbar parts were given two attributes: The shorter, pronged bar could be short or long and have upward- or downward-turned prongs; the longer, straight bar could be short or long and have a gap on the left or the right side. The participants reported pairs of these attributes that were located on either the same part or on different parts (see Duncan, 1984, 1993a, 1993b; Vecera & Farah, 1994). For example, reporting the length of the pronged bar and the direction of the prongs would involve reporting attributes from the same part; reporting the length of the straight bar and the direction of the prongs would involve reporting attributes from different parts. If attention can be restricted to the parts of an object, participants should be more accurate reporting attributes from the same part than from different parts.

One salient difficulty with this divided-attention procedure is that attributes on individual parts of an object are likely to be physically closer to one another than attributes on different parts of an object. Thus, any part-based effects observed in our paradigm could be due entirely to spatial proximity, not to part-based processing. Because of the strong coupling between a part's attributes and its spatial location, it is difficult to create stimuli in which attributes are the same distance from one another when they appear on the same part and when they appear on different parts (see Duncan, 1984). Therefore, to examine the possible role of spatial proximity in our study of part-based selection, we included a spatial manipulation in our stimuli: The parts of the objects could be either near one another or far from one another (Figure 1A). Although it is difficult to alter the spatial positions of attributes when they appear on the same part, our spatial manipulation allows us to alter the spatial positions of attributes when they appear on different parts. If part-based attention was due entirely to spatial proximity, the difference between the same-part and different-parts conditions should be small in the *near* displays, as compared with the *far* displays. That is, part (same part vs. different parts) should interact with spatial proximity (near vs. far).

This part \times spatial proximity interaction prediction follows from spatial accounts of visual attention in which adjacent stimuli tend to be selected more readily than distant stimuli (e.g., Eriksen & Hoffman, 1973; Posner, 1980). By contrast, if our results are due to a part-based effect beyond any location-based effect, the distance between the parts should influence performance little, if at all; the difference between the same-part and different-parts conditions would be similar for both near and far displays. Part condition (same part vs. different parts) would show a main effect, but there would be no interaction with distance.

Note that we are attempting to rule out a simple spatial proximity interpretation of our results (e.g., a spatial *spotlight* that is focused on one part at a time). Modified

spatial accounts (e.g., a *grouped array*; see Vecera, 1994; Vecera & Farah, 1994) may ultimately explain our present results. We make no assumptions regarding the shifting or indexing of this simple spatial selection mechanism. There has been extensive debate as to whether focal spatial attention moves in a continuous, analogue manner or if it “jumps” from one location to the next (see Eriksen & Murphy, 1987; Eriksen & Webb, 1989; Kramer, Weber, & Watson, 1997; Kwak, Dagenbach, & Egeth, 1991; Sperling & Weichselgartner, 1995; Vecera, 1997; Yantis, 1988). An interaction between part and spatial proximity could be predicted by either an analogue spatial spotlight or by a jumping spotlight. In the analogue case, the part effect would increase as the spatial distance increased, because of the time required to shift attention a greater distance. In the jumping case, a close spatial proximity of the object's parts could allow some of the unattended part to fall at or near the spatially attended region. As the spatial distance increased, little, if any, of the unattended part would fall at or near the spatially unattended region. On the basis of the accumulation of partial information from an unattended location, participants could be more accurate in reporting attributes from different parts when the parts were close than when they were distant, thereby allowing the part effect to increase as the distance between the parts increased.

EXPERIMENT 1

In Experiment 1, participants viewed multipart objects and reported attributes from these objects. The attributes fell either on the same part or on different parts. On the basis of our previous research (Vecera et al., 2000), we predicted more accurate performance in the same-part condition than in the different-parts condition. We also manipulated the spatial distance between the parts; the parts could be near to one another or far from one another. If part-based attention can be explained with a simple spatial attention process, we should observe a part \times spatial proximity interaction. If there exists a part-based component of attention that operates in addition to the selection of locations, we should observe a main effect for part, but no interaction between part and proximity.

In this experiment, the participants were cued to the two relevant attributes by two small colored arrows that appeared simultaneously with the stimulus object. Although this procedure represents a departure from previous research (e.g., Duncan, 1984; Vecera & Farah, 1994; but see Lappin, 1967), this procedure is advantageous because the participants cannot preset attention to a particular region or part in the display prior to the onset of the stimulus.

Method

Participants. Twenty undergraduate students at Carnegie Mellon University participated for course credit; all had normal or corrected vision.

Stimuli. The stimuli were multipart objects, similar to those developed by Tarr and Pinker (1990; Figure 1A). Each object was composed of a main vertical body with a base and two crossbars. Only the crossbars, one straight and one pronged, varied and were relevant for the participants. The straight part was always longer than the pronged part. The straight and pronged parts appeared equally in the upper and lower positions of the object.

The stimuli were viewed from a distance of approximately 60 cm. The parts were located 0.4 cm (0.38° of visual angle) apart from one another in the *near* condition or 2.4 cm (2.29°) apart from one another in the *far* condition. Each short object (parts near one another) measured 1.62 cm (1.55°) tall, and each tall object (parts far from one another) measured 3.53 cm (3.37°) tall. The straight bar measured 1.66 cm (1.58°) wide when long and 1.23 cm (1.17°) when short. The pronged part measured 0.95 cm (0.91°) wide when long and 0.74 cm (0.71°) when short. The prongs on the shorter part could be oriented upward or downward, and these arms measured 0.3 cm (0.29°) above the height of the bar. The gap on the side of the longer part was created by removing 3 pixels and was located 0.4 cm (0.38°) from the left or right of the midline of the object's main body. The pronged part never contained the gap, and the straight bar never contained prongs. The participants were told to ignore the relative positions of the parts and to focus only on the attributes they were to report. A pattern mask, which measured 2.5 cm (2.39°) wide \times 3.6 cm (3.43°) tall, followed the presentation of the stimulus object.

Two small arrows, one red and one green, each 1 cm (0.95°) in length and located 0.3 cm (0.29°) from the outside edge of the parts, also appeared with the stimuli. These arrows cued the participants about which attributes would be reported from the display (see Lappin, 1967, for a similar procedure). The arrows were presented simultaneously with the object, to diminish spatial precuing effects. When the arrows pointed at a part, both arrows were located on the same side of the object; half of the time the arrows were located to the left of the object, and half of the time they were located to the right of the object.

Procedure. The participants reported pairs of attributes. The attributes might both come from the same crossbar (same-part condition) or from different crossbars (different-parts condition). There were only two possible same-part decisions: (1) straight bar length and gap (whether the straight bar is long or short and whether the gap is on the left or right) or (2) pronged bar length and arms (whether the pronged bar is long or short and whether the arms face up or down). There were two possible different-parts decisions: (1) pronged bar length and straight bar gap (whether the pronged bar is long or short and whether the gap is on the left or right of the straight bar) and (2) straight bar length and pronged bar arms (whether the straight bar is long or short and the arms on the pronged bar point up or down). The two attributes not reported on a trial were held constant, and the participants were aware of this. When not being reported, the gap was held constant on the left side, the arms were held constant in the upward position, and both the shorter and the longer crossbars had a relatively shorter length.

The participants were assigned to one of two groups. One group reported gap side and length (*gap* group), and the other group reported prong direction and length (*prong* group). In both groups, length could be from the same crossbar as the other feature (same-part condition) or from a different crossbar (different-parts condition). The red and green arrows cued the participants as to which attributes to report. One arrow cued either the gap or the prong attribute, and the other arrow cued the length attribute (i.e., it informed the participants about which part's length to report). The red arrow indicated length attribute for half of the participants, and the green arrow indicated this attribute for the other half of the participants. Thus, a participant in the *gap* group in which the green arrow indicated the length attribute would view displays in which the red arrow always pointed at the straight bar that contained the gap and the green arrow would point at either the straight bar or the pronged bar. The green arrow would inform the participant as to which part's length should be reported.

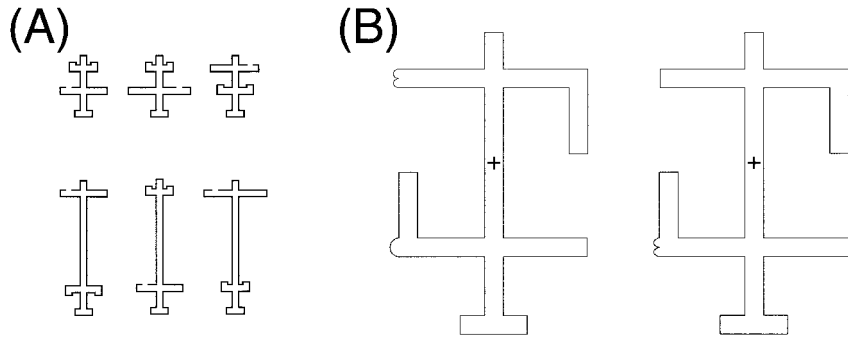


Figure 1. (A) Examples of the stimuli used in Experiment 1. The top row depicts three objects with the parts near to one another, and the bottom row depicts three objects with the parts far from one another. These stimuli show all of the possible attributes: short pronged bar and long pronged bar, upward-turned prongs and downward-turned prongs, short straight bar and long straight bar, and gap on left and gap on right. (B) Examples of stimuli used in Experiment 4, in which the critical features (the *bumps*) can appear on the same part or on different parts. The critical features are the same spatial distance from one another in the same-part and different-parts conditions.

The comparison between the same-part and different-parts conditions was a within-subjects factor. For example, for participants in the prong group, if the arrows pointed at the pronged part, the participant was instructed to report the prong direction (up or down) and the length of the pronged bar (short or long). If one arrow pointed at the pronged part and the other pointed at the straight part, a participant in the prong group would report the prong direction and the length of the straight bar; a participant in the gap group would report the side of the gap (left or right) on the straight part and the length of the pronged part. The participants used the 1 and 2 keys on a standard keyboard to report the first attribute and the 9 and 0 keys to report the second attribute. Prior to starting the experimental trials, the participants completed 12 practice trials drawn randomly from the same-part and different-parts conditions. There were 336 experimental trials; the participants were given a rest midway between these trials. The near and far distances between the parts appeared equally often in a random fashion.

Before the experiment, each participant engaged in a threshold-setting session to determine an appropriate exposure duration. The participants saw two horizontal lines, one shorter and one longer, and a red arrow and a green arrow; the participants reported two attributes from the same line or from different lines. The procedure was the same as that in the experimental condition. The participants' feature judgments (e.g., length and prong direction) and order of report (length first and prong direction second, or vice versa) were exactly the same in the threshold session and in the experimental trials. In the threshold session, the participants completed blocks of 24 trials. After each block, a participant's accuracy was determined, and exposure duration was varied until the participant reached a level of approximately 70% correct performance in two consecutive blocks. The starting exposure duration was 215 msec. The selected exposure duration was then kept constant throughout the experimental session.

Results and Discussion

The mean exposure duration across participants was 123 msec (range = 80–150 msec), which did not differ between the prong and the gap groups [$t(18) = 1.64, p > .10$]. We combined the groups because preliminary analyses indicated no differences between them. The accuracy data were analyzed with a two-way analysis of variance (ANOVA), with part (same vs. different) and dis-

tance (near vs. far) as within-subjects factors. The mean accuracy rates appear in Figure 2A.

There was a significant main effect for part, with higher accuracy in the same-part condition (83.45%) than in the different-parts condition [75.01%; $F(1,19) = 18.34, p < .0004$]. Planned comparisons revealed that accuracy in the same-part condition was higher than accuracy in the different-parts condition when the parts were near one another [$t(19) = 4.06, p < .0007$] and when the parts were far from one another [$t(19) = 3.71, p < .002$]. There also was a significant main effect for distance, with higher accuracy in the near condition (80.76%) than in the far condition [77.69%; $F(1,19) = 11.92, p < .0001$]. These two factors did not interact with one another [$F(1,19) < 1$], indicating that the spatial separation did not influence the cost of dividing attention between parts.

These results suggest that visual attention is limited in the number of parts that can be processed simultaneously. Performance is poorer when attention is divided across different parts of an object than when attention is restricted to a single part, indicating that there is a part-superiority effect analogous to the object-superiority effect observed in several studies. Attention may not only be directed to single objects; it also may be directed to the individual parts of single objects. Importantly, we observe these part-based effects when the location of the attributes (same part or different parts) is unknown in advance of the stimulus presentation. Attentional selection presumably can occur from multiple levels of the visual hierarchy, either the object level or the part level.

Selection by a simple spatial attention mechanism cannot fully explain our results. If spatial selection was the cause of the part-based effect, the difference between the same-part and the different-parts conditions should have been smaller in the near displays than in the far displays. However, the part-based effect was not larger in the far condition than in the near condition, contradicting a purely spatial account of part-based attention. If our re-

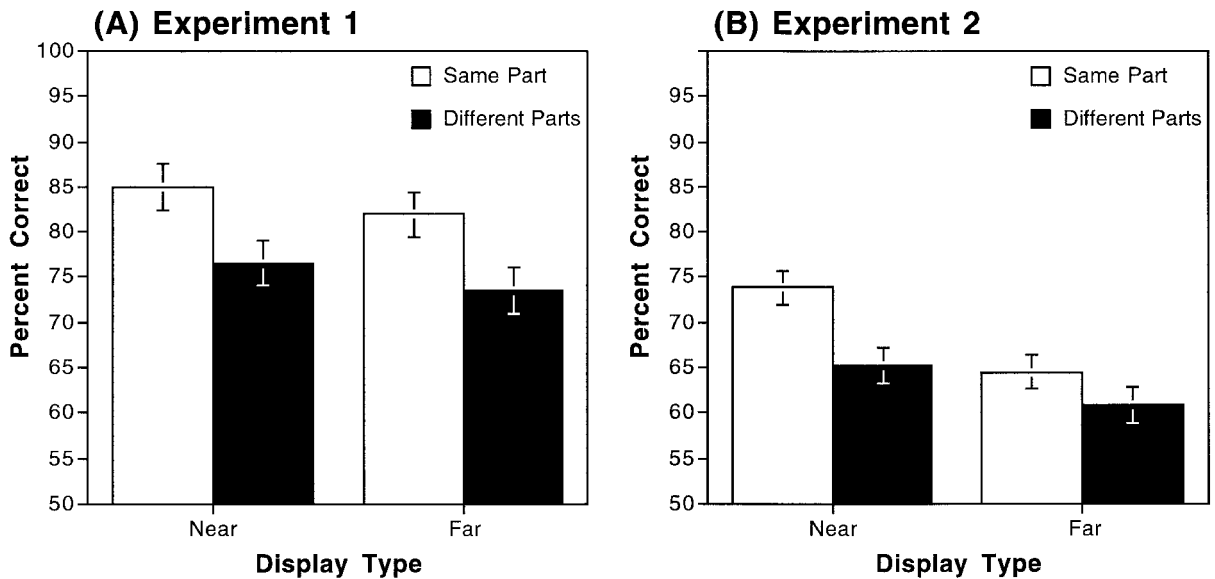


Figure 2. (A) Results from Experiment 1. The participants were more accurate reporting attributes from the same part than from different parts, and this part effect is not statistically larger in the far condition than in the near condition. (B) Results from Experiment 2. Again, the participants were more accurate reporting attributes from the same part than from different parts. The part-based effect is not larger in the far condition. Note: All error bars are within-subjects 95% confidence intervals computed with the subject \times part \times distance interaction mean square (Loftus & Masson, 1994).

sults were due to the attentional selection of the parts themselves and not the locations of the parts, the distance between the parts should have influenced performance little, if at all, which is exactly what we found. There was a highly significant main effect of part condition; the part effect was of similar magnitude in the near and far displays.

Although selective attention appears capable of selecting the individual parts of an object, there are two important issues that warrant discussion. First, could our present results simply be due to a remapping of the part-object hierarchy? That is, might participants be treating the parts as objects? If so, our results could be explained by object-based selection, and no mention of parts or part-based attention would need to be made. Second, what visual processes are the cause of our part-based effect? We have argued that selective attention is the process that exhibits the part-based effect. However, the cuing method used in Experiment 1 may have involved more complex stimulus-response mappings for arrows in the different-parts condition than for those in the same-part condition. Thus, our part-based limitation could have been due to response compatibility or working memory differences between the same-part and different-parts conditions, not to attentional differences between these conditions. We will discuss each of these issues in turn.

We have addressed the remapping of the part-object hierarchy in another study (Vecera et al., 2000, Experiment 2). To demonstrate that part-based selection involves the selection of parts per se, we required the participants to report the attributes of parts from displays

that contained two multipart objects. The participants reported attributes of either the same part or different parts, and these parts were located either on the same object or on different objects. If the part-based effect we have observed is due to remapping the part-object hierarchy, we should fail to find an object-based effect, because the parts would be coded as the relevant "objects" and the global object would not be represented. However, in addition to a part-based effect, we found that the participants were more accurate when the attended attributes were on a single object than when they were on different objects. The participants never knew whether the attributes would be on a single object or on two objects, thereby preventing them from knowing the appropriate level of analysis (part vs. object) in advance of the stimulus presentation. Thus, our results from this part-based attention task appear to be due to attentional selection of parts per se, and not to a remapping of parts to an object level (Vecera et al., 2000). Although our previous results indicate that attention to parts does not involve a remapping of the part-object hierarchy, we would not deny the possibility that the same attention mechanism may give rise to both object- and part-based effects.

Although Experiment 1 appears to demonstrate part-based attentional selection, our use of the colored arrows to cue the participants to the relevant attributes may raise problems. The part-based attentional cost that we observed might have been caused by the stimulus-response mappings between the same-part and different-parts conditions. In the same-part trials, the participants knew exactly which attributes to report, because both colored arrows

pointed to the same part; there was no decision error or memory load that would hinder performance. By contrast, in the different-parts condition, the participants had to use the colors of the arrows to determine which attribute should be reported for each part; the participants had to remember, for example, that the red arrow indicated the length attribute and that the green arrow indicated the gap attribute. Therefore, in the different-parts condition, there was possibly a greater decision error or an increased memory load that was absent in the same-part condition. The results of Experiment 1 could have been caused by increased decision error or an increased memory requirement in the different-parts condition.

We addressed this issue in Experiment 2. Our previous research suggested a procedure for eliminating the decision error or memory load differences between the same-part and different-parts conditions: We informed the participants which attributes to report at the beginning of a block of trials (Vecera et al., 2000). This is a standard procedure used in the object-based attention literature (e.g., Duncan, 1984; Vecera & Farah, 1994). To rule out decision or memory demands, the arrow cues were not used in Experiment 2. Instead, the participants were required to remember the two attributes they were reporting in a block of trials, thereby equating the decisions and memory requirements for the same-part and different-parts conditions.

EXPERIMENT 2

Method

Participants. The participants were 22 undergraduate students at the University of Iowa, who participated for course credit. All had normal or corrected vision.

Stimuli. The stimuli were multipart objects identical to those used in Experiment 1. The arrows used in Experiment 1 were not presented, however.

Procedure. The procedure was similar to that in Experiment 1, with the following exceptions. The first procedural change, employed to streamline the testing session, was the use of a single exposure duration for the object stimuli. Instead of varying the exposure duration for each participant individually, the object appeared for 100 msec for each participant. This exposure duration was chosen on the basis of our previous studies using similar tasks (e.g., Vecera et al., 2000; Vecera & Farah, 1994) and pilot data. The exposure durations of the fixation and mask were identical to those used in Experiment 1.

The second procedural change was in the instructions to the participants regarding the attributes to report. The participants were told the two attributes to report at the beginning of a block of trials (see Duncan, 1984, and Vecera & Farah, 1994, for this procedure). An instruction screen appeared at the beginning of each block to inform the participants about the attributes to report and the order in which to report those attributes. The instruction screen also provided examples of the stimulus objects and the reported attributes. In both the same-part and the different-parts conditions, the participants had to remember which two attributes they were reporting in a specified block of trials.

Using this procedure, the participants reported all possible pairs of attributes. Each participant completed eight blocks of trials, each containing 32 trials, 16 in the near condition and 16 in the far con-

dition. Because the same-part condition had only two possibilities (both attributes on the pronged bar or both attributes on the straight bar) and the different-parts condition had four possibilities, the participants performed each of the same-part reports twice, in order to equate the total number of trials between the same-part and different-parts conditions. In previous research (Vecera et al., 2000), we reported that repeating the same-part conditions in this manner does not inflate or overestimate the part-based attention effects in this task.

The eight blocks were counterbalanced across participants to prevent any order effects. The order of report of the attributes also was counterbalanced across participants (e.g., if one participant reported gap side first followed by prong direction, another participant would report the same two attributes in the opposite order). The dependent measure was the percentage of attributes correctly reported on the same part and on different parts.

Results and Discussion

Because of the preestablished exposure duration, 6 participants responded near chance (average accuracy of 49.25%), and these participants were excluded from the analyses. The remaining 16 participants' data were analyzed with a two-factor, within-subjects ANOVA, with part (same vs. different) and distance (near vs. far) as factors.

The mean accuracy rates appear in Figure 2B. There was a main effect for part [$F(1,15) = 24.19, p < .0002$], with higher accuracy in the same-part condition (69.17% correct) than in the different-parts condition (63.01% correct). Planned comparisons revealed that accuracy in the same-part condition was higher than accuracy in the different-parts condition when the parts were near one another [$t(15) = 5.30, p < .0001$], and when the parts were far from one another [$t(15) = 2.50, p < .03$]. These findings replicate the part-based attention effect found in Experiment 1 and reported by Vecera et al. (2000). There also was a main effect of distance, with higher accuracy in the near condition (69.51% correct) than in the far condition [62.67% correct; $F(1,15) = 33.69, p < .0001$]. Finally, there was an interaction between the part and distance factors [$F(1,15) = 7.39, p < .02$], although this interaction was *opposite* to that predicted by a spatial proximity model of selection. Inspection of Figure 2B shows that the part effect (i.e., the difference between the same-part and different-parts conditions) was slightly smaller in the far condition than in the near condition.

The results of Experiment 2 indicate that the part-based attentional effect cannot be attributed to other processes that might differ between the same-part and different-parts conditions, such as working memory differences or the stimulus-response mapping between the arrow cues and the responses. Our findings in Experiment 2 support an attentional interpretation of the superior performance in the same-part condition, as compared with the different-parts condition.

The interaction we observed between distance and part appears to be due to the high accuracy in the near, same-part cell. The near, same-part cell may exhibit higher accuracy relative to the other conditions because of joint effects of part-based attention and retinal acuity: Focus-

ing on a single part that is near fixation may provide for high levels of accuracy in reporting the attributes of that part. Also, this interaction is not replicated in Experiment 1, and we have no theoretical interpretation of the underadditive interaction observed in Experiment 2. The theoretically important result is that the part-based attention effect was not larger in the far condition than in the near condition.

The results of Experiments 1 and 2 indicate that selective attention can be restricted to the parts of an individual object. We see qualitatively similar effects when the attributes are cued concurrently with the presentation of the stimulus object (Experiment 1) and when the attributes are known in advance of the stimulus (Experiment 2). Neither of the present experiments exhibited a larger part effect with increasing spatial proximity. Thus, in our attribute report task, attending to parts may not involve attending to the locations of the parts.

Of course, the failure to find a larger part-based effect as the distance between the parts increased is a null result, and there are many possible reasons why we did not observe an interaction. The most straightforward possibility is that our distance manipulation was insufficient to observe the operation of spatial selective attention. We should note, however, that there was a sixfold difference in the distance between the parts in the near and far conditions, making it unlikely that our spatial manipulation was insufficient. Nevertheless, before we can accept the results of Experiments 1 and 2 as demonstrating a part-based selection mechanism that does not rely on spatial location, we must demonstrate that our spatial manipulation was adequate. We perform this manipulation check in Experiment 3.

EXPERIMENT 3

In Experiment 3, we required the participants to perform a task that is known to tap spatial selection. If the distance manipulation used in Experiments 1 and 2 was sufficient, we should observe statistically significant effects of distance in Experiment 3. We used the same multipart objects as those used in Experiments 1 and 2. However, we changed the task to a spatial precuing task, similar to Posner's classic task (Posner & Cohen, 1984) that involves shifts of covert spatial attention. In our experiment, the participants were asked to identify the color (red or magenta) of a small target that appeared within one of the two parts of the object. Prior to the appearance of the target, one of the object's parts was cued by having it brighten, as is shown in Figure 3. The target could be either validly cued, in which case the cue and target appeared in the same spatial region (i.e., same part), or invalidly cued, in which case the cue and target appeared in spatially separate regions (i.e., different parts). If spatial attention underlies selection in this task, the separation of the parts should influence performance. Specifically, the difference between reaction times (RTs) to validly and invalidly cued targets

should be smaller in the near condition than in the far condition. However, if our spatial manipulation was insufficient to observe spatial precuing effects, the RT difference between validly and invalidly cued targets would be similar for near and far displays.

Method

Participants. The participants were 16 University of Iowa undergraduates, who participated for course credit. All had normal or corrected vision.

Stimuli. The stimuli were identical to the multipart objects used in Experiment 1, with one exception: The gap that could appear on the straight object was filled in to create a smooth contour for the object. Only the objects that had long parts were used to increase the total display width. The pronged part had upward-turned prongs on half of the trials and downward-turned prongs on the other half of the trials.

Procedure. All the stimuli were presented via a Power Macintosh G3, viewed from a distance of approximately 60 cm. The participants' task was to discriminate a small target that appeared after a spatial precue. An individual trial began with an 800-msec presentation of a fixation cross, which was followed by a stimulus object that appeared for 500 msec. One of the parts was then cued; the cue consisted of brightening one of the parts (Figure 3). The brightening was achieved by having the thickness of the line that defined the part change from 1 pixel wide to 4 pixels wide. The part's width changed for 50 msec. Finally, the target, either a red or a magenta patch, appearing at the intersection of the part and the main body of the object. The participants responded to the target by pressing the left key on a Cedrus button box if the target was red and the right key if the target was magenta. The target was presented for 100 msec in order to keep the total cue and target exposure duration below 200 msec, to avoid eye movements toward the cued location. The participants were instructed to respond quickly but accurately.

The precue was predictive of the target's location; 75% of the trials were validly cued, in which case the cue predicted the target's location, and 25% of the trials were invalidly cued, in which case the cue did not predict the target's location. The participants received 128 practice trials, which were not analyzed, followed by eight blocks of 128 trials each, for a total of 1,024 trials. Half of the trials were from the *near* condition, and half were from the *far* condition.

Results and Discussion

Only correct RTs were used in the analyses. The participants' median RTs for each condition were analyzed with a two-factor within-subjects ANOVA, with cue type (valid vs. invalid) and distance (near vs. far) as factors.

The mean RTs and error rates appear in Figure 4. There was a significant main effect for cue type, with faster RTs to validly cued targets (344.01 msec) than to invalidly cued targets [361.56 msec; $F(1,15) = 24.10, p < .0002$]. Planned comparisons demonstrated that validly cued targets were responded to more quickly than invalidly cued targets in both the near displays [$t(15) = 3.62, p < .003$] and the far displays [$t(15) = 4.40, p < .0005$]. There was also a main effect of distance, with faster RTs to targets appearing in near displays (345.34 msec) than to targets appearing in far displays [361.13 msec; $F(1,15) = 25.29, p < .0001$]. Most important, there was a significant interaction between validity and distance [$F(1,15) = 5.68, p < .04$]. The participants demonstrated a larger validity effect in the far displays than in the near displays. The par-

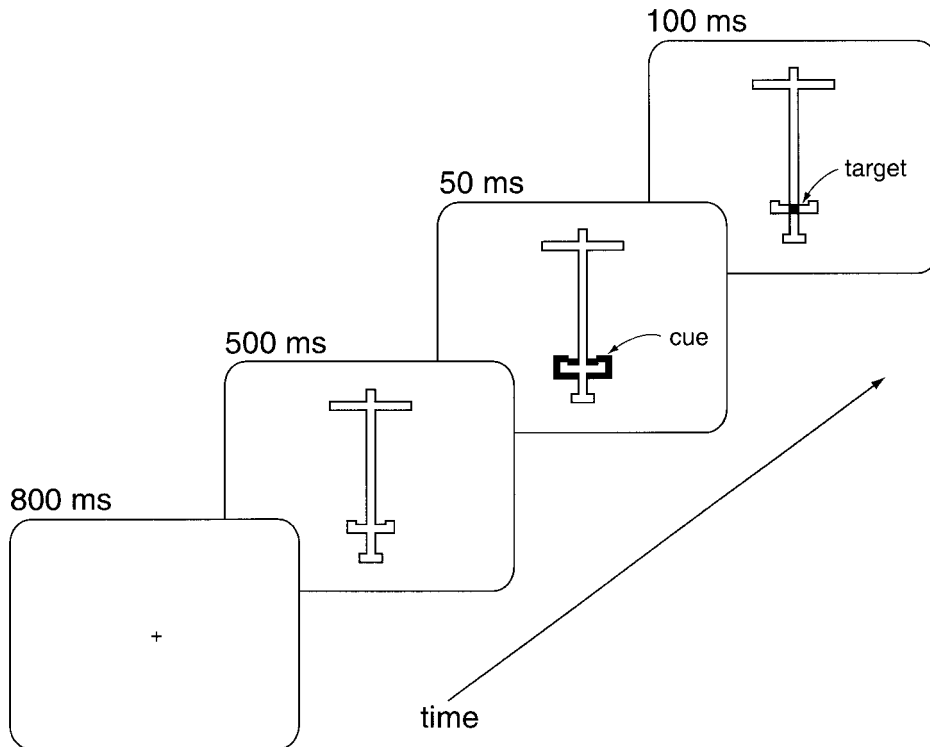


Figure 3. Sequence of events in Experiment 3. Following the appearance of the object, one of the parts was cued by having it brighten. A target then appeared, and the participants had to determine whether the target was red or magenta. This figure depicts a validly cued target. If the straight part had been cued, the target would have been invalidly cued.

participants found it more difficult to switch or divide attention between the cue and the target when they were far from one another than when they were near to one another. The mean error rates were extremely low (less than 4% errors in all conditions) and showed no systematic effects. Analysis of the error data revealed no statistically significant effects (all $ps > .25$).

These results demonstrate larger attentional costs for switching attention between two distant parts than between two closer parts. Thus, the distance between the parts was sufficient to observe the effects of spatial attention. Presumably, if spatial attention alone provided the basis for the part-based effects observed in Experiments 1 and 2, we should have observed a part \times distance interaction. Also, the results of Experiment 3 cannot be explained by appealing to preferential eye movements toward the cued location (overt spatial attention); the cue and the target were presented too briefly to permit eye movements.

As we discussed in the introduction, the presence of a distance effect (i.e., larger attention costs in the far condition than in the near condition) does not necessarily implicate an analogue shift of spatial attention. The distance \times cue type interaction observed in Experiment 3 can be explained without appealing to an analogue shift of attention (see Vecera, 1997): In the near condition, the

close spatial proximity of the object's parts could allow some of the unattended part to fall under or near the spatially attended region. In the far condition, little (if any) of the unattended part would fall under or near the spatially unattended region. On the basis of the partial information accumulated from the unattended location in the near condition, the participants could be faster to detect invalidly cued targets in the near condition than in the far condition. This would appear as a smaller difference between invalidly and validly cued targets in the near condition than in the far condition, which is exactly what we observed in Experiment 3. We make no claims as to the nature of the attentional shift in these studies (analogue or digital); our only claim is that spatial shifts appear to have occurred in Experiment 3, but not in Experiments 1 and 2.

We cannot claim that Experiment 3 confirms an absence of spatial selection in Experiments 1 and 2. The tasks we used in Experiments 1 and 2 differed dramatically from that used in Experiment 3, and it may be reasonable to assume that different tasks have different thresholds for spatial sensitivity. The attribute report task may need a larger manipulation of spatial distance than does the cued detection task to impact participants' performance. Importantly, however, we can conclude that the type of spatial selection indexed by the cued detection task does not

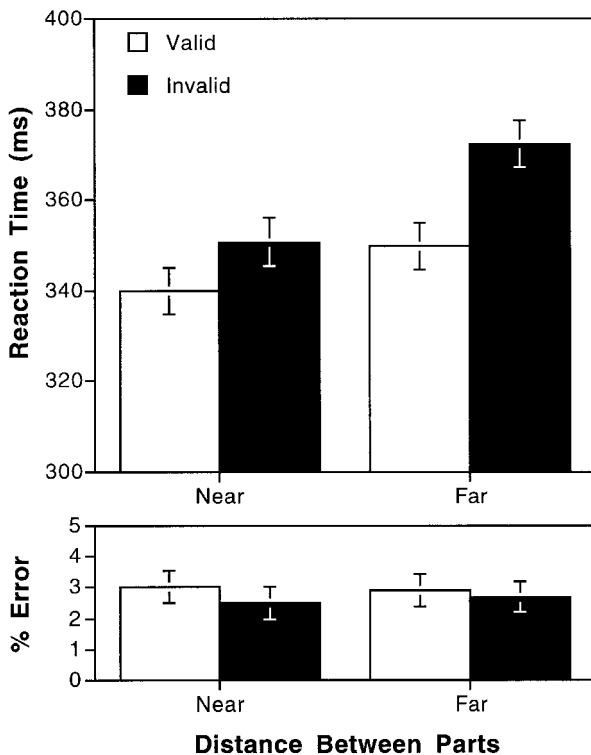


Figure 4. Results from Experiment 3. The cuing condition (valid vs. invalid) interacted with spatial proximity (near vs. far), implicating selection by spatial attention. Note: Error bars are within-subjects 95% confidence intervals computed with the subject \times cue type \times distance interaction mean square.

appear to mediate part-based selection in the attribute report task. Whether performance in the attribute report task is due to selection of parts or to a previously unspecified spatial attention mechanism will require further research.

Because of the inherent difficulties in distinguishing selection of parts from selection of the part's location, in our final experiment we sought converging evidence for part-based attentional selection. Instead of increasing the distance between parts to determine whether distant parts are more difficult to select than closer parts, we held the distance between parts constant. When the spatial distance both within and between parts is constant, a simple spatial selection account predicts no difference in attending to features on the same part or on different parts, because the spatial separation is identical in these conditions. A part-based account, however, predicts that attending to a single part should be easier than attending to different parts.

EXPERIMENT 4

In Experiment 4, the participants viewed multipart objects such as those shown in Figure 1B. The participants were asked to determine whether two features, two small bumps, were the same or different. These features were

independent of the two parts; that is, the small bumps did not distinguish one part from the other. The two features were located on either the same part or different parts, and the features were the same spatial distance from one another in these two conditions. If attention is influenced by the parts that compose the object, features on the same part should be discriminated more efficiently than features on different parts, despite the equivalent separation of the features in the same-part and different-parts conditions. No mention need be made of parts in order to perform this task, unlike in Experiments 1 and 2, in which the parts of the object are implied. Thus, the task used in Experiment 4 may provide a bias-free measure of part-based selection.

Method

Participants. The participants were 16 University of Iowa undergraduates, who received course credit for their involvement. All had normal or corrected-to-normal vision.

Stimuli. The stimuli were multipart geometric shapes similar to those shown in Figure 1B. As with the stimuli used in our previous experiments, naive participants decomposed these stimuli into a main body with a foot and two cross arms. Thus, the two arms are perceived as being separate parts, consistent with Hoffman's proposal of using minima of curvature to delineate parts (Hoffman & Richards, 1984; Hoffman & Singh, 1997).

There were two basic shapes, which differed in the single prongs that were attached to each cross arm. One shape, shown in Figure 1B, had the descending prong on the left side of the object and the ascending prong on the right side of the object. The other shape was the mirror image of the shape shown in Figure 1B. Each object was viewed from approximately 80 cm and measured 16 cm tall and 9.9 cm wide ($11.31^\circ \times 7.05^\circ$, respectively). The width of the bars used to create the objects was 1.0 cm (0.72°); the lines used to draw the objects were 1 pixel wide.

The critical features were curved bumps that were attached to the ends of the parts. There was a single bump that was composed of a single convex protrusion from the part (Figure 1B, top left feature); there also was a double bump that was composed of two convex protrusions from the part (Figure 1B, top right). The distance between the critical features was 9.9 cm (7.05°) in both the same-part and the different-parts conditions.

Procedure. The participants were instructed to report whether the two bumps were the same (i.e., both single bumps or both double bumps) or different (i.e., one single bump and one double bump). No mention was made of the different parts of the object. All the stimuli were presented via a Power Macintosh G3, viewed from a distance of approximately 80 cm. Each trial began with a fixation cross, present for 500 msec. The fixation point was followed by a stimulus object, which was present for 200 msec. The object was followed by a pattern mask, which was visible for 183 msec. The participants indicated their response following the mask; responses were made using a button box. The participants pressed the left key if the bumps were the same and the right key if the bumps were different. Following a response, there was a 200-msec intertrial interval. The participants received four blocks of 40 trials.

Results and Discussion

Only correct RTs were used in the analyses. Participants' median RTs for each condition were analyzed by comparing the RTs between the same-part and the different-parts conditions. Preliminary analyses revealed no systematic effects for same bumps versus different bumps. The mean RTs revealed a statistically significant

difference between the same-part and the different-parts conditions: The participants were faster to discriminate the features when they were located within the same part (484.16 msec \pm 24.25, 95% within-subjects confidence interval; see Loftus & Masson, 1994) than when they were located within different parts [522.03 msec \pm 24.25; $t(15) = 2.35$, $p < .04$].

The accuracy rates paralleled the RT results. The participants were more accurate at discriminating the features in the same-part condition (85.15% correct \pm 2.64) than in the different-parts condition [77.40% correct \pm 2.64; $t(15) = 4.43$, $p < .0005$]. The accuracy data demonstrate that the results from this experiment are not due to a speed-accuracy tradeoff and that part-based attentional effects can be observed in both RT and error measures.

The results of Experiment 4 are consistent with Experiments 1–3 in demonstrating a part-based component of visual attention that cannot be explained by simple spatial selection: The cost associated with switching attention between two parts does not increase with an increasing distance between the parts (Experiments 1 and 2), and a part-based cost is present when attention must move the same distance between parts as within a part (Experiment 4).

One possible explanation of the results of Experiment 4 could involve curve tracing (e.g., Jolicoeur, Ullman, & Mackay, 1986; Pringle & Egeth, 1988; Roelfsema, Scholte, & Spekrijse, 1999; Ullman, 1984), in which attention traces along the curve of the object. In our stimuli, the distance between parts was greater than the distance within a part when measured along the inside of the object, consistent with the curve-tracing hypothesis. We would agree that curve tracing may play a role in some forms of visual selection, including part-based selection. However, a key difficulty with this hypothesis is that it does not explain the results from Experiments 1 and 2, in which we found no effect of spatial distance. The curve-tracing hypothesis would predict increased RTs and errors as the distance between the parts increased (see Jolicoeur et al., 1986, for results consistent with this prediction), which we did not observe in our earlier experiments. Thus, although curve tracing may play a role in many forms of attentional selection, curve tracing alone cannot explain all of the results we have presented. Further experimentation would be required to understand the specific relationship between part attention processes and curve-tracing processes.

GENERAL DISCUSSION

In four experiments, we investigated whether visual attention can selectively process individual parts of a single object. Experiments 1 and 2 showed part-based effects on the division of attention; the participants were more accurate reporting attributes of the same part than reporting attributes of different parts. This part-based effect was not entirely due to the spatial distance between the parts, indicating that the part-based effect could not be

explained only by spatial selective attention. Also, in the first two experiments, a part-based effect was observed when the reported attributes were cued simultaneously with the onset of the stimulus object (Experiment 1) and when the reported attributes were known to each participant in advance of the stimulus object (Experiment 2). In contrast, Experiment 3 showed a spatial selection cost with the same stimulus displays in a spatial precuing task. The results of the third experiment indicate that the spatial manipulation employed in our stimuli was sufficient for observing selection by spatial attention. Finally, in Experiment 4, we found converging evidence for part-based selection. When the distance between two task-relevant features was identical in the same-part and different-parts conditions, the participants were nevertheless faster and more accurate to attend the features on the same part than on different parts.

These results, in conjunction with those of Vecera et al. (2000), suggest a part-based component of visual selective attention. Attention can select the parts of a single object, and selection is more efficient when attention is restricted to a single part than to multiple parts. However, one important issue regarding our results is whether our stimulus objects were processed as a collection of unbound features or parts and not processed as multipart objects. Our results could be explained by the fact that our task required the participants to code separate parts, and coding separate parts may have biased the participants' attentional allocation. Thus, our effects may be due to the strategic deployment of attention, not to the influence of parts on attention. Part-based attentional costs may occur only with stimuli in which participants preferentially process parts and do not represent the objects created by the parts.¹ As was discussed earlier, we addressed this potential criticism in a previous paper (Vecera et al., 2000), in which participants reported attributes of parts from displays containing multiple objects. The use of multiobject displays permitted us to examine both part-based attentional effects and object-based attentional effects. Our previous results replicated the part-based effect we have reported here and also demonstrated an object-based effect. This object-based effect occurred along with a part-based effect within the same blocks of trials in which participants did not know whether they were required to attend to the entire object or to a part. Thus, we observed both the part-based and the object-based attentional effects concurrently. These results indicate that our stimulus objects were perceived as objects per se and not as a conglomeration of parts or features. Furthermore, the results of Experiment 4 in the present paper suggest that part-based effects can be observed in a task that does not require parts to be mentioned explicitly.

We also have explored several alternative explanations of the part-based effect (Vecera et al., 2000). For example, we asked whether preferential practice effects for the same-part condition could have produced the part-based effect. A reanalysis of our data indicated that the part effect was not abolished when the first same-part blocks

were compared with the different-parts blocks. We also addressed the differences between accuracy of reporting the different attributes, such as higher accuracy reporting the prong direction than the gap side. Although the attributes differed in their perceptual salience, overall the attributes were reported more accurately in the same-part condition than in the different-parts condition. Thus, part-based attention does not appear to be dependent on the types of attributes participants report. This conclusion is consistent with Duncan's (1984, 1993a, 1993b) results, in which he finds an object-based component of attention that does not appear to be dependent on the specific types of attributes reported (e.g., brightness, texture, tilt of a line, etc.).

Given that part-based visual attention appears to be due to attention's operating on parts themselves, there are several theoretical issues that we must address. First, how does part-based attention relate to other results in the visual attention literature? Second, how does part-based attention relate to the part decomposition required for object recognition? Third, what is the fate of the entire object's identity when attention is restricted to the part of that object? Fourth, what is the locus of part-based attentional selection—does it occur early in the visual processing hierarchy, or does it occur late in this hierarchy? We briefly discuss each of these issues in turn.

The first issue for discussion concerns the relationship between our demonstrations of part-based attention and previous results from the attentional literature. Our results make a novel contribution to the attention literature, because previous research has not studied attentional selection of the parts that define an object's shape and because previous research has not examined selective attention to a single part versus divided attention across multiple parts. There is a relevant empirical result from the visual search literature that is consistent with our part-based attentional limitation. Bilsky and Wolfe (1995) reported that visual search for a part-whole relationship was more efficient than search for a part-part relationship, using stimuli that were closer to *parts*, as we have defined them. For example, one of the search tasks was to look for a snowman with a big body and a little head (part-part condition) or with a big body and fine checkerboard pattern on the body (part-whole condition; see Figure 1 in Bilsky & Wolfe, 1995). The distinction between part-whole and part-part searches may correspond to a difference between a within-part search and a between-parts search, making Bilsky and Wolfe's results appear to be consistent with our part-based attentional limitation.

Our second theoretical question centers on how part-based attention relates to the part decomposition required for some object recognition processes. Our results are compatible with theories of object recognition that require that objects be decomposed into parts prior to recognition (e.g., Biederman, 1987; Marr, 1982). Part decomposition occurs in our stimulus objects, and parts influence the allocation of attention. Our research has not addressed the direction of the influence between decom-

position and attention. On the one hand, part decomposition may influence both object recognition and visual attention. Indeed, decomposing an object into its parts may require object-based visual attention, which would suggest that object-based attention is required for object recognition. On the other hand, part-based attention and part decomposition may involve the same processes, and *attention* need not be invoked to explain the present results. Part-based costs may spontaneously emerge from a system that decomposes objects into their parts on the basis of image cues (e.g., minima of curvature). Additional studies will be required to resolve such issues.

The question arising from our third theoretical consideration is, how can an entire object be represented (or recognized) when attention is directed selectively to a part? It seems unlikely that object processing would cease when attending to an individual part. We do not seem to lose the identity of an object when attending to the parts of that object. How can part-based attention be reconciled with whole-object processing? Part processing and object processing need not be mutually exclusive, as indicated by connectionist models of part and object processing. Mozer (1999) has demonstrated that a connectionist network can represent hierarchical structure, allowing for the simultaneous representation of an object and its parts. The part representation observed in this model codes for the differences between parts by assigning each part an activation *tag*. Each part has a specific tag, and different parts have different tags. Although the object is *decomposed* into different parts via activation tags, the structure of the object remains represented in the network. This object representation could allow object recognition to occur despite the parts' being differentially represented.

Finally, our fourth theoretical question concerns the locus of part-based selection. Are parts selected early or late in the visual-processing stream? Although our present results are consistent with selection from a structural description of an object that does not represent the precise retinal location of objects and parts, we remain neutral about the visual representations from which attention might select in our paradigm. Because the attributes may be coded with respect to the part (e.g., the gap is coded to the left or right of the crossbar), selection could occur from a later, object-based structural description that represents objects and parts with little reliance on the retinal or spatiotopic locations of the objects or parts (although such a representation would code parts with respect to the object, as in coding a person's head as being located above the torso; see Marr, 1982). By contrast, selection could occur early in the visual-processing hierarchy, mediated through an array-format (or spatiotopic) representation. Parts could be coded in terms of their spatial locations, and parts could be distinguished from one another by image-level cues, such as concave cusps or minima of curvature (Braunstein et al., 1989; Hoffman & Richards, 1984; Hoffman & Singh, 1997). These early-level, bottom-up cues could then direct the allocation of visual attention.

The present results follow earlier results (Vecera & Farah, 1994) in demonstrating that simple space-based costs do not arise during attribute-reporting tasks in which the attributes can be coded with respect to an object-centered (or part-centered) reference point. Although this position has been somewhat controversial (see Kramer et al., 1997; Lavie & Driver, 1996), we would not argue that there is a single attentional bottleneck for visual attention. Instead, attentional limitations can arise from different levels of the processing hierarchy, and the type of task performed by participants may influence the representation underlying selection (Vecera & Farah, 1994). In masked displays, such as those we have used, the primary processing limitation may occur at the level of visual working memory (Vogel & Luck, 2000). Because visual working memory represents stimuli in object-centered coordinates, spatial selection may not occur. Our results are important because they demonstrate that attentional selection need not be mediated by spatial proximity in all tasks. Whether a modified spatial selection mechanism, such as a grouped array, could explain our results awaits further research. However, from our present results, it is clear that the organization of visual information at different levels of a stimulus hierarchy can guide the allocation of visual attention.

REFERENCES

- BAYLIS, G. C., & DRIVER, J. (1992). Visual parsing and response competition: The effect of grouping factors. *Perception & Psychophysics*, **51**, 145-162.
- BAYLIS, G. C., & DRIVER, J. (1993). Visual attention and objects: Evidence for hierarchical coding of location. *Journal of Experimental Psychology: Human Perception & Performance*, **19**, 451-470.
- BEHRMANN, M., ZEMEL, R., & MOZER, M. C. (1998). Object-based attention and occlusion: Evidence from normal subjects and a computational model. *Journal of Experimental Psychology: Human Perception & Performance*, **24**, 1011-1036.
- BIEDERMAN, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, **94**, 115-147.
- BILSKY, A. B., & WOLFE, J. M. (1995). Part-whole information is useful in visual search for size \times size but not orientation \times orientation conjunctions. *Perception & Psychophysics*, **57**, 749-760.
- BRAUNSTEIN, M. L., HOFFMAN, D. D., & SAIDPOUR, A. (1989). Parts of visual objects: An experimental test of the minima rule. *Perception*, **18**, 817-826.
- BROOKS, R. A. (1981). Symbolic reasoning among 3-D models and 2-D images. *Artificial Intelligence*, **17**, 205-244.
- DRIVER, J., & BAYLIS, G. C. (1995). One-sided edge assignment in vision: 2. Part decomposition, shape description, and attention to objects. *Current Directions in Psychological Science*, **4**, 201-206.
- DUNCAN, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, **113**, 501-517.
- DUNCAN, J. (1993a). Coordination of what and where in visual attention. *Perception*, **22**, 1261-1270.
- DUNCAN, J. (1993b). Similarity between concurrent visual discriminations: Dimensions and objects. *Perception & Psychophysics*, **54**, 425-430.
- EGLY, R., DRIVER, J., & RAFAL, R. D. (1994). Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General*, **123**, 161-177.
- ERIKSEN, C. W., & HOFFMAN, J. E. (1973). The extent of processing of noise elements during selective encoding from visual displays. *Perception & Psychophysics*, **14**, 155-160.
- ERIKSEN, C. W., & MURPHY, T. D. (1987). Movement of attentional focus across the visual field: A critical look at the evidence. *Perception & Psychophysics*, **42**, 299-305.
- ERIKSEN, C. W., & WEBB, J. M. (1989). Shifting of attentional focus within and about a visual display. *Perception & Psychophysics*, **45**, 175-183.
- FARAH, M. J. (1991). Patterns of co-occurrence among the associative agnosias: Implications for visual object representation. *Cognitive Neuropsychology*, **8**, 1-19.
- FELDMAN, J. (1999). The role of objects in perceptual grouping. *Acta Psychologica*, **102**, 137-163.
- HOFFMAN, D. D., & RICHARDS, W. A. (1984). Parts of recognition. *Cognition*, **18**, 65-96.
- HOFFMAN, D. D., & SINGH, M. (1997). Salience of visual parts. *Cognition*, **63**, 29-78.
- JOLICEUR, P., ULLMAN, S., & MACKAY, M. (1986). Curve tracing: A possible basic operation in the perception of spatial relations. *Memory & Cognition*, **14**, 129-140.
- KIMCHI, R. (1992). Primacy of wholistic processing and global/local paradigm: A critical review. *Psychological Bulletin*, **112**, 24-38.
- KIMCHI, R. (1998). Uniform connectedness and grouping in the perceptual organization of hierarchical patterns. *Journal of Experimental Psychology: Human Perception & Performance*, **24**, 1105-1118.
- KRAMER, A. F., & JACOBSON, A. (1991). Perceptual organization and focused attention: The role of objects and proximity in visual processing. *Perception & Psychophysics*, **50**, 267-284.
- KRAMER, A. F., WEBER, T. A., & WATSON, S. E. (1997). Object-based attentional selection—grouped-arrays or spatially-invariant representations?: Comment on Vecera and Farah (1994). *Journal of Experimental Psychology: General*, **126**, 3-13.
- KWAK, H.-W., DAGENBACH, D., & EGETH, H. (1991). Further evidence for a time-independent shift of the focus of attention. *Perception & Psychophysics*, **49**, 473-480.
- LAPPIN, J. S. (1967). Attention in the identification of stimuli in complex displays. *Journal of Experimental Psychology*, **75**, 321-328.
- LAVIE, N., & DRIVER, J. (1996). On the spatial extent of attention in object-based visual selection. *Perception & Psychophysics*, **58**, 1238-1251.
- LOFTUS, G. R., & MASSON, M. E. J. (1994). Using confidence intervals in within-subjects designs. *Psychonomic Bulletin & Review*, **1**, 476-490.
- MARR, D. (1982). *Vision*. San Francisco: Freeman.
- MOORE, C. M., YANTIS, S., & VAUGHAN, B. (1998). Object-based visual selection: Evidence from perceptual completion. *Psychological Science*, **9**, 104-110.
- MOZER, M. C. (1999). A principle for unsupervised hierarchical decomposition of visual scenes. In M. S. Kearns, S. A.olla, & D. Cohn (Eds.), *Advances in neural information processing systems 11* (pp. 52-58). Cambridge, MA: MIT Press.
- NAVON, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, **9**, 353-383.
- PALMER, S. E. (1977). Hierarchical structure in perceptual representation. *Cognitive Psychology*, **9**, 441-474.
- PALMER, S. E., & ROCK, I. (1994). Rethinking perceptual organization: The role of uniform connectedness. *Psychonomic Bulletin & Review*, **1**, 29-55.
- PENTLAND, A. P. (1986). Perceptual organization and the representation of natural form. *Artificial Intelligence*, **28**, 293-331.
- PETERSON, M. A., & HOCHBERG, J. (1983). Opposed-set measurement procedure: A quantitative analysis of the role of local cues and intention in form perception. *Journal of Experimental Psychology: Human Perception & Performance*, **9**, 183-193.
- POSNER, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, **32A**, 3-25.
- POSNER, M. I., & COHEN, Y. (1984). Components of visual orienting. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 531-556). Hillsdale, NJ: Erlbaum.
- PRINGLE, R., & EGETH, H. E. (1988). Mental curve tracing with elementary stimuli. *Journal of Experimental Psychology: Human Perception & Performance*, **14**, 716-728.
- ROELFSEMA, P. R., SCHOLTE, H. S., & SPEKREIJSE, H. (1999). Temporal constraints on the grouping of contour segments into spatially extended objects. *Vision Research*, **39**, 1509-1529.

- SCHYNS, P. G., & MURPHY, G. L. (1994). The ontogeny of part representation in object concepts. In D. L. Medin (Ed.), *The psychology of learning and motivation* (Vol. 31, pp. 305-349). San Diego: Academic Press.
- SIDDIQI, K., TRESNESS, K. J., & KIMIA, B. B. (1996). Parts of visual form: Psychophysical aspects. *Perception*, **25**, 399-424.
- SPERLING, G., & WEICHSELGARTNER, E. (1995). Episodic theory of the dynamics of spatial attention. *Psychological Review*, **102**, 503-532.
- TARR, M. J., & PINKER, S. (1990). When does human object recognition use a viewer-centered reference frame? *Psychological Science*, **1**, 253-256.
- ULLMAN, S. (1984). Visual routines. *Cognition*, **18**, 97-159.
- VECERA, S. P. (1994). Grouped locations and object-based attention: Comment on Egly, Driver, & Rafal (1994). *Journal of Experimental Psychology: General*, **123**, 316-320.
- VECERA, S. P. (1997). Grouped arrays versus object-based representations: Reply to Kramer et al. (1997). *Journal of Experimental Psychology: General*, **126**, 14-18.
- VECERA, S. P., BEHRMANN, M., & MCGOLDRICK, J. (2000). Selective attention to the parts of an object. *Psychonomic Bulletin & Review*, **7**, 301-308.
- VECERA, S. P., & FARAH, M. J. (1994). Does visual attention select objects or locations? *Journal of Experimental Psychology: General*, **123**, 146-160.
- VOGEL, E. K., & LUCK, S. J. (2000, April). *Selective storage in visual working memory*. Poster presented at the Annual Meeting of the Cognitive Neuroscience Society, San Francisco.
- WATSON, S. E., & KRAMER, A. F. (1999). Object-based visual selective attention and perceptual organization. *Perception & Psychophysics*, **61**, 31-49.
- WOLFE, J. M., & BENNETT, S. C. (1996). Preattentive object files: Shapeless bundles of basic features. *Vision Research*, **37**, 25-43.
- WOLFE, J. M., FRIEDMAN-HILL, S. R., & BILSKY, A. B. (1994). Parallel processing of part-whole information in visual search tasks. *Perception & Psychophysics*, **55**, 537-550.
- YANTIS, S. (1988). On analog movements of visual attention. *Perception & Psychophysics*, **43**, 203-206.

NOTE

1. We would not deny the possibility that part-based attention may emerge through multiple mechanisms, such as a *top-down* processing of parts (see Peterson & Hochberg, 1983, for an example) on the basis of participants' expectancies of what the parts might be versus the *bottom-up* processing of parts on the basis of image-level cues (e.g., minima of curvature). Thus, there may be both an endogenous and an exogenous influence of parts on attentional selection.

(Manuscript received March 23, 1999;
revision accepted for publication June 1, 2000.)