Journal of Experimental Psychology:
Human Perception and Performance
1993, Vol. 19, No. 3, 451–470

0096-1523/93/$3.00

Visual Attention and Objects: Evidence for Hierarchical Coding of Location

Gordon C. Baylis and Jon Driver

In 5 experiments, it was found that judging the relative location of 2 contours was more difficult when they belonged to 2 objects rather than 1. This was observed even when the 1- and 2-object displays were physically identical, with perceptual set determining how many objects they were seen to contain. Such a 2-object cost is consistent with object-based views of attention and with a hierarchical scheme for position coding, whereby object parts are located relative to the position of their parent object. In further experiments, it was shown that in accord with this hierarchical scheme, the relative location of objects could disrupt judgments of the relative location of object parts, but the reverse did not occur. This was found even when the relative position of the parts could be judged more quickly than that of the objects.

Selective attention operates to direct our actions toward individual objects in a cluttered world. Many researchers have suggested that visual attention operates analogously to a spotlight, "illuminating" areas of interest in the visual field (e.g., Broadbent, 1982; Posner, 1980). An alternative to this purely spatial view suggests that the visual world is parsed into objects or groups defined by gestalt principles of organization (e.g., Duncan, 1984; Prinzmetal, 1981) and that attention is directed to these objects rather than to unparsed regions of space.

The object-based view derives support from two types of experiment. On the one hand are experiments which show that it is relatively difficult to ignore distracting information that belongs to the same object or group as relevant information (Baylis & Driver, 1992; Baylis, Driver, & McLeod, 1992; Driver & Baylis, 1989; Kahneman & Henik, 1981; Kramer & Jacobson, 1991). On the other hand are studies which suggest that it is difficult to attend to different objects simultaneously. The present studies are concerned with the latter prediction of the object-based view. Duncan (1984) reported supportive evidence. He presented subjects briefly with two visual objects (an outline box and a line struck through it). Subjects had to make judgments about one or two of the following attributes: the size of the box, the location of a gap in the box, the orientation of the line, and the texture of the line. They were able to make two judgments concerning the same object (e.g., the orientation and texture of the line) simultaneously without loss of accuracy. However, they showed a cost in making two judgments rather than one if the two judgments concerned attributes from different objects (e.g., the orientation of the line and the size of the box). Duncan's (1984) interpretation was that attention selects information about one object at a time.

However, Watt's computational MIRAGE algorithm (1988, pp. 133–136) provides an alternative explanation of Duncan's (1984) finding. This algorithm combines the output of different spatial frequency filters, which are considered to be analogous to those in the early human visual system. By examining MIRAGE's response to Duncan's stimuli, Watt found that information specifying both attributes of the box could be picked up when outputs from all the filters were combined, whereas information about the attributes of the line became available only when the coarser filters were switched out. He therefore suggested that Duncan's data could be explained without recourse to the notion of objects. One might argue instead that visual attention can operate by switching particular spatial frequency channels in or out (see Julesz & Papathomas, 1984, for a related proposal).

In the present experiments, our subjects have to judge information belonging to one or two objects as in Duncan's (1984) studies, and we look for an impairment in performance for the two-object case. In Experiment 2, we examine a situation in which the one- and two-object displays are physically identical. Whether the information is parsed as belonging to one or two objects is manipulated by the subjects' perceptual set. Any cost for the two-object case in these circumstances could not be attributed to stimulus differences between the one- and two-object displays, such as spatial frequency composition.

Our subjects' task was to judge which of two contours, presented on a computer monitor, was lower on the screen. The contours could be parts of a single object or parts from two different objects. Object-based accounts of attention

Gordon C. Baylis, Department of Psychology, University of California, San Diego; Jon Driver, Department of Experimental Psychology, University of Cambridge, Cambridge, England.

This research was supported by U.S. Office of Naval Research Contract N00014–91–J-1735 to Gordon C. Baylis; Jon Driver was supported by a Fellowship from the McDonnell–Pew Center for the Cognitive Neuroscience of Attention at the University of Oregon, and by the British Medical Research Council.

We are grateful to Rich Ivry and Bela Julesz for discussions of this work, to Nicole Shidler for running subjects in Experiment 7, and to John Duncan, Glyn Humphreys, and an anonymous reviewer for their many helpful comments on earlier versions of this article.

Correspondence concerning this article should be addressed to Gordon C. Baylis, Department of Psychology, 0109, University of California, San Diego, La Jolla, California 92093.

(e.g., Duncan, 1984) predict that the judgment should be easier in the single-object case. Although Watt (1988) criticized such accounts, his MIRAGE algorithm also suggests that relative position judgments should be easier for elements that belong to one object or perceptual group, for reasons given below. A similar prediction can also be derived by considering the manner in which the visual system is thought to code "what" and "where" information. We argue that these three different approaches to the same prediction are complementary rather than in competition.

Watt emphasizes that coding visual location is not entirely straightforward despite the inherent spatial organization of sensory input to the visual system. The filters that the visual system is considered to use inevitably create distortions. The error this introduces can be minimized by a process of constraint relaxation, in which the coded position of each contour contributes to coding the position of every other contour. However, this approach would lead to a combinatorial explosion as the number of positions to be derived increases. One possible solution, used by MIRAGE, is to code position hierarchically (see Figure 1). The scene (Figure 1a) is initially parsed into "groups" (which we take to be analogous to candidate objects), and the position of each group is derived (Figure 1b). The positions of elements in each group (which we take to be analogous to object parts) are then described relative to the position of the group in question (Figure 1c). This reduces the number of iterations required to code position because individual elements from one group do not contribute to the coding of elements in another group.

How is the relative position of elements from different groups–objects encoded under this scheme? Watt (1988, p. 95) gives the following example (see Figure 1): "Consider the representation of the relative position of two dots in different groups. The relative positions of the two groups are known from the initial stage of calculation. The position of each dot with respect to its own group is known from the second stage of calculation. Therefore, the position of each dot with respect to the other is known, *albeit indirectly*" (our emphasis). If human vision uses a similar scheme, judging the relative position of contours from different groups or objects should be more difficult than judging relative positions within an object or group, because only the within-object positions are coded directly. Of course, object-based models of attention also predict that between-object comparisons should be harder than within-object comparisons, because only the between-object judgments require attention to more than one object.

Watt (1988) proposed hierarchical position coding on purely computational grounds and provided no direct evidence for its operation in human vision. In this article, we provide experimental tests of the hypothesis that a similarly hierarchical scheme applies to human vision, whereby the positions of object parts (analogous to individual dots in Figure 1) are described relative to the location of the object they belong to (analogous to dot groups in Figure 1). As we have seen, this hypothesis can be derived from Watt's computational considerations. In addition, the hypothesis is consistent with a reformulation of the widely cited distinction between "what" and "where" information in vision.

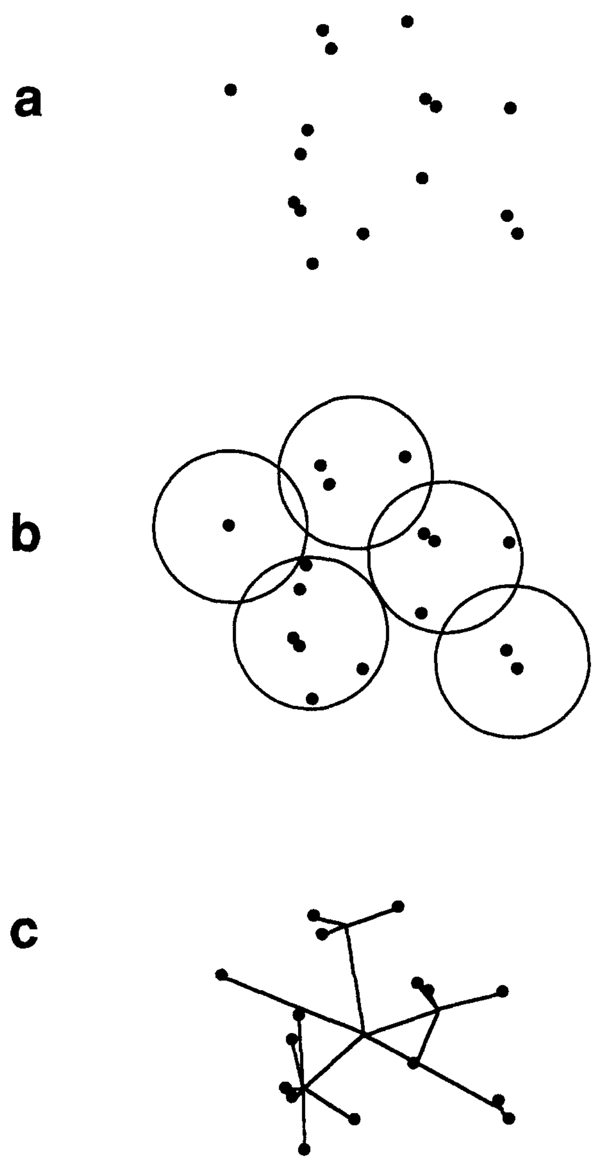

Figure 1. Illustration of the computational economy introduced by coding positions hierarchically. (A set of points [Panel a] is divided into localized groups [Panel b], and the location of each point is coded with respect to the central location of the group it belongs to [Panel c]. Adapted from *Visual Processing: Computational, Psychophysical, and Cognitive Research,* p. 951, by R. J. Watt, 1988, Hove, UK: Erlbaum. Copyright 1988 by Erlbaum. Adapted by permission.)

A range of data from neuroscience and human performance suggest that different components of the primate visual system (the ventral and dorsal extrastriate pathways, respectively) are involved in extracting "what" information and "where" information (e.g., Mishkin, Ungerleider, & Macko, 1983). *Where* information refers to the location of objects within a scene. *What* information refers to object identity. This terminology carries a tacit implication that identity ("what") information is nonspatial, in contrast to "where" information. However, visual identity is usually given by shape information, and shape clearly depends on the relative location of contours. In other words, "what" information is inherently spatial in vision (Farah, Brunn, Wong, Wallace, & Carpenter, 1990).

This point does not undermine the conventional what–where distinction but suggests that it should be restated in terms of a distinction between different kinds of location information. "Where" information is given by the location of objects within a scene. "What" information is given by the relative locations of the constituent contours of an object, regardless of their absolute scene position. Thus, "where" information refers to locations in a scene-based frame of reference, and "what" information refers to locations in an object-based frame of reference (cf. Marr & Nishihara, 1978). According to such a view, the ventral visual pathway represents spatial information in an object-centered manner (Hasselmo, Rolls, Baylis, & Nalwa, 1989).

The hierarchical scheme for position coding that we propose provides both "what" and "where" information naturally. Objects are located relative to each other in a scene-based description (as for the dot groups in Figure 1b). The parts of an object are located relative to each other in an object-centered description (as for the individual dots within a group in Figure 1c). Such a scheme would provide both "where" (Figure 1b) and "what" information (Figure 1c) on our account. However, it contains no explicit coding of the relative position of contours from different objects.

Our first experiment examines the major prediction of this hypothesis: Comparing the location of contours from different objects should be harder than comparing two contours from the same object. The relative position of contours within an object should be coded explicitly in the routine derivation of "what" (i.e., shape) information. However, the relative position of contours from different objects should not be given directly in this description.

Experiment 1

The subjects' task was to make a speeded judgment of the relative height on a computer screen of two apices, formed at angles on the outline of solid shapes. The apices were either parts of a single object (e.g., Figure 2a) or parts of two separate objects (e.g., Figure 2b). Our hypothesis was that the perceptual judgment would be more difficult when apices from different objects had to be compared, even though the between- and within-object comparisons concerned the same points in space. This prediction is derived from Duncan's (1984) object-based account of human attention, Watt's (1988) computational model, and our own conceptualization of the distinction between scene-based and object-centered information in the primate visual system. The fact that such a prediction can be derived from these three very different approaches strengthens its plausibility.

In addition to varying the number of objects to be judged, we also manipulated the vertical offset of the objects in two-object displays. This object offset could be congruent with the apex offset that had to be judged. For example, the object on the left might be lower than the object on the right, whereas the apex on the left object was also lower than the apex on the right object (e.g., Figure 2c). Alternatively, the offset of the objects could be incongruent; for example, the object on the right might be the lower object, whereas the apex on the left object was the lower of the two apices (e.g., Figure 2d). Finally, the object offset could be neutral with respect to the apices; that is, the objects had the same vertical position as each other (e.g., Figure 2b). This condition provides a baseline for comparison with the single-object case.

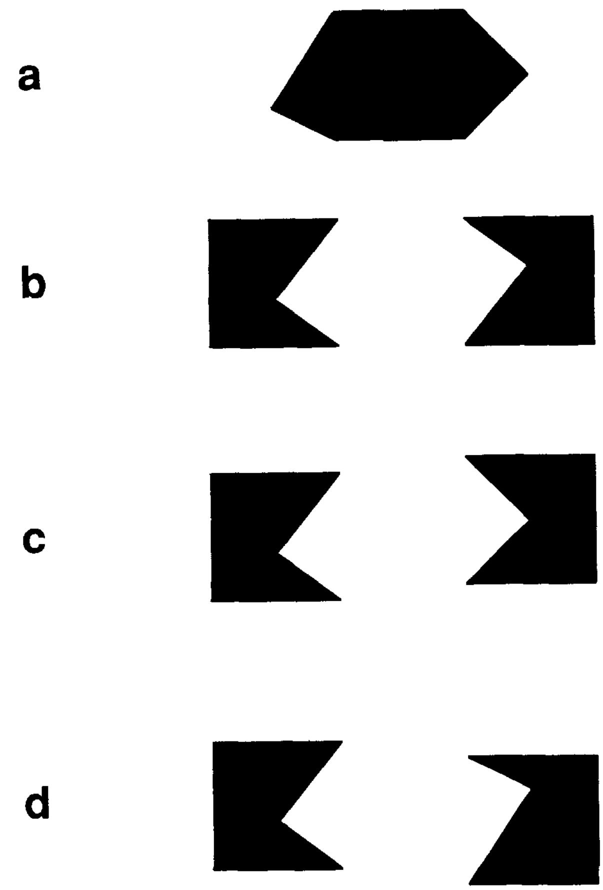

Figure 2. Examples of displays from the four conditions of Experiment 1: (a) single, (b) double baseline, (c) double congruent, (d) double incongruent. (Note that although slight departures from exact scale exist in this and other figures, the relative offsets of apices and objects have been preserved exactly in every case.)

These manipulations of object offset allow us to examine whether the relative location of the objects affects judgments concerning the relative location of parts from different objects. According to our hypothesis they should, as the position of object parts is coded relative to the position of the objects to which they belong. This could produce a conflict when the relative positions of objects and parts disagree (e.g., Figure 2d).

Method

Subjects. The 10 subjects (7 women and 3 men) were lower division psychology undergraduates at the University of California, San Diego. All had normal or corrected-to-normal vision. Subjects received course credit for their participation.

Apparatus and materials. The experiment was conducted on a Compaq 386/20 microcomputer. Stimuli were presented on a color EGA monitor (Samsung). EGA double-page graphics mode was used to ensure that the onset or the offset of stimuli occurred within a single frame.

Example displays are shown in Figure 2. They were presented as white shapes on a black background. Viewing distance was 70 cm with the result that a single-object display (e.g., Figure 2a) subtended a maximum of 4.4° horizontally × 2.6° in height, whereas two-object baseline displays (e.g., Figure 2b) were 6.0° horizontally × 2.6° vertically. The pentagons used in the double condition were 1.2° in maximum width and 2.6° in height. These were formed around the outer contours of imaginary hexagons similar in shape to those in the single condition (exactly the same for the double-baseline condition only). The task was to decide which apex was lower on the screen and respond on key [Z] with the left index finger if the left apex was lower and key [/] with the right index finger if the right apex was lower, using the microcomputer's standard extended keyboard, where [Z] is to the left of [/]. The vertical offset of the apices was always 0.8°. Reaction times (RTs) were recorded in milliseconds.

To preclude the unlikely strategy of comparing the target apices to fixed points on the video monitor or retina, the display elements were positioned on the screen with a degree of random variation.[1]

Design. The design was within subjects, with the independent variable being the number of objects and their relation to the target apices, leading to four conditions: *single*, in which a single hexagon was displayed (e.g., Figure 2a); *double baseline*, in which displays comprised two objects (K-shaped pentagons) that were vertically aligned (e.g., Figure 2b); *double congruent*, in which displays comprised two objects that were vertically offset by 0.4° in the same direction as the offset of the target apices (e.g., Figure 2c); *double incongruent*, in which displays were composed of two objects that were vertically offset by 0.4° in the opposite direction of the offset of the target apices (e.g., Figure 2d).

The single condition occurred three times more than each of the equiprobable double conditions, with the result that subjects saw one-object displays as often as two-object displays. Note that the distance between the target apices and their vertical offset is exactly the same in all four conditions.

Procedure. Subjects were shown a diagram of a typical display from each condition and told to respond to the vertical offset of the apices, pressing the response key on the side of the lower apex. They were instructed to ignore the relative positions of the objects themselves because these were uninformative and could sometimes be misleading.

The sequence of events on any one trial was as follows: The fixation cross was presented for 500 ms, followed by the display for 150 ms. An incorrect response produced a loud beep, whereas no feedback was given on correct trials. An interval of 700 ms followed the subject's response, and then the sequence was repeated to produce the next trial. There were six blocks of 100 trials each. Within each block, condition was interleaved in a different pseudorandom sequence for each subject, with the constraint that there were equal numbers of the various double conditions and equal numbers of single displays and double displays overall. Half the trials in each block required a response with the left hand, half with the right. At the end of each block, subjects' mean RT for correct responses was displayed on the monitor, together with their mean error rate and a message that requested that they be more accurate in the next block if their error rate had exceeded 15% or that they respond more quickly if their error rate had been below 5%.

Treatment of results. The first block of 100 trials was discarded as practice, as were the first 2 trials of each block. Thus 490 trials were available for each subject. All these contributed to the accuracy analyses. However, the data were trimmed for RT analysis as follows for this experiment and all others in this article. Error trials were excluded. Trials immediately following an error were also omitted because of the variability they typically introduce (Rabbitt, 1966). An upper and lower RT cutoff for acceptable trials was calculated by using the method of Driver and Baylis (1991), which trims the tails of the RT distribution to remove those trials with extremely short or extremely long RTs, beyond which performance was at chance (i.e., failed a chi-square test at the 5% level). The combination of these upper and lower RT criteria excluded 10.7% of the recorded data; the pattern of the results is the same if these trials are included. All analyses were carried out with SYSTAT (Wilkinson, 1986).

Results

The means of subjects' median RTs are shown in Figure 3 for the four conditions, together with the associated mean error rates. A one-way within-subjects analysis of variance (ANOVA) on the RT data showed a highly significant effect of condition, $F(3, 27) = 62.8, p < .001$. Wilcoxon matched-pair signed-rank tests were used to make planned pairwise comparisons between the conditions, with the following results: Single versus double baseline, $t = 0, p < .01$; double congruent versus double incongruent, $t = 0, p < .01$.

A similar analysis was carried out on the error data, showing a highly significant effect of condition in a one-way ANOVA, $F(3, 27) = 6.91, p < .001$. Subsequent Wilcoxon tests gave the following results: Single versus double baseline, $t = 5, p < .02$; double congruent versus double incongruent, $t = 2, p < .01$.

Discussion

There are two findings in this experiment. The effect of the number of objects suggests that comparing the relative location of two points is easier when these points belong to the same object rather than two different objects. This finding is consistent with the proposed hierarchical scheme for position coding, on which the locations of object parts are coded explicitly only in relation to their parent object in the routine

[1] The fixation cross could appear centered at any point within a 0.5° × 0.5° area in the center of the screen. The vertical position of the center point of the single object or pair of objects varied by 0.8° around the center of the screen, independent of fixation. The spatial relations between the objects in two-object displays were constant in a given condition (no vertical offset in the double baseline condition, and 0.4° vertical offset in the double congruent and double incongruent conditions). Finally, the vertical position of the two apices (which bore a constant relation to each other of 0.8° vertical offset) was covaried through 1.6° around the center of the monitor, independent of the variations in object location and fixation location. The apices were always separated by 1.2° horizontally from the nearest end of the horizontal contours in single displays. Of course, their vertical separation from these contours varied with the pseudorandom fluctuation in their own vertical location and that of the objects. The positions taken by the apices were the same for the single and double conditions, and the relation of the apices to the vertical position of the objects was identical in the single and double baseline conditions.

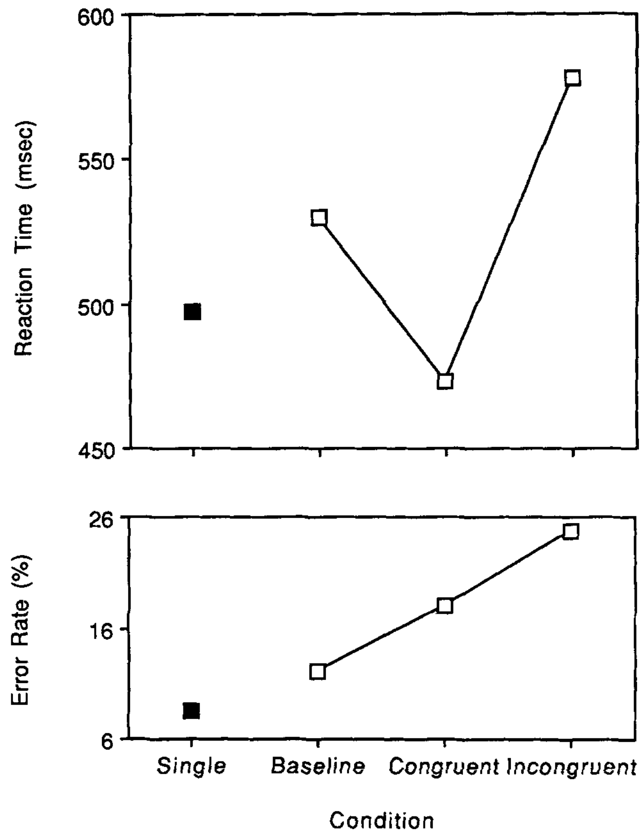

Figure 3. Mean reaction times (top panel) and error rates (bottom panel) for Experiment 1. (Data from one-object displays are shown with filled symbols; data from two-object displays are shown with open symbols.)

process of deriving "what" and "where" information. The result also favors object-based views of attention over purely spatial accounts; a cost of attending to two objects rather than one was obtained even though the points that had to be compared in the two cases had identical positions in space.

The second finding was that judgments about the relative position of contours from two different objects were impaired if the objects were offset in the opposite direction. This effect is also consistent with our hierarchical scheme for visual position coding. Object positions should influence judgments concerning the relative position of parts from different objects because these part positions are described relative to the object positions. We consider the interfering effects of object offsets on apex judgments further in Experiments 3–7.

Our second experiment concentrates on the two- versus one-object cost observed in Experiment 1. We have attributed this effect to the number of objects involved. However, it is possible that the effect is caused by some physical difference between the single and double displays other than the number of objects presented. For example, the spatial frequency composition of the single and double displays is not matched, so an explanation analogous to Watt's (1988) account of Duncan's (1984) data may be possible.

In Experiment 2 we used a similar task to that used in the first study, while attempting to completely preclude stimulus-bound explanations for the two-object cost. We used physically identical displays in the one- and two-object cases, manipulating subjects' perceptual set so that the very same display was parsed as containing one or two objects.

Experiment 2

As before, the task was to compare the vertical position of two apices. Red and green displays were used rather than the monochrome displays of Experiment 1. Each display comprised a colored hexagon (the same shape as single-condition objects from Experiment 1) flanked on either side by K-shaped pentagons of a different color (shaped like double-condition objects from Experiment 1). Examples are given in Figures 4a and 4b (in which the different shading indicates the different colors, i.e., red or green). It can be seen that the resulting displays yield ambiguous figures analogous to Rubin's (1921) celebrated vase–faces picture. Like Rubin's picture, the displays in Figures 4a and 4b can be perceived in two ways; either as two outer figures (the pentagons) against a central ground or as a single central figure (the hexagon) with surrounding ground.

Subjects' interpretation of these ambiguous figures was manipulated by means of their perceptual set. One group of subjects (the red group) was instructed to respond to the red contours, whereas the other group (the green group) had to respond to the green contours. The red group should parse displays like Figure 4a as comprising one relevant object, whereas Figure 4b should be seen to contain two relevant objects. On the object-based account of attention (and our own hierarchical position-coding hypothesis), the red group should therefore find judging the relative position of the apices more difficult for displays like Figure 4b than Figure 4a. However, the reverse should apply for the green subjects (i.e., Figure 4a should be parsed as two relevant objects and Figure 4b as one relevant object), so for them displays like Figure 4a should be harder than Figure 4b. No stimulus-bound explanation could be advanced if this pattern were found because the very same displays would be relatively easy for one group of subjects to judge and relatively difficult for the other group.

To ensure that subjects attended to their specified color, we included displays in which the outer pentagons were horizontally separated from the central hexagon rather than joined (e.g., Figures 4c–f). This permits the side of the lowest apex on the central hexagon to disagree with the side of the lowest apex on the outer pentagons (e.g., Figures 4e and 4f). For these incongruent displays, the correct response can be made only if the color instruction is followed. For example, the correct response for a red subject to a display like Figure 4e was the left key, whereas the correct response for a green subject was the right key. Incorrect responses produced a loud beep. Half the trials were of the separated incongruent type (e.g., Figures 4e and 4f). These were included to ensure that subjects adhered to their color instructions, to influence their parsing of the critical joined displays.

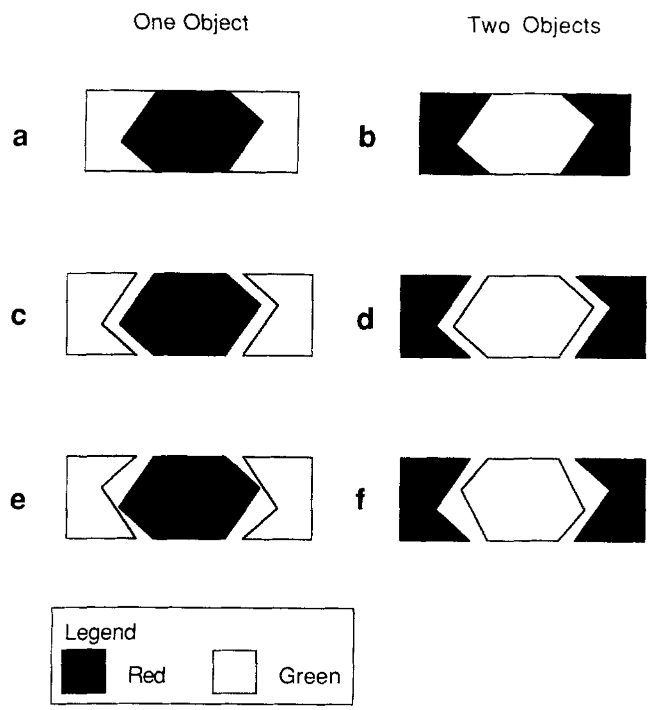

Figure 4. Examples of displays from the six conditions used in Experiment 2: (a) Joined one object, (b) joined two object, (c) separated congruent one object, (d) separated congruent two object, (e) separated incongruent one object, (f) separated incongruent two object. (The displays are labeled as containing one or two objects from the perspective of a subject in the red group; for subjects in the green group, the number of objects would be the reverse.)

Method

Methodological details were as for Experiment 1 except where indicated.

Subjects. The 20 subjects (12 women and 8 men) were lower division psychology undergraduates at the University of California, San Diego. Ten were randomly assigned to the red group, and the remainder were assigned to the green group. All had normal or corrected-to-normal vision and normal color vision by self-report, and they received course credit for their participation.

Apparatus and materials. The apparatus was as before. The shapes were presented in red and green on a black background. Viewing distance was 70 cm with the result that an entire joined display (e.g., Figure 4a or 4b) subtended 6.0° horizontally × 2.6° in height. The dimensions of the outer pentagons and central hexagon were as for Experiment 1. The horizontal separation of each outer pentagon from the central hexagon was 0.8° in separated displays. The task was to decide which of the apices in the target color was lower on the screen, responding with keypresses as before.

Design. The design was mixed, with one between-subjects factor and two within-subjects factors. The between-subjects factor was the target color, red or green. The first within-subjects factor, number of relevant objects (alternatively, display coloring), had two levels: There was either one object in the target color (the central hexagon) or two objects (the outer pentagons). These two patterns of coloring occurred equally often. The second within-subjects factor, display type, had three levels: joined, congruent–separated, or incongruent–separated. The incongruent conditions occurred twice as often as either the joined or congruent conditions to ensure that subjects attended to their target color. The six within-subject conditions are illustrated in Figure 4.

Procedure. Subjects were shown illustrations of the six display types and were told to respond on the side of the lowest apex in their target color. The importance of following the color instruction was emphasized because it was the only guide to the correct response for incongruent displays. The timing of events and the feedback provided were as in Experiment 1. There were six blocks of 96 trials each. Within each block, the six types of display appeared in the appropriate overall proportions in a different pseudorandom sequence for each subject. Half the trials in each block required a response with the left hand, half with the right.

Treatment of results. This was as in Experiment 1. The first block of 96 trials was discarded as practice, as were the first 2 trials of each block. Thus 470 trials were available for each subject. All these contributed to the accuracy analyses. Data were trimmed for RT analyses as in Experiment 1, and the upper and lower RT criteria excluded 7.8% of the recorded data.

Results

The means of subjects' median RTs are shown in Figure 5 for the six within-subject conditions, together with the associated mean error rates, pooled across color group. Re-

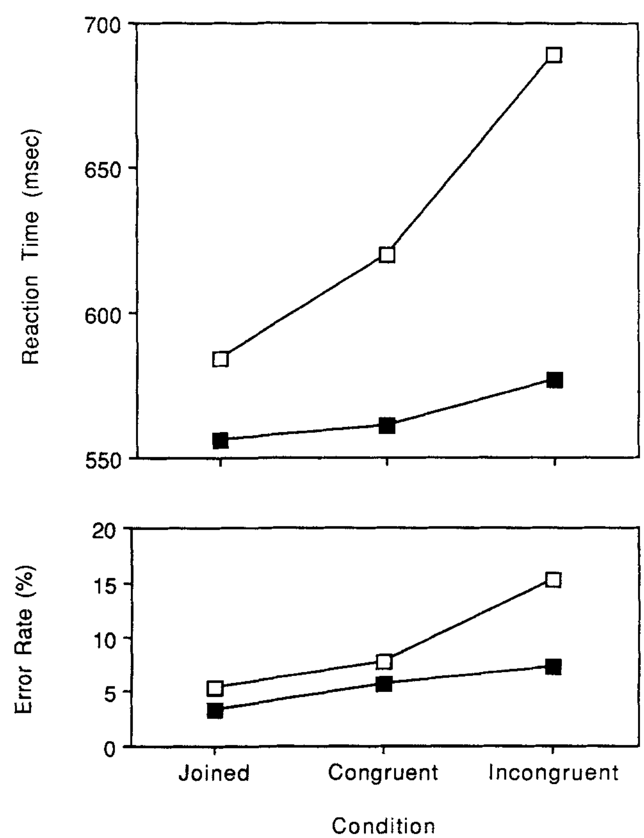

Figure 5. Mean reaction times (top panel) and error rates (bottom panel) for Experiment 2, pooled across subject group. (Data for displays that should be parsed as one object are shown as filled symbols, and data for displays parsed as two objects are shown as open symbols.)

sponses were clearly above chance in the incongruent conditions, demonstrating that subjects followed their color instructions. A three-way mixed ANOVA on the RT data found no main effect, $F(1, 18) = 0.0$ or interaction involving the between-subjects factor of color group, $F(1, 18) = 0.0$; for the interaction with number of relevant objects, $F(1, 18) = 0.1$; for the interaction with display type, $F(2, 36) = 0.4$; for the three-way interaction, $F(2, 36) = 0.4$.

There were main effects of the number of relevant objects, $F(1, 18) = 77.3$, $p < .0001$, and display type, $F(2, 36) = 58.6$, $p < .0001$, and an interaction between them, $F(2, 36) = 22.8$, $p < .0001$. Planned Wilcoxon tests showed that in every case the single-object conditions were faster than their two-object equivalents: Single joined versus double joined, $t = 0$, $p < .001$; single congruent versus double congruent, $t = 0$, $p < .001$; single incongruent versus double incongruent, $t = 0$, $p < .001$.

Similar analyses were carried out on the accuracy data. A three-way mixed ANOVA again showed no effects involving color group: For the main effect, $F(1, 18) = 1.5$; for the interaction with number of objects, $F(1, 18) = 0.1$; for the interaction with display type, $F(2, 36) = 0.2$; for the three-way interaction, $F(2, 36) = 0.6$. There were effects of the number of relevant objects, $F(1, 18) = 13.9$, $p < .002$, and of display type, $F(2, 36) = 24.9$, $p < .001$, and a significant interaction, $F(2, 36) = 6.4$. Wilcoxon tests gave the following results for the number of objects: Single joined versus double joined, $t = 17$ (with 3 ties), $p < .02$; single congruent versus double congruent, $t = 103$, *ns*; single incongruent versus double incongruent, $t = 13$, $p < .01$.

The congruency of the flanking pentagons with the central hexagon affected performance for the separated displays.[2] In the current experiment, the purpose of these separated displays was to ensure that subjects adhered to the color instructions for the critical joined displays. The prediction for the joined displays was a cost for displays parsed as containing two objects rather than one. This effect of the number of relevant objects was found in the analyses presented above, but those analyses pooled the data across subject group and across joined and separated displays. A stronger test would examine the red and green subjects' data separately for just the joined displays. The red subjects should show a cost for displays like Figure 4b compared with Figure 4a. However, the reverse should apply for the green subjects, who should find displays like Figure 4a more difficult because these contain two objects in their target color.

Figure 6 presents the mean RTs and error rates for the two different joined-display colorings, shown separately for the red and green subjects. A two-way mixed ANOVA (Color Group × Display Coloring) was conducted on the RT data from the joined displays only. There were no main effects of group or of display type, $F(1, 18) = 0.0$ in both cases, but there was a significant interaction, $F(1, 18) = 67.3$, $p < .001$. The red subjects responded faster to the joined displays with a red center than the displays with the green center, but the reverse was found for the green subjects. Thus, the relative difficulty of the two types of joined display did not depend on their physical properties but rather on how many relevant objects they were seen to contain, according to the subjects' perceptual set. A similar analysis of the error data (from the joined displays only) found no main effects of group or of display coloring, $F(1, 18) = 2.7$, *ns*, and $F(1, 18) = 1.6$, *ns*, respectively, and no interaction, $F(1, 18) = 1.6$. However, inspection of Figure 6 shows that the accuracy data showed a pattern similar to that of the RT data, so speed–accuracy trade-offs can be discounted.

Discussion

In Experiment 1, comparing the relative location of two contours was easier when the contours belonged to the same object rather than two different objects. In the joined displays of Experiment 2, we observed the same cost for judging contours from two objects even though there was no physical difference between one- and two-object displays. Identical stimuli were parsed as two-object displays for one group of subjects and as one-object displays for the other group, depending on subjects' perceptual set. The two-object cost observed for the joined displays in the present experiment must therefore depend on top-down attentional factors rather than any bottom-up stimulus factor, such as differences in spatial frequency composition. These data are consistent with previous suggestions of a difficulty in attending to two objects simultaneously (Duncan, 1984) and agree with our suggestion that object parts are located explicitly only in relation to their parent object in the routine process of deriving "what" and "where" information, so comparison of parts from different objects is necessarily less direct.

The dramatic impact of subjects' perceptual set on the relative difficulty of the joined displays (see Figure 6) is consistent with the phenomenology of ambiguous figures such as Figure 4a and 4b, or Rubin's (1921) vase–faces drawing. The interesting property of such displays is that the dividing contour can be seen as belonging to one side of the divide or the other but never both (Hoffman & Richards, 1984). We have exploited this aspect of the visual system's parsing procedures in the present experiment. Subjects in the green

[2] Wilcoxon tests on the RT data found the following effects of congruency for the separated displays: Single congruent versus single incongruent, $t = 18$, $p < .025$; double congruent versus double incongruent, $t = 1$, $p < .001$. The effects of congruency on the error rates were similar: Single congruent versus single incongruent, $t = 22.5$, $p < .05$; double congruent versus double incongruent, $t = 7$, $p < .01$. Thus subjects were affected by the agreement between the offset of the pentagons' apices and that for the hexagon's apices. The interaction between display type and number of relevant objects arose because congruency had more substantial effects in the separated two-object displays than in the separated one-object displays. This interaction may simply reflect reduced acuity for apices on the outer pentagons relative to the central hexagon in the separated displays. Double congruent displays are slower than double joined displays, presumably because the information to be judged is more peripheral in the former case. Such acuity-based differences in saliency in the separated displays could result in the hexagon apices exerting more of an effect on judgments of the pentagon apices than vice versa, so the effect of congruency is larger in the double displays than in the single displays.

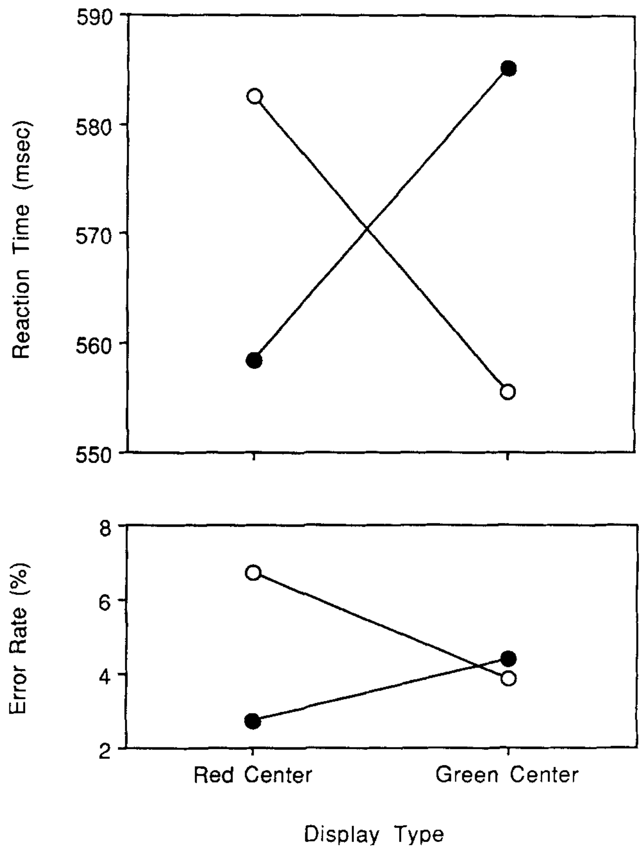

Figure 6. Mean reaction times (top panel) and error rates (bottom panel) for the critical joined displays of Experiment 2, separated by whether the displays had a red center (e.g., Figure 4a) or a green center (e.g., Figure 4b). (Data for the red group of subjects are shown as filled circles, and those for the green group are shown as open circles.)

group were apparently unable to code the apices in displays like Figure 4a as belonging to both the outer pentagons and the central hexagon. If they had been able to do so, their performance with such displays should have been the same as for the red subjects. The complementary argument applies to the red subjects' performance for displays with a green center, such as Figure 4b.

Experiment 3

Having demonstrated that the two-object cost initially observed in Experiment 1 is found even when there are no physical differences between one- and two-object displays, we now turn to further consideration of the other findings from Experiment 1. In particular, we focus on the finding that incongruent object offsets (as in Figure 2d) interfered with judgments of part offset, and we consider various hypotheses about how the offset of parts from different objects might be compared in the two-object tasks.

We have suggested that object parts are only located explicitly relative to their parent object in the routine process of deriving "what" and "where" information and that this coding scheme is responsible for both the two-object cost and the interference from incongruent object offsets. When faced with a two-object display, according to our hypothesis the visual system should routinely code the relative location of the objects and describe the location of each object's parts relative to that object's location. However, this provides no explicit code for the relative locations of parts from different objects. These could be derived by a combination of the object-location and part-location descriptions (analogously to the indirect routes between dots from different groups in Figure 1c). Both the two-object cost and the interference from incongruent object offsets are consistent with the use of an indirect scheme such as this for between-object comparisons.

However, it might be suggested that the two-object tasks are performed by some operation applied to the space intervening between the two objects. For example, subjects might "fill in" an imaginary hexagon between the presented white pentagons (or perceive a subjective hexagon analogous to Kanizsa (1979) figures) and then proceed as for a true hexagon. Alternatively, they might construct an imaginary contour between the apices of the two pentagons and then judge this contour's orientation.[3] These possible strategies do not challenge our central claim that within-object part locations are coded more directly than between-object part locations. Indeed, the proposed strategies would provide support for this hypothesis because they would represent an attempt to turn the two-object tasks into single-object tasks. Presumably this would only be attempted if single-object judgments were indeed easier than two-object judgments.

Though the possibility of such strategies does not undermine our account of the two-object cost, it does challenge our interpretation of the interference from incongruent object offsets. We attributed this interference to the mismatch between apex positions and the object positions each apex is described relative to on our hypothesis. However, it is possible that the interference reflects distortion of the area between the objects rather than reflecting the offset in object locations per se. In the double incongruent condition of Experiment 1 (see Figure 2d), the top and bottom edges of the imaginary hexagon between the K-shaped pentagons have a slant that is opposed to the apex offset. This might produce the interference we observed if subjects followed a strategy of filling in an imaginary hexagon and were affected by the slant of its edges when comparing its apices.

If the interference was due to distortion of the central space rather than to object position per se, comparable interference should be found when judging the apices of an actual distorted hexagon. Indeed, the interference should be at least as great and probably larger because the disruption produced by an actual irrelevant contour will presumably be greater than that produced by an imaginary irrelevant contour. Accordingly, we carried out a replication of Experiment 1 that included distorted single-hexagon displays to match the imaginary shapes formed between the pentagons in congruent and incongruent two-object displays. As before the task was to

[3] We are grateful to John Duncan for these suggestions.

compare the vertical position of two apices. Each display comprised either a white central hexagon or two white K-shaped pentagons.

Note that our hierarchical coding account does not preclude interference for the single distorted hexagons. The position of all parts should be coded relative to the same parent object in this case, and we have no grounds for excluding the possibility of interference from irrelevant contours within an object description. The only prediction that can be made is as follows. If the interference we observed in Experiment 1 from incongruent object offsets was caused by within-object interference arising in an imaginary distorted hexagon, it must be as great or larger with real distorted hexagons.

Method

All methodological details were as for Experiment 1 unless indicated.

Subjects. The 20 new subjects (11 women and 9 men) were lower division psychology undergraduates at the University of California, San Diego. All had normal or corrected-to-normal vision and normal color vision by self-report, and they received course credit for their participation.

Design. The design was within subjects, with two variables. The first variable was the number of objects: There was either one object (a central hexagon) or two objects (outer pentagons). The second variable, offset, had three levels: congruent, baseline, and incongruent. In congruent two-object displays, the pentagons were vertically offset by 0.4° in the same direction as the apices. In congruent one-object displays, the top and bottom contours of the hexagon displays had a corresponding tilt (0.4° offset between endpoints) in the same direction as the apex offset. Pentagon offset and hexagon tilt were opposed to apex offset in the incongruent conditions, and there was no pentagon offset or hexagon tilt in the baseline conditions. The two variables were crossed to yield six conditions that occurred equally often (see Figure 7).

Procedure. Subjects were shown illustrations of the six display types and were told to respond on the side of the lowest apex. The timing of events and the feedback provided were as in Experiment 1. There were eight blocks of 72 trials each. Within each block, the six types of display were equiprobable and interleaved in a different pseudorandom sequence for each subject. Half the trials in each block required a response with the left hand, half with the right.

Treatment of results. This was as for Experiment 1. The first block of 72 trials was discarded as practice, as were the first 2 trials of each block. Thus 490 trials were available for each subject. All these contributed to the accuracy analyses. Data were trimmed for RT analyses as in Experiment 1, and the upper and lower RT criteria excluded 5.3% of the recorded data.

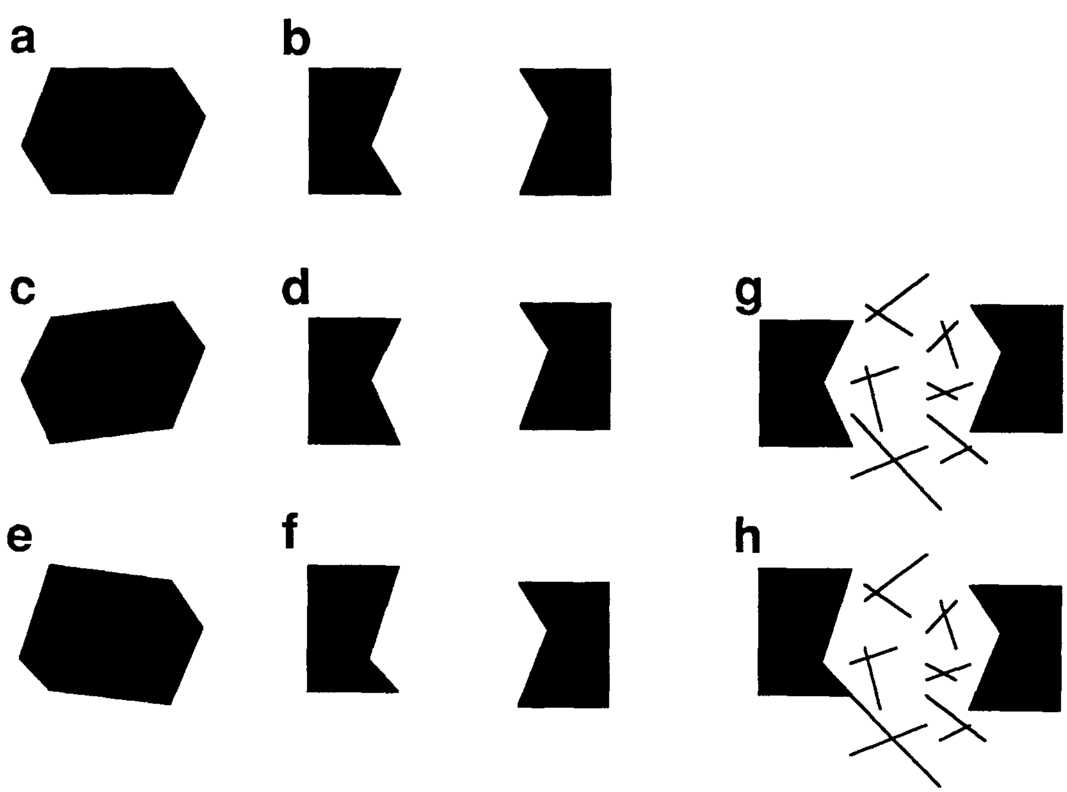

Figure 7. Sample displays from Experiment 3. ([a] Single baseline, [b] double baseline, [c] single congruent, [d] double congruent, [e] single incongruent, [f] double incongruent. Panels c–h illustrate the six conditions from Experiment 4. The single baseline and double baseline conditions [Panels a and b] were omitted and replaced with double congruent with crosses [Panel g] and double incongruent with crosses [Panel h].)

Results

The means of subjects' median RTs are shown in Figure 8 for the six conditions, together with the associated mean error rates. A two-way within-subjects ANOVA on the RT data found a significant effect of number of objects, $F(1, 19) = 64.1$, $p < .001$, a significant effect of congruency condition, $F(2, 38) = 102.5$, $p < .001$, and a significant interaction, $F(2, 38) = 10.6$, $p < .001$. Planned Wilcoxon tests showed that in every case the single-object conditions were faster than their two-object equivalents: Single baseline versus double baseline, $t(20) = 1$, $p < .001$; single congruent versus double congruent, $t(20) = 29$, $p < .01$; single incongruent versus double incongruent, $t(20) = 4$, $p < .001$. Further Wilcoxon tests were used to investigate the effects of congruency: Single baseline versus single congruent, $t(20) = 41.5$, $p < .05$; single baseline versus single incongruent, $t(20) = 1$, $p < .001$; double baseline versus double congruent, $t(19) = 0$, $p < .001$; double baseline versus double incongruent, $t(20) = 0$, $p < .001$.

Similar analyses were carried out on the accuracy data. A two-way within-subjects ANOVA on the RT data found a significant effect of number of objects, $F(1, 19) = 32.2$, $p < .001$, a significant effect of congruency condition, $F(2, 38) = 56.1$, $p < .001$, and a significant interaction, $F(2, 38) = 20.2$, $p < .001$. Planned Wilcoxon tests showed that in every case the single-object conditions were more accurate than their two-object equivalents: Single baseline versus double baseline, $t(19) = 8.5$, $p < .01$; single congruent versus double congruent, $t(17) = 35.0$, $p < .05$; single incongruent versus double incongruent, $t(20) = 10.0$, $p < .01$. Further Wilcoxon tests were used to investigate the effect of congruency: Single baseline versus single congruent, $t(16) = 46.0$, $p < .1$; single baseline versus single incongruent, $t(20) = 3$, $p < .001$; double baseline versus double congruent, $t(18) = 18.5$, $p < .01$; double baseline versus double incongruent, $t(19) = 0$, $p < .001$.

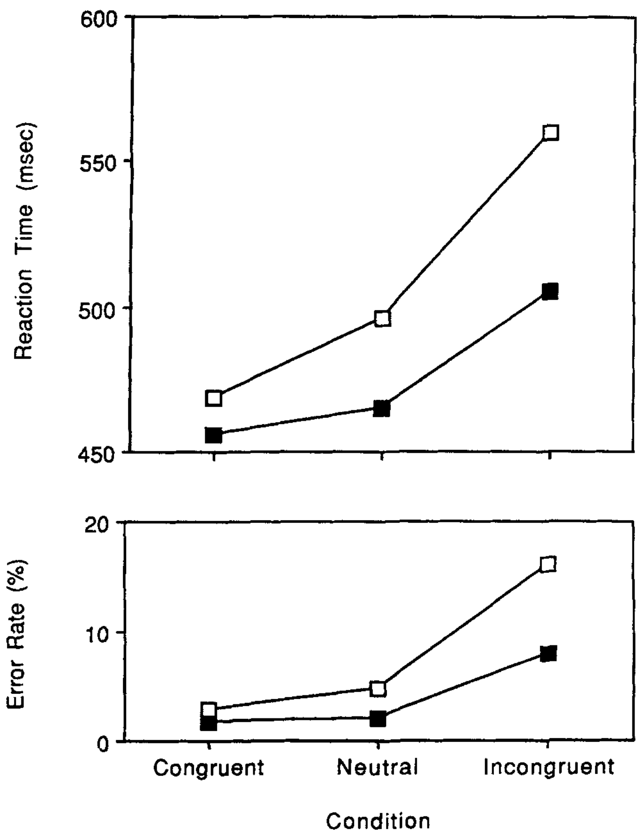

Figure 8. Mean reaction times (top panel) and error rates (bottom panel) for Experiment 3. (One-object display data are shown with filled symbols, and two-object display data are shown with open symbols.)

Discussion

Judging the relative position of two points was harder when they belonged to two objects rather than one, as we had found in Experiments 1 and 2. We also replicated the interference effect from incongruent object offsets on judgments for parts from different objects, as previously observed in Experiment 1. Distortion of the central hexagon (in the single congruent and incongruent conditions) also produced interference, but this was smaller than the interference caused by object offsets. The effect of object offsets cannot, therefore, solely be the product of a distorted imaginary hexagon between the two pentagons, unless one accepts the implausible proposition that imaginary irrelevant contours are more disruptive than actual irrelevant contours. At least some of the object-offset interference must be due to the relative location of the objects per se, as our hierarchical coding interpretation suggests.

To assess the likelihood that an imaginary or subjective central figure was responsible for any part of the congruency effect in two-object displays, all subjects were asked at the end of the experiment if they had perceived any imaginary or subjective central figures. Subjects were shown a display like that in Figure 7b, presented on the computer for the duration of this debriefing. They were asked "Do you perceive any shape in the middle of the display?" If they answered affirmatively, they were asked to name the shape they could see. The results are shown in the first column of Table 1. Those subjects who reported no central shape were then prompted with an example of a Kanisza triangle taken from Rock (1984, p. 132), and the question was repeated as they continued to view the computer display. The final reports of these subjects were added to the reports of figures from the unprompted subjects to give the numbers in Column 2 of Table 1.

It is clear that few subjects were aware of any subjective central figure in two-object displays without prompting, that the reports that were elicited varied in shape, and that many of these shapes (e.g., diamonds) are implausible candidates for producing interference in the incongruent displays. Our conclusions in this article do not rest on these phenomenal report data alone. However, these data further undermine the suggestion that imaginary or subjective central contours played a central role in performing the two-object tasks and in producing interference under conditions of object offset. In the next experiment, we address the suggestion more di-

Table 1
Numbers of Each Type of Kanisza Figure Perceived in Experiments 3 and 4

Figure	Experiment 3 (n = 20) Initial	Experiment 3 (n = 20) Including prompts	Experiment 4 (n = 18) Without crosses	Experiment 4 (n = 18) With crosses
Nothing	14	7	8	15
Diamond or square	4	10	6	2
Rounded diamond[a]	1	1	2	0
Hexagon	1	1	2	1
Triangle[b]	0	1	0	0

[a] The subjects in this category described a figure with apices to the left and right, with an upper and lower curved line between them. [b] When further quizzed about the shape of the reported triangle, the subject described it as having one apex at each side and the third pointing either up or down. It seems likely that this subject was overly influenced by the figure shown in Rock (1984).

rectly by adding disruptive visual information between the shapes in two-object displays specifically to obstruct the emergence of any central subjective contours.

Experiment 4

As before, the task was to compare the vertical position of two apices. Each display comprised either a central hexagon or two K-shaped pentagons. In some conditions, irrelevant contours were introduced between the pentagons to obstruct the emergence of any subjective or imaginary central figure. If performance of the two-object task depends on constructing a central shape, the intervening irrelevant contours should be very disruptive and lead to an increase in the difficulty of this task. Similarly, if subjects were performing this task by constructing a line between the two apices, the presence of intervening contours will render the task more difficult.

Method

The methodological details followed Experiment 1 unless indicated.

Subjects. The 18 new subjects (9 women and 9 men) were lower division psychology undergraduates at the University of California, San Diego. All had normal or corrected-to-normal vision and normal color vision by self-report. Subjects received course credit for their participation.

Apparatus and materials. The apparatus was as before. The shapes were presented in red on a black background. Display dimensions were as for Experiments 1 and 3. The types of display were as for Experiment 3, except that the baseline displays were omitted and replaced with congruent and incongruent two-object displays that had irrelevant green crosses placed between the pentagons. The intention was that these intervening contours would preclude the emergence of subjective contours between the pentagons. The strokes of the distractor crosses had random orientation and were approximately 1.2° long. Notionally, they were each centered at locations horizontally equidistant from their neighboring cross and the nearest contour of the closest pentagon. Vertically, they were notionally centered at either the midpoint of the display or at the top or bottom margins of the pentagons. The crosses could deviate randomly up to 0.8° from these nominal centers. When present, there were always six crosses in total. The task was to decide which of the two apices on the pentagons was lower on the screen, as before.

Design. The design was within subjects, with two variables. The first variable was the number of objects and had three levels: There was either one object (the central hexagon), two objects (the outer pentagons), or two objects plus distractor crosses. The second variable had two levels: congruent and incongruent. These factors were crossed to yield six conditions (see Figure 7).

Procedure. Subjects were shown illustrations of the six display types and were told to respond on the side of the lowest apex. The timing of events and the feedback provided were as in Experiments 1–3. There were eight blocks of 72 trials each. Within each block, the six types of display were equiprobable and interleaved in a different pseudorandom sequence for each subject. Half the trials in each block required a response with the left hand, half with the right.

Treatment of results. The first block of 72 trials was discarded as practice, as were the first 2 trials of each block. Thus 490 trials were available for each subject. All these contributed to the accuracy analyses. Data were trimmed for RT analyses as in Experiments 1–3, and the upper and lower RT criteria excluded 7.4% of the recorded data.

Results

The means of subjects' median RTs are shown in Figure 9 for the six conditions, together with the associated mean error rates. A two-way within-subject ANOVA on the RT data found no main effect of number of objects, $F(1, 17) < 1$, a significant effect of congruency condition, $F(2, 34) = 13.9$, $p < .001$, and a significant interaction, $F(2, 34) = 93.5$, $p < .001$. An equivalent analysis of the accuracy data found a significant effect of number of objects, $F(1, 17) = 6.0$, $p < .05$, a significant effect of congruency condition, $F(2, 34) = 4.9$, $p < .05$, and a significant interaction, $F(2, 34) = 39.1$, $p < .001$.

The data were next analyzed in two separate two-way ANOVAs to investigate, analogously to Experiment 3, the congruency effects in single versus double displays and to investigate the effect of distracting crosses on two-object judgments. The RT data from displays without crosses showed a significant effect of number of objects, $F(1, 17) = 16.1$, $p < .01$, a significant effect of congruency condition, $F(1, 17) = 159.3$, $p < .001$, and a significant interaction, $F(1, 17) = 12.2$, $p < .01$. Congruency effects were larger in the two-object displays, replicating Experiment 3. The same pattern was found in an equivalent analysis of the accuracy data: for number of objects, $F(1, 17) = 12.0$, $p < .01$; for congruency, $F(1, 17) = 62.2$, $p < .001$; and for the interaction, $F(1, 17) = 7.4$, $p < .02$.

The RT data from just the two-object displays showed no significant effect of the presence of distracting crosses, $F(1, 17) = 3.7$, a significant effect of congruency, $F(1, 17) = 117.1$, $p < .001$, and no interaction, $F(1, 17) < 1$. Thus, the intervening crosses neither disrupted the two-object judgment substantially nor modulated the interference effects. The same pattern was found in an equivalent analysis of the accuracy data: no effect for the presence of crosses, $F(1, 17)$

< 1, a significant effect of congruency, $F(1, 17) = 22.1$, $p < .001$, and no interaction, $F(1, 17) < 1$.

As in Experiment 3, all subjects were asked at the end of the experiment if they perceived any subjective central shapes in two-object displays. In this experiment, all subjects were first shown a diagram of a Kanisza triangle from Rock (1984, p. 132). Subjects were then shown a display like that of Figure 7d presented continuously on the video monitor and asked the same question as in Experiment 3. Subjects were then shown a display like Figure 7g, and the question was repeated. The results are shown in the third and fourth columns of Table 1. It can be seen that the presence of the intervening crosses significantly decreases the reports of subjective shapes ($p < .05$).

Discussion

As in Experiments 1–3, we again found that judging the relative position of two points in visual space was harder when they belonged to two objects rather than one. We also replicated the interference effect from incongruent object offsets previously observed in Experiments 1 and 3. Distortion of the central hexagon in the single congruent and incongruent conditions again produced some interference, but as in Experiment 3 this was smaller than the interference caused by object offsets. Finally, the inclusion of distracting contours between the shapes in two-object displays neither increased the difficulty of the task significantly nor modulated the interference produced by incongruent object offsets, although it did reduce reports of subjective contours between the shapes.

The suggestion that subjects solve the two-object tasks by constructing an imaginary or subjective hexagon between the two pentagons seemed implausible to us a priori because the briefly presented figures in two-object displays are much smaller and brighter than the area between them. The possibility did not in any case challenge our account of the two-object cost, as discussed earlier. However, it did raise doubts about our attribution of the interference effect from incongruent object offsets to the mismatching object locations themselves, as this interference might have been produced by distortion of the central space between the objects.

We now have considerable grounds for rejecting this suggestion. First, many subjects denied the existence of subjective central figures, even when heavily prompted. This seems unlikely if they had been using such figures to perform the two-object task on each trial. The subjective shapes reported under prompting varied considerably, and most bore no plausible relation to the observed interference from incongruent object offsets. Second, both Experiments 3 and 4 found larger congruency effects in two-object displays than in single-object displays. The suggestion that both of these effects reflect distortion of a central hexagon therefore requires one to accept that imaginary contours are more disruptive than actual contours. Third, adding distracting random contours between the two-object displays to obstruct imaginary or subjective contours did not affect the difficulty of the task and did not modulate the interference effects in Experiment 4.

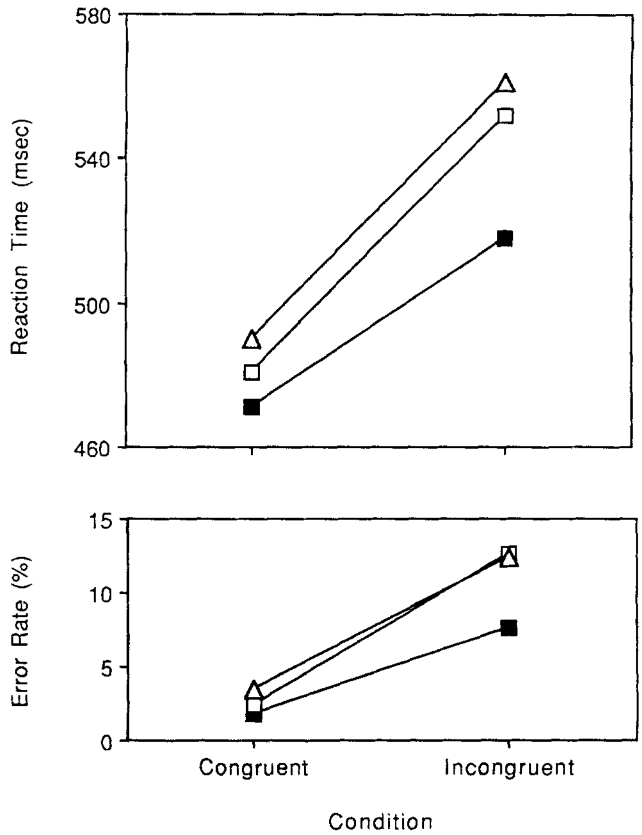

Figure 9. Mean reaction times (top panel) and error rates (bottom panel) for Experiment 4. (One-object displays are shown as filled symbols, and two-object displays are shown as open symbols. The open triangles are for two-object displays with distractor crosses, and the open squares are for two-object displays without distractor crosses.)

At this point, it seems safe to conclude that at least some of the interference from incongruent object offsets genuinely involves object-position information rather than just distortion of the central space. This intrusion of object locations into part judgments accords with our hierarchical position-coding hypothesis. On this view, the derivation of part positions is logically dependent on the derivation of object positions because the former are described relative to the latter (Figure 1c). It is therefore natural that incongruent object offsets can interfere with part offset judgments. However, the reverse interference (incongruent part offsets interfering with object-position judgments) should not be observed according to our hypothesis. Relative object positions are described independently of relative part positions on the hierarchical scheme. The next three experiments test for this asymmetry in the pattern of object–part interference.

Experiment 5

To test the prediction that object positions will interfere with judgments of part positions but not vice versa, one has to ensure that the pattern of interference from object positions on judgments of part position, and vice versa, is not solely a function of their relative physical salience. Accordingly, we compared the effects of two different but relatively large part offsets with a relatively small object offset. The parameters were chosen so that the larger part offset was likely to be more salient and the smaller apex offset less salient than the object offset. To allow for larger apex offsets, it was necessary to use taller displays than the hexagons and pentagons from Experiments 1–4. The purpose of this experiment was to confirm the relative salience of the apex and object offsets with the chosen stimulus parameters.

The displays consisted of two white pentagons on a black background, similar to those used in the two-object displays from Experiments 1, 3, and 4. There were two possible tasks. In the apex task, subjects had to press the key on the side of the lower apex (as in Experiments 1–4) while the pentagons were aligned. In the object task, subjects had to press the key on the side of the lower pentagon while the apices were aligned. Half the subjects had three blocks of the apex task followed by three blocks of the object task, whereas the remainder received the reverse sequence.

Method

Subjects. The 20 subjects (12 women and 8 men) were paid volunteers from the University of Oregon subject panel. All had normal or corrected-to-normal vision by self-report and received $5 for their participation.

Apparatus and materials. The experiment was conducted on an IBM PS/2 Model 30 microcomputer with a color VGA monitor. The onset and the offset of stimuli occurred within a single frame by changing the palette look-up table. Example displays are shown in Figure 10. In the apex task, the apices either had a vertical offset of 0.8° (as in Experiments 1–4; see Figure 10a) or 2.4° (see Figure 10b), whereas the pentagons were vertically aligned. In the object task, the pentagons had a vertical offset of 0.4° (as in the congruent and incongruent conditions from Experiments 1, 3, and 4), whereas the apices were vertically aligned (see Figure 10c). The objects were 3.6° in height rather than the previous 2.6°, but all other visual angles were as for Experiments 1–4. The subjects had to decide which of the pentagons (or apices, as appropriate to their current task) was lower on the screen and press key [Z] if the left object (or apex) was lower or key [/] if the right-hand object (or apex) was lower.

Design. The design was mixed, with one between-subject variable and one within-subject variable. The within-subject variable was the type of display leading to three conditions: small apex offset, in which the pentagons were aligned while the apices were 0.8° offset (Figure 10a); large apex offset, in which the pentagons were aligned while the apices were 2.4° offset (Figure 10b); object offset, in which the apices were aligned while the pentagons were 0.4° offset (Figure 10c).

The between-subjects variable was whether the apex or object task was performed first, with subjects randomly divided into two equal groups.

Procedure. Subjects were shown illustrations of displays from their first task and told to respond on the side of the lowest object or apex as appropriate. New instructions and diagrams were given after three blocks when the task changed. The timing of events and the feedback were as before. There were six blocks of 100 trials each. The first and fourth blocks were discarded as practice, as were the first two trials of each block, leaving 392 trials per subject.

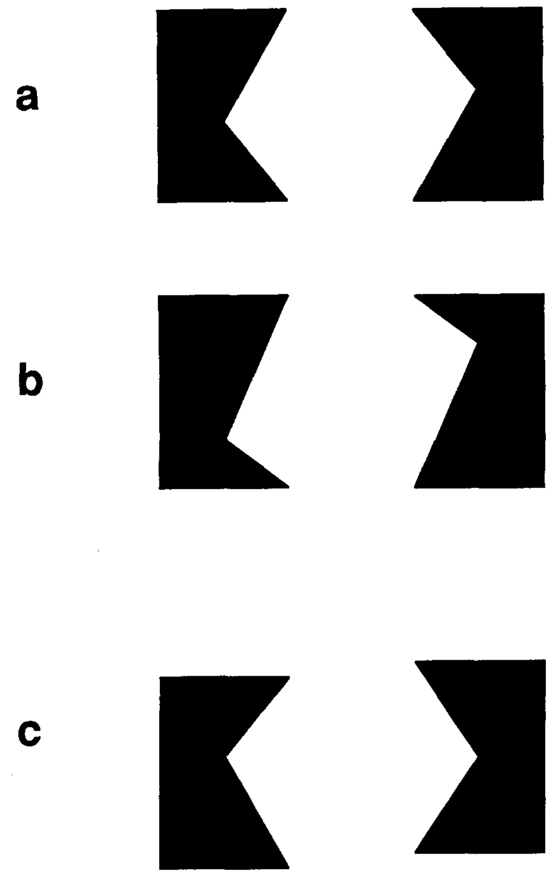

Figure 10. Example displays from Experiment 5: (a) small apex offset, (b) large apex offset, (c) object offset.

Within each block of the apex task, the two apex offsets were equiprobable and interleaved in a different pseudorandom sequence for each subject. In all conditions, left and right responses were required equally often. The upper and lower cutoffs excluded 7.9% of the RT data.

Results

The means of subjects' median RTs are shown in Figure 11 for the three within-subject conditions, together with the appropriate error rates, pooled across the between-subject variable of task order. A two-way mixed ANOVA (Task Order × Condition) on the RT data showed no effect of task order, $F(1, 18) = 2.1$, a significant effect of condition, $F(1, 18) = 23.3$, $p < .001$, and no interaction, $F(1, 18) = 3.0$. Planned Wilcoxon tests on the data pooled across groups confirmed the differences in task difficulty that are apparent in Figure 11: The small apex offset was judged more slowly than the object offset, $t = 36$, $p < .01$. The object offset was judged more slowly than the large apex offset, $t = 43$, $p < .02$. The small apex offset was also slower than the large apex offset, $t = 0$, $p < .01$.

In similar analyses of the error data, a two-way mixed ANOVA found no effect of task order, $F(1, 18) = 0.4$, a significant effect of condition, $F(1, 18) = 35.2$, $p < .001$, and

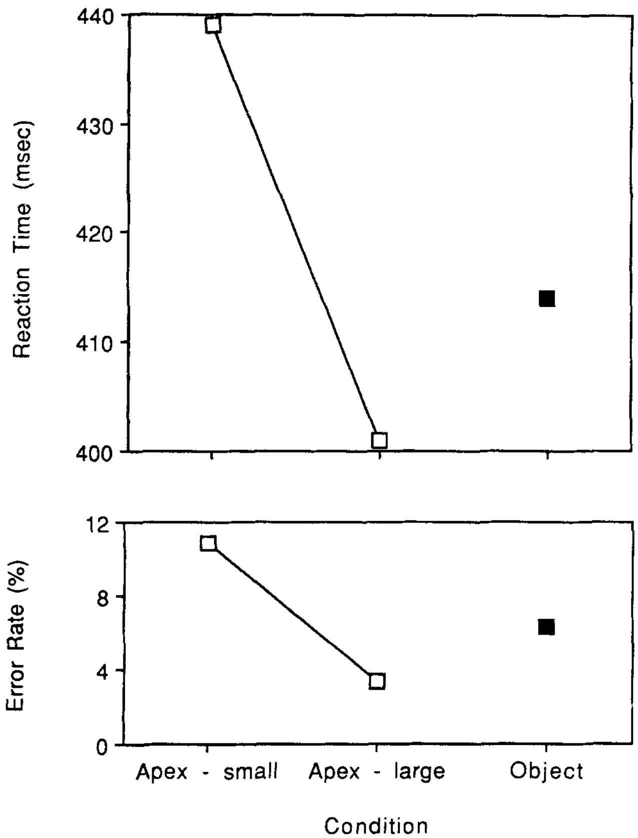

Figure 11. Mean reaction times (top panel) and error rates (bottom panel) for Experiment 5.

no interaction, $F(1, 18) = 2.8$. Planned Wilcoxon tests found the same differences in task difficulty as in the RT data. The small apex offset was less accurate than the object offset, $t = 11.5$, $p < .01$. The object offset was less accurate than the large apex offset, $t = 17$ (with 1 tie), $p < .01$. The small apex offset was also less accurate than the large apex offset, $t = 0$, $p < .01$.

Discussion

With the parameters chosen, the large apex offset was most salient (i.e., judged fastest and most accurately), followed by the object offset and then the small apex offset. In Experiments 1, 3, and 4, incongruent object offsets of the size used here interfered with judgments of the small apex offset used here. We attributed this interference to the hierarchical coding of position, whereby parts of an object are located relative to the location of their parent object. However, the interference from object offset on judgments of apex offset may have arisen in our previous studies simply because the object offset was more salient than the small apex offset with the particular parameters used. The present experiment confirms that this object offset can indeed be derived more rapidly than the small apex offset. Thus, a simple horse-race account of the interference in Experiments 1, 3, and 4 is possible. The object offset may have interfered with apex offsets simply because it was derived more quickly because of greater physical salience.

This horse-race account makes specific predictions about the conditions under which incongruent object offsets should interfere with apex offset judgments and about whether the reverse interference could also be observed (i.e., apex offsets disrupting object offset judgments). For the current parameters, the predictions are as follows. The object offset should interfere with the small apex offset (as we found in Experiments 1, 3, and 4), but it should not interfere with judgments of the large apex offset, which can be derived more rapidly. The latter prediction is tested in Experiment 7. A further prediction of the horse-race account is that the large apex offset should interfere with judgments of the object offset because the large apex offset is derived faster. This would conflict with our hierarchical coding account, which predicts that apex offsets should never interfere with object offsets, regardless of their relative saliency. On the hierarchical scheme, the location of parts is described relative to the locations of their parent objects, but there is no analogous reverse dependency. The object location codes required for the object offset task should be impervious to relative part positions. Experiment 6 tested whether the large apex offset could disrupt judgments of the present object offset, as predicted by the horse-race account in contradiction to our hierarchical account.

Experiment 6

The task was to make speeded judgments about the relative position of two pentagons, pressing the key on the side of the lower object. The displays were similar to the two-object displays used previously, each comprising two white K-shaped pentagons (see Figure 12).

The question was whether incongruent large (or small) apex offsets could affect judgments of object offset. According to our hierarchical account, they should not. According to the horse-race account, apex offsets should interfere provided they are derived more quickly than object offsets. Given the results of Experiment 5, the horse-race theory therefore predicts that the large apex offset should interfere, whereas the small apex offset should not.

Method

Subjects. The 21 subjects (11 women and 10 men) were psychology undergraduates at the University of Oregon. All had normal or corrected-to-normal vision by self-report and received $5 for their participation.

Apparatus and materials. The apparatus was as for Experiment 5. Each display comprised two K-shaped pentagons (see Figure 12) presented as white shapes on a black background, and the visual angles were as for Experiment 5. The subjects responded on key [Z] if the left object was lower and key [/] if the right object was lower.

Design. The design was within subjects, with two variables. The first was the offset of the apices: small (0.8°) or large (2.4°). The second was whether the lowest apex was on the same side as the lowest pentagon (congruent; see Figures 12a and 12b) or on the opposite side (incongruent; see Figures 12c and 12d). These variables were crossed to yield the four conditions shown in Figure 12

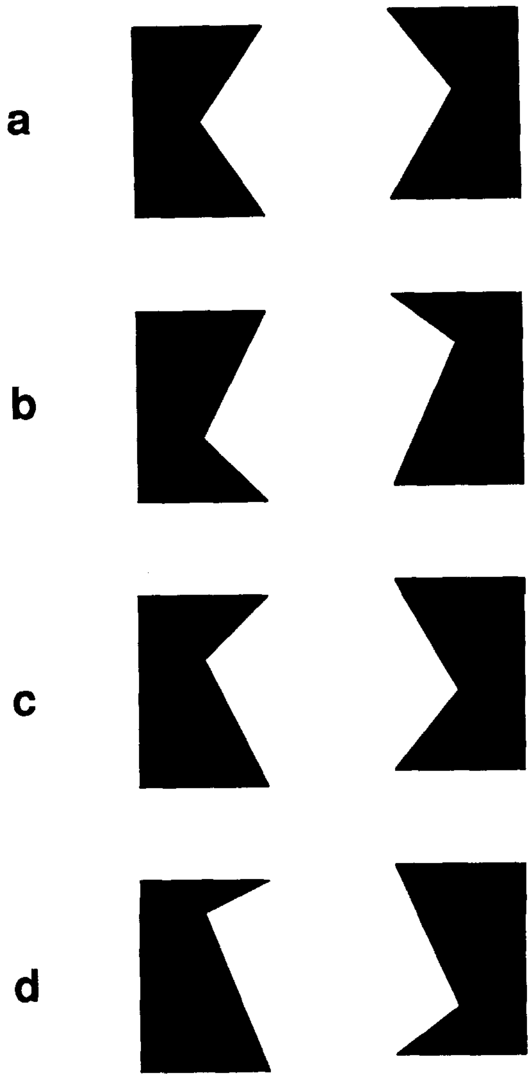

Figure 12. Example displays from the four conditions of Experiment 6: (a) small-offset congruent, (b) large-offset congruent, (c) small-offset incongruent, (d) large-offset incongruent.

(small offset congruent, large offset congruent, small offset incongruent, and large offset incongruent).

If object offset judgments are affected by apex offset, a cost should be found in the incongruent conditions. If such interference depends on the relative discriminability of the apex offset in the manner suggested by the horse-race account, it should be found only with the large apex offset. The hierarchical scheme for position coding suggests that apex offset should not exert any effect on judgments of object offset.

Procedure. Subjects were shown a diagram of the displays and told to press the key on the side of the lowest object. They were told to ignore the apex positions because these were uninformative. The timing of events and the feedback were as before. There were six blocks of 100 trials each. Within each block, the conditions were equiprobable and interleaved in a different pseudorandom sequence for each subject. Half the trials in each block required a response with the left hand, half with the right. The upper and lower RT cutoffs removed 7.4% of the RT data.

Results

The means of subjects' median RTs for the four conditions (small offset congruent, large offset congruent, small offset incongruent, large offset incongruent) were 406, 409, 406, and 410 ms respectively, with error rates of 6.6%, 6.9%, 5.8%, and 7.5%. It is clear that there were no effects of apex offset. Two-way (Congruency × Apex Offset) ANOVAs on both the RT data and error data found that no main effects or interactions approached significance.

Discussion

Irrelevant apex offsets had no effect on judgments of object offset, regardless of their size. Even the large apex offset, which Experiment 5 found to be discriminated more rapidly than the current object offset, produced no interference. This is problematic for a simple horse-race account of interference. The absence of any interference contrasts with the substantial interference that the same object offset produced in Experiments 1, 3, and 4 on judgments of the small apex offset. This asymmetry in the pattern of part and object position interference is consistent with a hierarchical scheme for position coding, whereby the coding of part positions depends on object positions but not vice versa.

However, it might be suggested that even though the large apex offset was more salient than the object offset (as we found in Experiment 5), this difference, though in the right direction, was insufficient to produce interference. An elaborated horse-race theory could plausibly be proposed, on which incongruent distracting information does not interfere whenever it is derived more quickly than the target information, but only when it is ahead by a certain (unspecified) amount. Although the large apex offset was derived more rapidly than the object offset in Experiment 5, the difference was smaller than that between the small apex offset and the object offset (see Figure 11). One could therefore speculate that the object offset interfered with small apex offset judgments in Experiments 1, 3, and 4 simply because the difference in salience for this pair of attributes is so great rather than because of an inherent asymmetry in part and object position interference.

With the modification that differences in processing speed must exceed a certain amount, the horse-race theory no longer predicts that interference is inevitable from an incongruent attribute extracted faster than the target attribute and could therefore accommodate the present experiment. However, this revised account still predicts that interference should never be found from an incongruent attribute that is extracted slower than the target attribute, for instance, from the current object offset on judgments of the large apex offset. Such interference remains a possibility on the hierarchical coding account because the derivation of part locations is held to depend logically on the derivation of object locations.

Accordingly, our final experiment examined whether the current object offset could interfere with judgments of the large apex offset. The design of Experiment 1 was repeated, with the difference that the shapes in each display were taller (i.e., the same height as in Experiments 5 and 6) so that the large apex offset (2.4°) could be used for one group of subjects. An additional group of subjects was run with the small apex offset (.8°) to ensure that our earlier findings with this offset (in Experiments 1, 3, and 4) could be replicated with the taller displays.

Experiment 7

Method

The methodological details were as for Experiment 1 with the following exceptions.

Subjects. Twenty-four new subjects (11 men and 13 women) came from the same source as Experiments 1–4. All reported normal or corrected-to-normal vision.

Design. A mixed design was used with one between-subjects variable (whether the apex offset was small, i.e., 0.8°, or large, i.e., 2.4°) and one within-subjects variable that had the same four conditions as Experiment 1: single, double baseline, double congruent, and double incongruent (see Figure 13). The shapes in each display were 3.6° in height, as for Experiments 5 and 6.

Results

The RT cutoffs discarded 7.8% and 4.8% of the data from the small and large apex offset groups, respectively. The means of subjects' median RTs are shown in Figures 14a and 14b for the four within-subject conditions, together with the associated mean error rates. Data from the small apex offset group are shown in Figure 14a, and those from the large apex offset group are shown in Figure 14b. A two-way mixed ANOVA on the RT data showed a highly significant effect of group, $F(1, 22) = 25.3$, $p < .001$, of condition, $F(3, 66) = 48.2$, $p < .001$, and a significant interaction, $F(3, 66) = 12.0$, $p < .001$. This confirms that subjects were faster to respond to the larger apex offsets and that the effects of congruency in the two-object conditions were reduced when judging the larger apex offset (cf. Figures 14a and 14b).

Wilcoxon tests were used for planned pairwise comparisons between conditions, with the following results in the small apex-offset group: Single versus double baseline, $t = 0$, $p < .01$; double congruent versus double incongruent, $t = 2$, $p < .01$. Similar results were also found in the large apex offset group: Single versus double baseline, $t = 1$, $p < .01$; double congruent versus double incongruent, $t = 8$, $p < .05$.

A two-way ANOVA on the error data showed no effect of group, $F(1, 22) = 0.1$, a highly significant effect of condition, $F(3, 66) = 12.9$, $p < .001$, and no interaction, $F(3, 66) = 0.8$. Thus, the error data followed the same general pattern in the two groups. Subsequent Wilcoxon tests gave the following results in the small apex offset group: single versus double baseline, $t = 3$ (with 1 tie), $p < .01$; double congruent versus double incongruent: $t = 18$, $p = .06$. Similar results were found in the large apex offset group: single versus double baseline, $t = 0$, $p < .01$; double congruent versus double incongruent, $t = 13.5$, $p < .05$.

Discussion

The results of Experiment 1 were completely replicated in the small apex offset group, as would be expected because the only change was our use of taller displays. More important, a qualitatively similar pattern was observed in the easier task performed by the large apex offset group. A cost of attending to two objects rather than one object was found in both cases, as previously shown in Experiments 1–4. For current purposes, the more important result was that object offsets affected judgments of apex position, as in Experiments 1, 3, and 4. This applied even for the large apex offset group, in which the distracting object offset was less salient than the target apex offset (as shown in Experiment 5). This result is inconsistent with any horse-race account of interference.

The finding that incongruent object offsets disrupt judgments of apex offset (Experiments 1, 3, 4, and the present study), although the reverse does not occur (Experiment 6), suggests that the interference arises from the logical dependency of part position coding on object positions, as envisaged by our hierarchical coding hypothesis. Certainly, the pattern of results is not explicable solely in terms of the relative processing speed for the target attribute and the distracting attribute. In Experiment 5, we observed that the large apex offset could be derived more rapidly than the current object offset. The pattern of interference between these two attributes is the opposite of the pattern predicted by this difference in processing speed alone. The slower derived object offset interfered with judgments of the large apex offset in the present study, whereas the more rapidly derived large apex offset did not interfere with judgments of object offset in Experiment 6.

Although any complete explanation of the pattern of interference in terms of relative salience is impossible, salience clearly had an effect in the present experiment; the object offset produced less disruption for judgments of the larger and more salient apex offset than for the small apex offset. This is scarcely surprising. We are not proposing that relative salience has no role in the amount of interference which is observed. For example, a tiny and indiscriminable object offset would obviously produce no interference with any judgment. Our claim is simply that relative salience is inadequate as an explanation for the asymmetry in object–part interference, whereby object locations can affect judgments of part locations but not vice versa. Our results join a growing body of evidence that suggests that interference from distracting information cannot be explained simply in terms of the relative speed of extraction for relevant and irrelevant information (e.g., Dunbar & MacLeod, 1984; Glaser & Dungelhoff, 1984; Pfaf, Van der Heijden, & Hudson, 1990; Robertson & Lamb, 1991).

General Discussion

Position judgments about parts of one object were more rapid than equivalent judgments about two objects even though the positions to be compared were the same for one- and two-object displays. This two-object cost was found in

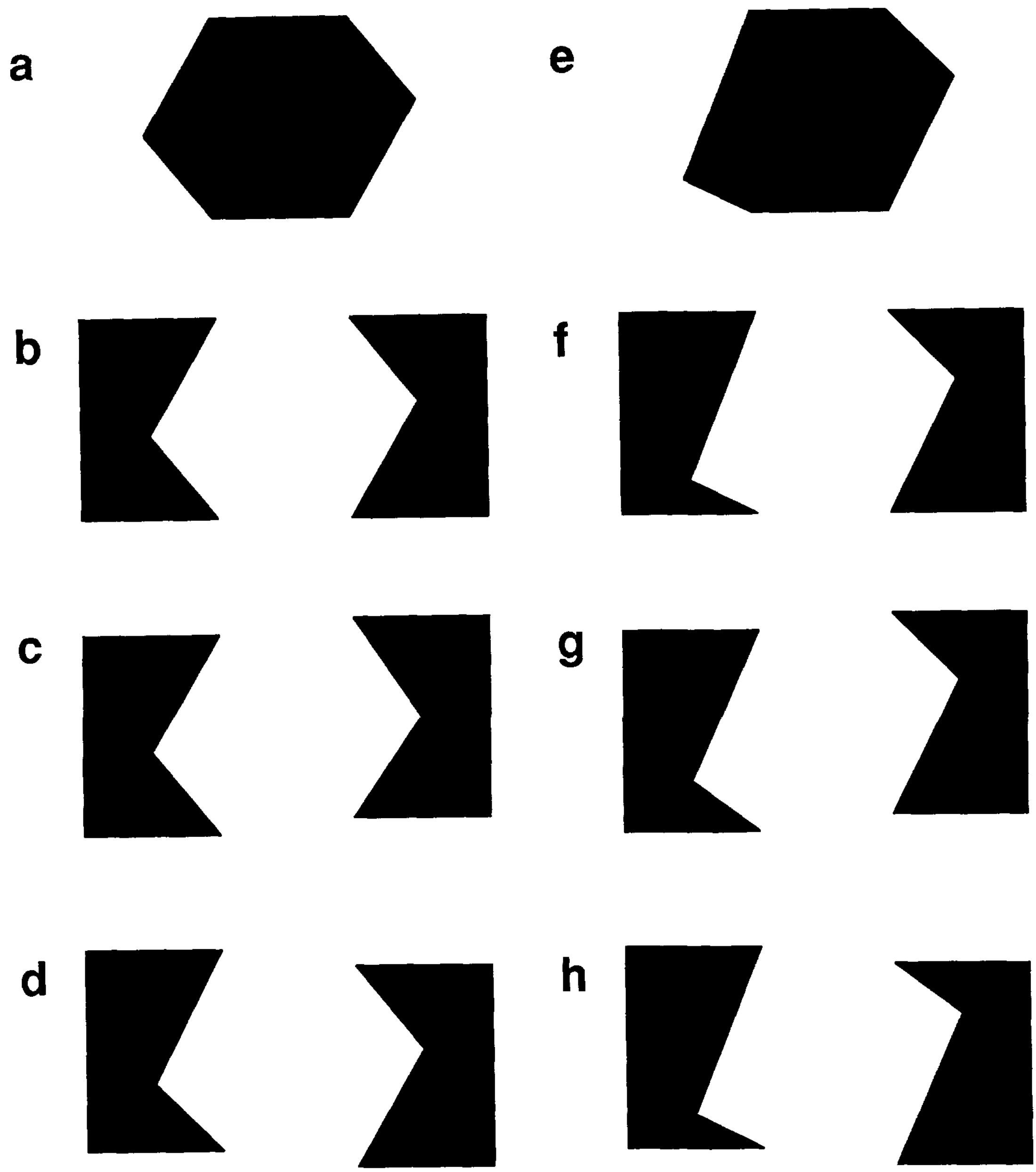

Figure 13. Example displays from the four within-subject conditions for the small apex offset group of subjects (Panels a–d) and the large apex offset group (Panels e–h) in Experiment 7. (For both groups, the within-subject conditions were as in Experiment 1: Panels a and e, single; Panels b and f, double baseline; Panels c and g, double congruent; Panels d and h, double incongruent.)

each of five experiments. Moreover, this effect was even found when the one- and two-object displays were physically identical in every respect but parsed as one or two objects according to the subjects' perceptual set (Experiment 2).

Judgments of the relative position of two points require that subjects establish a reference frame or anchor point for the comparison (Garnham, 1989; Levelt, 1984). According to our account, the natural reference points for the visual system when judging part positions are provided by the global locations of each object. We propose that spatial information is routinely represented in two different ways in the visual system. First, a scene-based description of space represents the location of objects within a scene. Second, an object-based description is produced to describe the relative positions of parts of each object. Such a hierarchical representation of space may parallel the division of the primate visual system into a scene-based dorsal stream and an object-based ventral stream. This is a reformulation of the traditional division into "what" and "where" visual systems. Such a model predicts the two-object cost as follows. Derivation

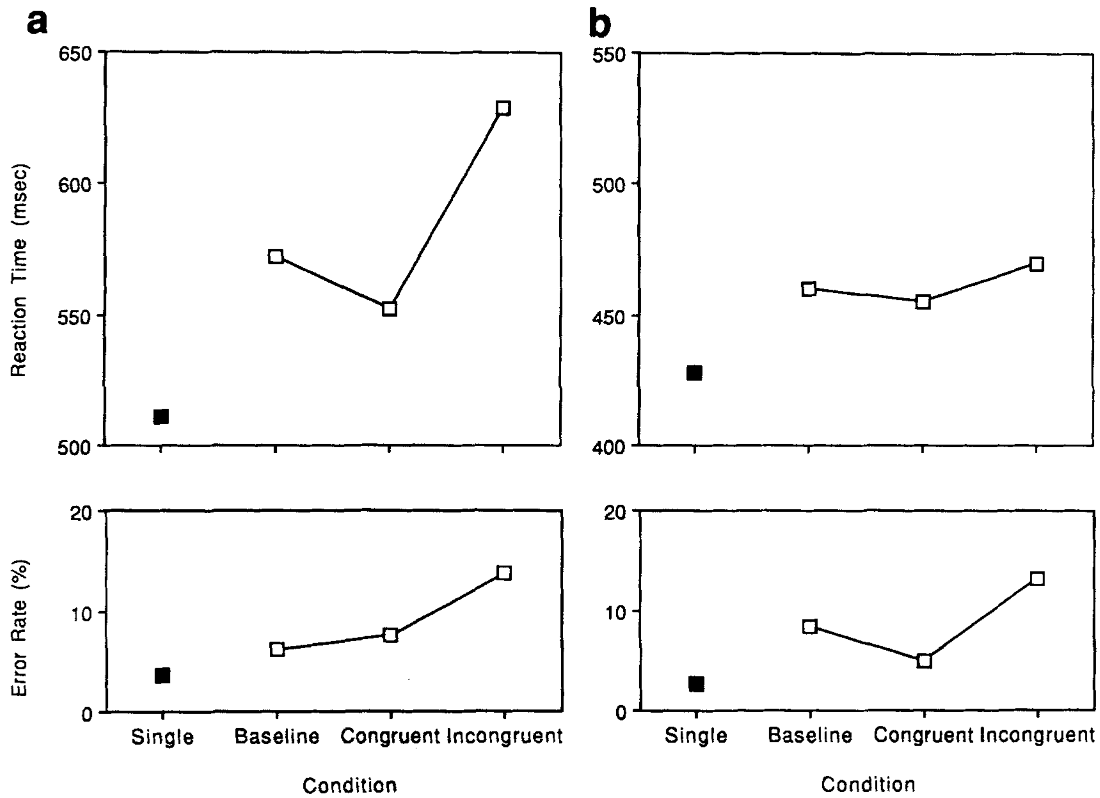

Figure 14. Mean reaction times (top panel) and error rates (bottom panel) for Experiment 7: (a) small apex offset group, and (b) large apex offset group.

of the relative position of parts of the same object is direct because such a representation is found within the object-based spatial system. However, determination of the relative location of parts of two objects is indirect because such information is not routinely derived by either the object-based or the scene-based systems. Such a two-object judgment requires that parts of objects be represented in scene-based terms, which is not normally the case. The indirectness of such a computation leads to an increased error rate and RT.

The second major finding of the current studies is that object locations can exert an interference effect on determining the relative location of parts from different objects. The proposed hierarchical nature of spatial representation clearly predicts such an effect because determining the scene-based locations of object parts is predicated on the scene-based descriptions of their parent objects.

The third major finding, the lack of interference from object part position on judgments of object location, is also consistent with the hierarchical view. The locations of objects are routinely computed in a scene-based fashion. No reference needs to be made to object-centered shape descriptions of the relative location of parts within each object. Therefore information in the object-based representation will not interfere with judgments of object position. The observed asymmetry in interference effects was not due to differences in the physical salience of the object and part positions (as indexed by processing speed). Object offsets that were less salient on this index than the apex offset to be judged nevertheless produced interference (Experiment 7). By contrast, judgments of object position were not susceptible to interference from apex offsets even when the latter were more salient than the object offset (Experiment 6).

Much previous work has examined the interfering effects of shape information at different levels of spatial scale. Through the use of large letters composed of smaller letters, initial work found that the identity of the large letter interfered with response to the identity of the smaller letters more than the reverse (Navon, 1977). This has been taken as evidence that global shape information is derived before local shape (although see Robertson & Lamb, 1991, for a summary of recent qualifications). This global precedence hypothesis is distinct from our account in several ways, especially the following two.

First, the relation between the local shapes and the global shape is entirely arbitrary in the Navon paradigm. The identity of the local letters can be changed without altering the identity of the global letter. Similarly, the identity of the global letter can be altered while retaining the identity of the local letters. In contrast, there is a logical dependency between the relative position of objects and their parts. Although an individual part can be moved without altering the position of an object, an individual object cannot be moved without altering the position of its parts relative to the parts of another object. For example, consider the objects in Figure

10. Altering the location of the apices (e.g., Figure 10a vs. 10b) need not alter the position of the parent objects. However, altering the location of any object would necessarily alter the location of its apices relative to other objects. This logical dependency accords with the dependency encapsulated in our hierarchical coding proposal, whereby the coded position of parts depends on the coded position of objects but not vice versa.

The second major difference is that the global precedence hypothesis refers only to shape, whereas our hierarchical coding scheme refers to position in both object- and scene-based coordinates. On our theory, the shape of an object is coded in a different description from its scene-based location. We have identified a performance difficulty in comparing the positions of contours from different objects; shape recognition does not require such comparisons.

As we have noted, the two-object costs we have found accord with object-based theories of attention (Duncan, 1984), as well as our hierarchical position-coding hypothesis. Are these two accounts rivals? We suggest that they are compatible, albeit expressed at different levels of specificity. Our hypothesis is a particular suggestion about the coding of position and shape in the visual system. Object-based views of attention capture one consequence of this coding, a relative difficulty in making position judgments for contours from separate objects. The two accounts might be distinguished by investigating whether comparable two-object costs are found for judgments of attributes other than position. However, such results might be obtained because the new attribute was also coded hierarchically. In other words, hierarchical coding of local part characteristics relative to global object characteristics might be a general principle that leads to object-based limits on attention.

Consistent with this view, Vecera and Farah (in press) have recently proposed that whether object-based versus space-based attentional limits are observed depends on whether the task in question taps object-centered representations such as shape descriptions. Certainly, although space-based and object-based models of attention are conventionally cast as rivals, it is possible that both are appropriate characterizations but of different selective mechanisms. Our experiments may have emphasized object- rather than space-based selection because, as we have argued, the within-object judgments could readily be made through object-centered shape descriptions. Our experiments do not rule out a role for selection by location in other situations. Nevertheless, they demonstrate limitations on visual processing which cannot be captured by any monolithic conception of attention as a spatial spotlight.

References

Baylis, G. C., & Driver, J. (1992). Visual parsing and response competition: The effects of grouping. *Perception & Psychophysics, 51,* 145–162.

Baylis, G. C., Driver, J., & McLeod, P. (1992). Movement and proximity constrain miscombinations of color and form. *Perception, 21,* 201–218.

Broadbent, D. E. (1982). Task combination and selective intake of information. *Acta Psychologica, 50,* 253–290.

Driver, J., & Baylis, G. C. (1989). Movement and visual attention: The spotlight metaphor breaks down. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 448–456.

Driver, J., & Baylis, G. C. (1991). Target–distractor separation and feature integration in visual attention to letters. *Acta Psychologica, 76,* 101–119.

Dunbar, K., & MacLeod, C. M. (1984). A horse race of a different color: Stroop interference patterns with transformed words. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 622–639.

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General, 113,* 501–517.

Farah, M. J., Brunn, J. L., Wong, A. B., Wallace, M. A., & Carpenter, P. A. (1990). Frames of reference for allocating attention to space: Evidence from the neglect syndrome. *Neuropsychologia, 28,* 335–347.

Garnham, A. (1989). A unified theory of the meaning of some spatial relational terms. *Cognition, 31,* 45–60.

Glaser, W. R., & Dungelhoff, F. J. (1984). The time course of picture-word interference. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 640–654.

Hasselmo, M. I., Rolls, E. T., Baylis, G. C., & Nalwa, V. (1989). Object centered encoding by face-selective neurons in the superior temporal sulcus of the monkey. *Experimental Brain Research, 75,* 417–429.

Hoffman, D. D., & Richards, W. A. (1984). Parts of recognition. *Cognition, 18,* 65–96.

Julesz, B., & Papathomas, T. V. (1984). On spatial frequency channels and attention. *Perception & Psychophysics, 36,* 398–399.

Kahneman, D., & Henik, A. (1981). Perceptual organization and attention. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 181–211). Hillsdale, NJ: Erlbaum.

Kanizsa, G. (1979). *Organization in vision: Essays on Gestalt perception.* New York: Praeger.

Kramer, A. F., & Jacobson, A. (1991). Perceptual organisation and focused attention: The role of objects and proximity in visual processing. *Perception & Psychophysics, 50,* 267–284.

Levelt, W. J. M. (1984). Some perceptual limitations on talking about space. In A. J. van Doorn, W. A. van de Grind, & J. J. Koenderink (Eds.), *Limits in perception: Essays in honour of Maarten A. Bouman.* Utrecht, The Netherlands: VNU Science Press.

Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organisation of three-dimensional shapes. *Proceedings of the Royal Society of London, B200,* 269–294.

Mishkin, U., Ungerleider, L., & Macko, K. A. (1983). Object vision and spatial vision: Two critical pathways. *Trends in Neuroscience, 6,* 414–417.

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

Pfaf, R. H., Van der Heijden, A. H. C., & Hudson, P. T. (1990). SLAM: A connectionist model for attention in visual selection tasks. *Cognitive Psychology, 22,* 273–341.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32,* 3–26.

Prinzmetal, W. (1981). Principles of feature integration in visual perception. *Perception & Psychophysics, 30,* 330–340.

Rabbitt, P. M. A. (1966). Errors and error correction in choice response tasks. *Journal of Experimental Psychology, 71,* 264–272.

Robertson, L. C., & Lamb, M. R. (1991). Neuropsychological contributions to theories of part–whole organization. *Cognitive Psychology, 23,* 299–330.

Rock, I. (1984). *Perception.* New York: Scientific American Press.

Rubin, E. (1921). *Visuell wahrgenommene figuren.* Copenhagen: Glydendalske.

Vecera, S. P., & Farah, M. J. (in press). Does visual attention operate on objects or locations? *Journal of Experimental Psychology: General.*

Watt, R. J. (1988). *Visual processing: Computational, psychophysical, and cognitive research.* Hillsdale, NJ: Erlbaum.

Wilkinson, L. (1988). *SYSTAT: The system for statistics.* Evanston, IL: Systat.

Received August 27, 1991
Revision received September 22, 1992
Accepted September 23, 1992 ■

Ethical Standards for the Reporting and Publishing of Scientific Information

The following ethical standards are extracted from the "Ethical Principles of Psychologists and Code of Conduct," which appeared in the December 1992 issue of the *American Psychologist* (Vol. 47, No. 12, pp. 1597–1611). Standards 6.21–6.26 deal with the reporting and publishing of scientific information.

6.21 Reporting of Results

(a) Psychologists do not fabricate data or falsify results in their publications.

(b) If psychologists discover significant errors in their published data, they take reasonable steps to correct such errors in a correction, retraction, erratum, or other appropriate publication means.

6.22 Plagiarism

Psychologists do not present substantial portions or elements of another's work or data as their own, even if the other work or data source is cited occasionally.

6.23 Publication Credit

(a) Psychologists take responsibility and credit, including authorship credit, only for work they have actually performed or to which they have contributed.

(b) Principal authorship and other publication credits accurately reflect the relative scientific or professional contributions of the individuals involved, regardless of their relative status. Mere possession of an institutional position, such as Department Chair, does not justify authorship credit. Minor contributions to the research or to the writing for publications are appropriately acknowledged, such as in footnotes or in an introductory statement.

(c) A student is usually listed as principal author on any multiple-authored article that is substantially based on the student's dissertation or thesis.

6.24 Duplicate Publication of Data

Psychologists do not publish, as original data, data that have been previously published. This does not preclude republishing data when they are accompanied by proper acknowledgment.

6.25 Sharing Data

After research results are published, psychologists do not withhold the data on which their conclusions are based from other competent professionals who seek to verify the substantive claims through reanalysis and who intend to use such data only for that purpose, provided that the confidentiality of the participants can be protected and unless legal rights concerning proprietary data preclude their release.

6.26 Professional Reviewers

Psychologists who review material submitted for publication, grant, or other research proposal review respect the confidentiality of and the proprietary rights in such information of those who submitted it.