Role of perceptual organization
while attending in depth

MARK T. MARRARA and CATHLEEN M. MOORE
Pennsylvania State University, University Park, Pennsylvania

Seven experiments were conducted in order to explore the conditions under which visual attention can be allocated in depth. In each experiment, observers were cued to the most likely target location in stereoscopic depth displays, and targets could appear in either the cued location or in another location. In Experiments 1 and 2, we show that previous failures to observe effects of cuing in depth may have depended on the specific timing characteristics of the displays. In Experiments 3, 4, and 5, we eliminated the hypothesis, suggested by Experiments 1 and 2, that in order to allocate attention in depth, attention must be allocated to a specific object token. Experiment 6 provides evidence that, in the absence of other organizing information (e.g., color), attention can be allocated in depth on the basis of the surface information available in the display. Finally, in Experiment 7, we demonstrate that, although sufficient, surface information is not necessary for the allocation of attention; color supported the allocation of attention across multiple items that failed to fall on a single coherent surface in depth. Together, these findings suggest that attention in depth, like attention in two-dimensional displays, is determined by the perceptual organization of the display.

Allocating attention to a given region in the visual field can aid in the processing of information that subsequently appears at that location. For example, knowledge of what region in the sky a comet is to be found in can increase the probability of seeing that comet. This general phenomenon has been demonstrated and studied extensively in the lab. For example, numerous experiments using simple detection as well as discrimination tasks have shown that people are faster and more accurate in responding to stimuli that appear in cued locations than to stimuli that appear in uncued locations (see, e.g., Posner, Nissen, & Ogden, 1978). Findings of this sort are thought to reflect the fact that attention is allocated to a cued location (e.g., Posner et al., 1978).

Visual attention has been conceptualized in several different ways. One common metaphor is that attention is like a "spotlight" that can be moved through the visual field (see, e.g., Broadbent, 1982; Posner, Snyder, & Davidson, 1980). All information located within the spotlight is selected for further processing, whereas information located outside the spotlight is ignored. Another conceptualization is that attention is like a zoom lens that

can be focused to different sizes (e.g., Eriksen & St. James, 1986). This metaphor was inspired by studies that have shown that the distribution of attention within the visual field can vary in size depending on the nature of the task. Finally, on the basis of studies in which reaction time to stimuli has been varied as a function of the distance from the cued location, attention has also been described as a gradient, with the highest processing efficiency at the center of the distribution and the least processing efficiency at peripheral locations (e.g., LaBerge & Brown, 1989).

An important issue concerning visual attention is the nature of the representation within which it is allocated. In particular, does the representation include information about all three spatial dimensions (left–right, up–down, near–far) or is the allocation of attention based entirely on a two-dimensional representation of the visual world? Because the world is projected onto a two-dimensional array of photoreceptors at the retina, one might hypothesize that attention is allocated without regard to the third dimension. On the other hand, because one of the presumed functions of the visual system is to facilitate navigation within a three-dimensional world, one might hypothesize that attention is allocated within a representation that includes depth information. One reason to suspect that attention is allocated within a three-dimensional representation is that there is evidence from visual-search tasks that three-dimensional interpretations of two-dimensional displays are extracted preattentively (Enns & Rensink, 1990, 1991; Humphreys, Keulers, & Donnelly, 1994; Kleffner & Ramachandran, 1992; Rensink & Enns, 1998). The cuing studies cited thus far, however, are mute with regard to the question, because all of the stimuli in those studies were presented within a single two-dimensional frontoparallel plane. The question comes down to the fol-

lowing: Can attention be cued to a spatial location that is defined by depth?

Downing and Pinker (1985) provided some of the earliest evidence suggesting that attention can be allocated in depth. They presented participants with two curved rows of lights, with each row set at a different distance in depth from the observer. The participants viewed the display monocularly and responded as quickly as possible to a light that appeared in one of eight possible locations. At the beginning of each trial, the participants fixated a liquid-crystal display that was located between the two rows of lights. A symbolic cue that indicated the location in which the target would most likely appear was then presented on this display. The participants were slower to detect the onset of the target light when the target appeared at uncued locations on a different depth plane than when the target appeared at uncued locations on the same depth plane. The magnitude of this effect increased with increasing retinal eccentricity, and the effects were greater when targets appeared at far uncued locations than when they appeared at near uncued locations. The most important implication of these results with regard to the present study is that attention seemed to be allocated to a given depth, suggesting that attention is allocated on the basis of a representation that includes information about all three spatial dimensions (see also Gawryszewski, Riggio, Rizzolatti, & Umiltà, 1987).

Although Downing and Pinker's (1985) results suggest that attention can be allocated in depth, there is a possibility that these results could have been caused by peripheral factors, rather than by attention. Ghirardelli and Folk (1996) argued that the distance between the stimuli in the displays that were used by Downing and Pinker and in similar studies were sufficiently large (the equivalent to 20°–30° in x, y space according to their calculations) that accommodation and convergent eye movements may have been necessary, or at least helpful. Moreover, the amount of time between the onset of the cue and the onset of the target was sufficiently long (400–800 msec) for both to occur on at least some of the trials. These possibilities make Downing and Pinker's results difficult to interpret: Were they caused by the allocation of attention or by the focusing of the eyes?

To eliminate the possibility of these alternative explanations, Ghirardelli and Folk (1996) ran a similar study using stereoscopic displays in which depth was defined by horizontal binocular disparity (a slight difference in position along the x-axis of two identical images between the two eyes). With these displays, it was possible to present stimuli at nearly identical x, y positions (except for the binocular disparity), but at perceptually different z positions. Because the stimuli were presented physically at the same depth plane (i.e., on the computer monitor) and because the perceptual depths achieved through binocular disparity were not great (less than 4°), accommodation and convergent eye movements were rendered irrelevant. Finally, in addition to using stereoscopically defined displays, Ghirardelli and Folk used

a short cue-to-target interval: The cue appeared for 100 msec, and after a 50-msec interval, the target appeared for 50 msec. There was, therefore, insufficient time between cue onset and target onset for peripheral changes to systematically affect performance. Thus, any cuing effects in depth that Ghirardelli and Folk observed could be safely attributed to attention having been allocated in depth.

Having thus controlled for peripheral explanations, Ghirardelli and Folk (1996) found no evidence that attention can be allocated in depth. Targets that appeared in the cued depth location were not reliably identified more quickly or accurately than targets that appeared in the uncued depth location. These results suggest that attention may be allocated on the basis of a representation that does not include information about all three dimensions.

Around the same time that Ghirardelli and Folk (1996) reported their null results, Hoffman and Mueller (1997) reported a study that included all of the same controls for peripheral explanations, but did find evidence for attention being allocated in depth. Ghirardelli and Folk, however, used a different type of target stimulus. In particular, whereas Ghirardelli and Folk presented alphanumeric characters for a very brief period of time (50 msec), Hoffman and Mueller used figure-eight place holders that could turn into alphanumeric stimuli (as on a digital clock). Although in Hoffman and Mueller's study the target character itself was present for only a short period of time (70 msec), the placeholder was present from the early stages of the trial.

Our purpose in the present study was to investigate how attention is allocated in depth (assuming that it can be) by identifying the critical differences between those studies in which cuing effects in depth have been found and those in which it has not. Experiment 1 was a replication of Ghirardelli and Folk's (1996) study. In Experiment 2, we introduced the placeholder figure eights that Hoffman and Mueller (1997) used. When the placeholders were presented for a sufficiently long period of time before the target was revealed, cuing effects in depth were observed. Otherwise, no cuing effects in depth were observed. These results suggest that the timing of the stimulus displays was critical to whether or not cuing effects were observed. In addition, however, the placeholders provided object tokens to which the participants could direct their attention, and such tokens may have been necessary for the allocation of attention in depth. In Experiments 3–6, we ruled out this hypothesis, but the data suggest that coherent surface information, as defined by common depth, might have been necessary for the allocation of attention in depth. Finally, in Experiment 7, we showed that, although surface information is sufficient, attention can be allocated to multiple items that fail to fall on a coherent surface in depth, when other organizing information, such as color, is present in the displays. Together these results imply that, like attention in two-dimensional displays, perceptual organization plays an important role in the allocation of attention in depth.

EXPERIMENT 1

The purpose of Experiment 1 was to replicate Ghirardelli and Folk's (1996) study in which cuing effects in depth failed to occur.

Method

Participants

Fifteen participants from Pennsylvania State University undergraduate subject pool were tested. In compensation they received extra credit in a psychology course. All reported normal or corrected-to-normal visual acuity and color vision, and all were naive as to the purpose of the experiment before participating. Individuals who could not see depth on the basis of binocular disparity were excluded from the study. This was tested by using a custom screening procedure that is described in more detail below.

Apparatus

Stimuli were presented on a 21-in. Nanao FlexScan F2-21 EX color monitor that was driven by a Number 9 Verge 3D PCI graphics card. Trial events and data collection were controlled by an IBM-compatible pentium-based personal computer. Depth was simulated with the use of a Stereographics CrystalEyes II stereo-display system. With this system, separate left- and right-eye images can be presented through a set of liquid-crystal shutter glasses synchronized with the monitor. The presentation rate was 60 Hz per eye.

Task

On each trial, a target letter appeared in one of two possible locations. The task was to report, as quickly as possible, whether the target letter was a T or an L, by pressing the left or right button on a hand-held button box, respectively.

Stimuli

The stimuli included two rectangles (one horizontal, one vertical) that subtended approximately $5.2° \times 1.6°$ from a typical viewing distance of 70 cm. All displays in this and the following experiments consisted of outlined color figures drawn on a black background. The vertical rectangle was magenta (IBM standard palette designation 13); the horizontal rectangle was green (IBM standard palette designation 10). A $0.4° \times 0.4°$ gray (IBM standard palette designation 7) square was presented at the center of the screen as a fixation marker. The target letters T and L subtended $0.4° \times 0.3°$ and were presented at one of four possible orientations (0°, 90°, 180°, or 270°).

Two types of display were used: two-dimensional (2-D) and three-dimensional (3-D) (see Figure 1). In the 2-D displays, both rectangles appeared at the horopter (i.e., the point of zero binocular disparity, which was the depth of the monitor) but to the left and right sides of the fixation square. The locations of the rectangles were counterbalanced within each block. The distance (center-to-edge) between the fixation square and the nearest side of each rectangle was 2°. The targets were presented 2.2° to the left or right of fixation within one of the two rectangles.

In the 3-D displays, the two rectangles were both centered at fixation, but they appeared in different depth planes, one in front and one behind the horopter. The location of the rectangles was counterbalanced within each block. The virtual distance of the rectangles from fixation was approximately 1.75°, which corresponded to a binocular disparity of $-0.15°$ and $+0.15°$ (Boff & Lincoln, 1988). The fixation square in the 3-D displays was presented at the horopter so that it appeared between the two rectangles. An additional square, which framed the stimulus configuration, was presented at the horopter to increase the likelihood that the rectangles would be perceived at different depths. This square was light gray (IBM standard palette designation 7) and subtended 7° horizontally and ver-
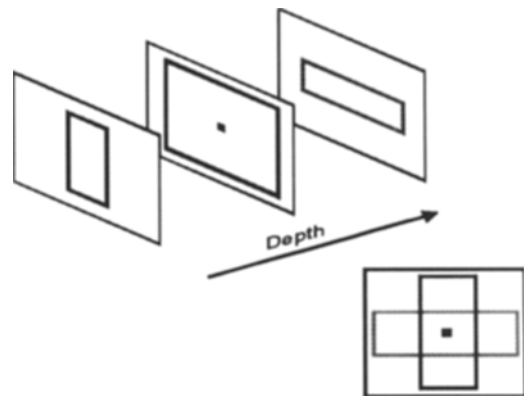
Figure 1. Three-dimensional displays from Experiment 1. Each rectangle appeared at a different depth. The lower right inset of the figure depicts how the displays would look if viewed from straight on.

tically. The targets were presented in the center of one of the two rectangles in either the front or the back depth plane.

Most of the cues consisted of one of the two rectangles flashing bright white (IBM standard palette designation 15) for 50 msec. On some trials (neutral trials see below) the cue consisted of the fixation square that increased in size by 0.1° for 50 msec.

Design

A 2 (display type: 2-D, 3-D) $\times$ 3 (validity: valid, neutral, invalid) within-subjects design was used. Display type was manipulated between blocks of trials, with the participants receiving three contiguous blocks of each of the two display types. The order of these blocks was counterbalanced across participants.

Validity was defined by the relationship between the location of the cue and the location of the target. One third of the trials were neutral trials, in which no location was indicated and the target appeared equally often in the two possible locations. Of the remaining trials, 80% were valid trials, in which the target appeared in the rectangle that was cued, and 20% were invalid trials, in which the target appeared in the rectangle that was not cued.

Each block consisted of 60 trials from which data were collected, and the participants completed six blocks. This resulted in 96, 24, and 60 observations in the valid, invalid, and neutral conditions, respectively, for each of the two display types.

Finally, each block began with three practice trials randomly selected from the set of possible trial types. If the participants pressed the wrong button on any trial, responded in less than 150 msec, or failed to respond within 1,500 msec, the trial was scored as an error. Each error trial was followed by a "recovery" trial that was chosen randomly from the set of possible trial types. The data from practice and recovery trials were not included in the analyses.

Procedure

Each session consisted of a screening procedure and the experiment proper. The screening procedure was designed to identify participants who were unable to perceive depth on the basis of binocular disparity. All participants completed both the screening and the experiment proper. The data for those who did not meet criterion in the screening procedure were subsequently excluded from further analyses.

Screening procedure. The screening procedure consisted of two blocks of 20 trials each. The first block was practice. The displays were the same as those used in the 3-D condition. The fixation display for each trial appeared for 1,500 msec. The fixation square was then replaced by a target dot that appeared in either the front or the

back depth plane. The target remained on the screen until a response was made. The task was to indicate in which depth plane the target appeared, by pressing the left or right button on the button box to indicate front or back, respectively. A response was followed by a blank screen, which was replaced 1,000 msec later by the fixation display for the next trial. The participants were instructed to respond as accurately as possible and to not worry about how fast they made their responses. Criterion performance was 90% in the second block.

Experimental procedure. The experiment proper began with a set of written instructions that described the task and told the participants that the target would appear in the cued location on most of the trials. The instructions emphasized the importance of trying to use the cue to allocate attention while maintaining fixation on the central fixation square throughout each trial.

After the instructions were presented, the participants completed two 30-trial practice blocks, one of each display type. The order in which the participants received these practice blocks was counterbalanced. After practicing, the participants completed six blocks of 60 trials each. As long as they maintained at least 95% correct accuracy, they were instructed to respond as quickly as possible. If mean accuracy for a given block dropped below 95%, the following message appeared on the screen: "You are making too many errors. Slow down and increase accuracy."

The trial events are illustrated in Figure 2. Each trial began with the fixation display, which appeared for 1,500 msec. The cue then flashed for 100 msec. Following the offset of the cue, the fixation display reappeared for 50 msec. The fixation square then disappeared, and at the same time, the target appeared for 50 msec. This display remained on the screen either until a response was made or for 1,500 msec. An intertrial interval of 1,000 msec, during which the screen was blank, followed the participants' responses.

RT was measured from the onset of the target letter to the execution of the buttonpress. Errors were followed by three 500-msec, 1000-Hz tones, each separated by 25 msec, and a 1,000-msec "time out" during which the screen was blank.
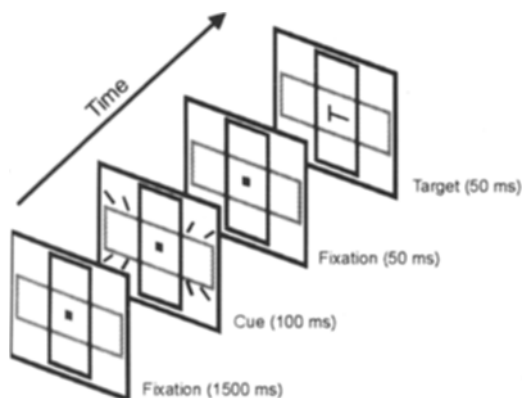
Results

Screening Procedure

All 15 participants met criterion in the screening procedure. The mean accuracy in the screening task for participants was 99.67%, with a range of 95%–100%.

Experimental Procedure

All RT analyses in this and the subsequent experiments were conducted on the basis of correct trials. If the overall mean error rate (ER) for a given participant was greater than 10%, that participant's data were excluded from further analyses. The data from all participants were included in the analyses for Experiment 1.

Response time. Figure 3 shows the mean correct RTs for each level of validity for each display type. The overall mean RT was 484 msec. A 2 (display type: 2-D, 3-D) $\times$ 3 (validity: valid, neutral, invalid) repeated measures analysis of variance (ANOVA) conducted on the mean RTs revealed a significant main effect of display type $[F(1,14) = 25.97, MS_e = 985.91, p < .001]$, as well as a significant main effect of validity $[F(2,28) = 3.93, MS_e = 443.11, p < .05]$. The display type $\times$ validity interaction was also significant $[F(2,28) = 8.69, MS_e = 196.41, p < .001]$.

Planned comparisons between the valid and invalid conditions for both display types confirmed that responses were faster on valid trials than on invalid trials for the 2-D displays [mean difference = 25 msec, $t(14) = 2.74$, $p < .01$] but not for 3-D displays [mean difference = -1.13 msec, $t(14) = 0.22$, n.s.].[1]

Error rate. The same analyses were run on the ERs. The overall mean ER was 1.64. Table 1 gives the mean ERs for each level of validity for each display type. No significant effects that were in a different direction from those in the RTs were revealed.

Discussion

Experiment 1 replicated Ghirardelli and Folk's (1996) main findings. Although a reliable cuing effect occurred with the 2-D displays, none occurred with the 3-D displays. One interpretation of these results is that visual attention is allocated within a representation that does not include information about depth. As described in the introduction, however, Hoffman and Mueller (1997) did observe reliable cuing effects in depth, under similar conditions. In Experiment 2, we tested one possible reason for these apparently conflicting results.

EXPERIMENT 2

Informal observation of the displays used in Experiment 1 suggests that, because of the short duration of the target stimulus (50 msec), it was difficult to resolve where in depth it appeared. If the participants failed to resolve the depth of the target on some or all of the trials, this could account for the lack of a cuing effect. This follows, because which depth location was cued would be irrelevant if the depth of the target could not be resolved. Notice that resolution of the target depth is not a necessary condition for the task; it is, in principle, possible to identify the letter without knowing where in depth it appeared.

As described in the introduction, one methodological difference between Ghirardelli and Folk's (1996) study and Hoffman and Mueller's (1997) study was that the latter used figure-eight placeholders that appeared before the presentation of the target stimuli. These placeholders were present from the beginning of the trial and then
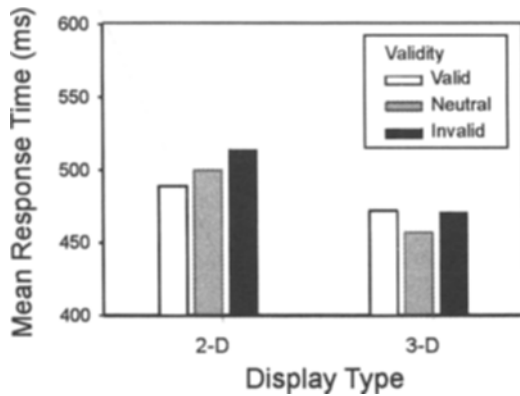

Figure 2. Timing and trial events for Experiment 1. The figure shows the three-dimensional condition.

Figure 3. Mean correct response times for Experiment 1.

changed into the target stimuli. There was plenty of time for the depth of the placeholders to be resolved, and therefore, for the depth of the target to be resolved. This might account for why Hoffman and Mueller obtained reliable cuing effects in depth, but Ghirardelli and Folk did not.

In Experiment 2, we tested this hypothesis by using figure-eight placeholders and by manipulating when they were presented. One of the two main conditions (i.e., the *long* condition) was like Hoffman and Mueller's (1997) study in that the placeholders were presented for a long time before they changed into the targets. The other condition (i.e., the *short* condition) was more like Ghirardelli and Folk's (1996) study in that the placeholders were present for only a short time before they changed into the targets. If our hypothesis is correct, we should find a reliable cuing effect in depth in the long condition but not in the short condition.

Method

Participants
Twenty-seven new participants from the same pool as that in Experiment 1 were tested.

Task
The task was the same as that in Experiment 1, except that the participants were to report whether the target was an S or an H, by pressing the left or right button on the button box, respectively.

Apparatus and Stimuli
The apparatus and stimuli were identical to those in Experiment 1, with the following exception: A placeholder object now appeared in each rectangle. These placeholders consisted of a figure eight from which certain line segments could be removed to form a target letter. The placeholders subtended 0.6° × 1.1°, and the color, depth, and orientation of a placeholder were the same as those of the rectangle in which it was located (see Figure 4). The target letters subtended 0.6° × 1.1° and could be either an S or an H.

In the 2-D displays, the two rectangles appeared at the horopter to the left and right sides of the fixation square. The distance (center to edge) between the fixation square and the nearest side of each placeholder was 2.2°. In the 3-D displays, each placeholder appeared at a different depth plane at the center of the screen, thereby replacing the fixation square. In both types of displays, the location of the rectangles and placeholders was counterbalanced within each block.

Design
A 2 (display type: 2-D, 3-D) × 2 (validity: valid, invalid) × 2 (display length: short, long) within-subjects design was used. Display type was manipulated between blocks of trials, with the participants receiving three contiguous blocks of each of the two display types. The order of these blocks was counterbalanced across participants. In each block, 80% of the trials were valid; the remaining trials were invalid. The order of trials was randomly determined for each block. Display length refers to when the placeholders appeared on the screen. Half the trials in each block were long displays, the order of which was randomly determined for each block.

Each block consisted of 40 trials from which data were collected; the participants completed eight blocks. This resulted in 64 observations in the valid conditions and 16 observations in the invalid conditions for each display length for each of the two display types.

Finally, each block began with three practice trials that were randomly selected from the set of possible trials for that block. If the participants responded incorrectly to any trial, responded in less than 150 msec, or did not respond within 1,500 msec, the trial was scored as an error. Each error trial was followed by a "recovery" trial that was chosen randomly from the set of possible trial types. Data from the practice and recovery trials were not included in the analyses.

Procedure
Screening procedure. The screening procedure was identical to that in Experiment 1.

Experimental procedure. The experimental procedure was the same as that in Experiment 1, except for the following changes in the trial events (see Figure 5): Each trial began with the fixation display, which appeared for 2,500 msec. During the long-display length trials, the placeholders appeared on the screen from the onset of this display. The cue flashed for 100 msec. During the short-display length trials, the placeholders appeared with the onset of the cue display. After the offset of the cue, the fixation display reappeared for 50 msec. This display contained the placeholders during both display length conditions. The fixation display then disappeared; at the same time, the target appeared for 50 msec. This display remained on the screen either until a response was made or for 1,500 msec. An intertrial interval of 1,000 msec, during which the screen was blank, followed the response.

Results

Screening Procedure
Two of the 27 participants failed to meet criterion in the screening procedure. The data for 5 other participants were eliminated because their mean ERs in the main task exceeded 10%. The mean accuracy for the screening procedure for the 19 remaining participants was 99.74%, with a range of 95% to 100%.

Experimental Procedure
Response time. Figure 6A shows the mean correct RTs for each level of validity and display length for the 2-D displays, and Figure 6B shows the corresponding data for the 3-D displays. The overall mean RT was

Table 1
Error Rates (in Percentages) for Experiment 1

Display	Validity		
	Valid	Neutral	Invalid
2-D	1.32	1.56	2.78
3-D	1.46	0.78	1.94
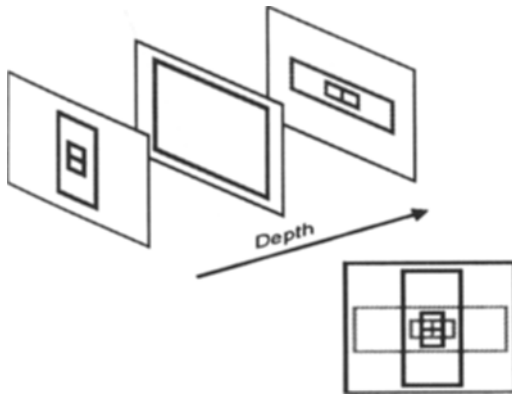
Figure 4. Three-dimensional displays from Experiment 2. Each rectangle appeared at a different depth, and each place-holder shared the color, depth, and orientation of the rectangle in which it appeared. The lower right inset of the figure depicts how the displays would look if viewed from straight on.

663 msec. A 2 (display type: 2-D, 3-D) × 2 (validity: valid, invalid) × 2 (display length: short, long) repeated measures ANOVA conducted on the mean RTs revealed a significant main effect of validity [$F(1,18) = 42.56$, $MS_e = 4{,}821.60, p < .0001$], as well as a significant display type × validity interaction [$F(1,18) = 11.57, MS_e = 3{,}225.92, p < .01$]. A significant main effect of display length was also found [$F(1,18) = 10.93$, $MS_e = 2{,}737.91, p < .01$], as well as a significant display type × display length interaction [$F(1,18) = 13.29, MS_e = 1{,}874.80, p < .01$]. Finally, the validity × display length interaction was also significant [$F(1,18) = 29.65$, $MS_e = 3{,}708.31, p < .001$].

Planned comparisons between the valid and invalid conditions for both display lengths of each display type confirmed that, with 2-D displays, responses were faster for valid trials than for invalid trials in the short display length condition [mean difference = 35 msec, $t(18) = 2.38, p < .05$], as well as in the long display length condition [mean difference = 175 msec, $t(18) = 6.35, p < .001$]. However, with 3-D displays, responses were faster for valid trails than for invalid trials in the long display length condition [mean difference = 80 msec, $t(18) = 4.75, p < .001$], but not in the short display length condition [mean difference = 5 msec, $t(18) = 0.29$, n.s.].

Error rate. The same analyses were run on the ERs. The overall mean ER was 8.65. Table 2 gives the mean ERs for each level of validity and display length for each display type. No significant effects that were in a different direction from those in the RTs were revealed.

Discussion

The results of Experiment 2 suggest that at least one reason that Ghirardelli and Folk (1996) failed to find evidence of attention in depth was that the targets may have been presented for too short a time to be easily resolved in depth. The placeholders in Hoffman and Mueller's

(1997) study eliminated this problem. This interpretation is supported by the present results in that the short and long conditions differed primarily with regard to how much time the participants were given to resolve the depth of the placeholders (which eventually turned into the targets). A reliable cuing effect was observed when a lot of time was given, but none was observed when only a short amount of time was given.

Another possible explanation for the findings from Experiment 2 is that the effectiveness of the cue was undermined on the short display length trials by the onset of the placeholders. Specifically, the appearance of the placeholders at the same time as the appearance of the cue may have caused transient attention (see, e.g., Nakayama & Mackeben, 1989) to be pulled to both locations, if not to both at once, perhaps to each equally often across trials. If this had occurred, it would have reduced the effectiveness of the cue itself for drawing attention to the intended depth location. This is certainly a possibility, and the data from the 2-D condition lend some credence to it. In particular, the cuing effect in the short display length condition is smaller than the effect in the long display length condition, indicating that the cue was not as effective in the short condition as in the long condition. However, the placeholders did not succeed in completely capturing attention because there was a reliable effect in
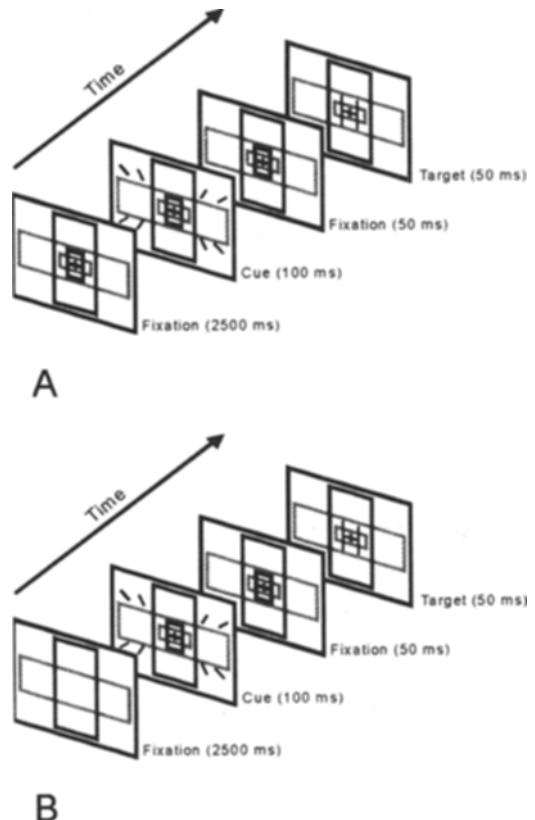
Figure 5. Timing and trial events for Experiment 2. Panel A, *long* display condition; Panel B, *short* display condition.
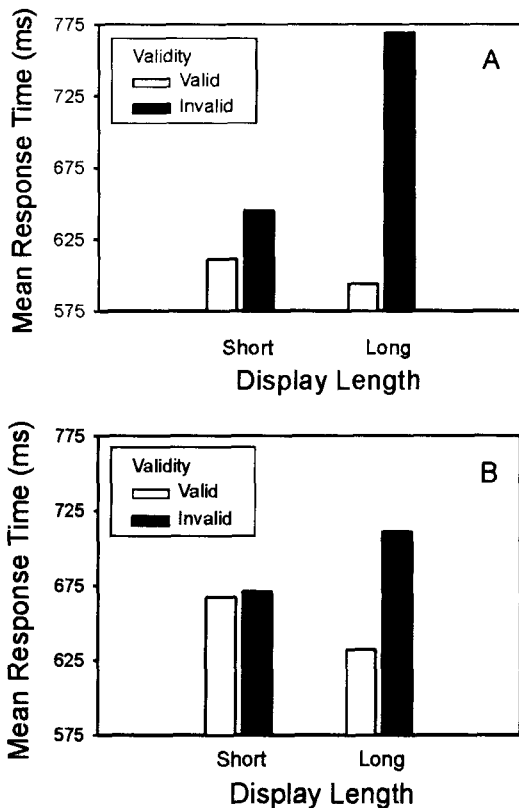
Figure 6. Mean correct response times for the two-dimensional condition (Panel A) and the three-dimensional condition (Panel B) for Experiment 2.

that short condition. This suggests that there was something more that caused problems for the short display length condition in the 3-D case. That the depth of the stimuli could not be resolved in the time available was likely the problem.

Although the results from this experiment suggest that the time that is available for the resolution of target depth is a critical variable, the introduction of the placeholders did more than allow sufficient time to resolve the depth of the targets. It also introduced object tokens into the display. The participants could have allocated their attention to the particular object that, on valid trials anyway, eventually became the target. It is possible, therefore, that this is what allowed for the allocation of attention in depth (see Hoffman & Mueller, 1997, for a similar account).

EXPERIMENT 3

In Experiment 3, we tested whether attention must be allocated to a particular object token that eventually becomes the target. To test this we used stimuli like those shown in Figure 7. As in Experiment 2, placeholders were used to allow sufficient time for the resolution of depth for all stimuli. In this case, however, many placeholders were used, which formed a "screen" of figure eights. As a result, a whole screen could be cued, and the target could emerge from any one of the placeholders.

Any reliable effect of cuing in depth found in this experiment could therefore not be due to the participants' having allocated their attention to the particular placeholder that eventually became the target. At the time of the cue, the participants had no way of knowing which placeholder might become the target.

Method

Participants
Twenty-one new participants from the same pool as that in Experiments 1 and 2 were tested.

Task
The task was the same as that in Experiment 2.

Apparatus and Stimuli
The stimuli were two overlapping 4 × 2 screens constructed of placeholders identical to those in Experiment 2 (see Figure 7). Each screen subtended 5.3° × 7.7° and was presented in the center of the computer screen. One screen was colored light blue (IBM standard palette designation 9); the other was colored light orange (IBM standard palette designation 12). The location in depth of each screen was counterbalanced within a block. The screen appearing at the back depth plane was displaced laterally and vertically by 0.2°. Each placeholder making up the screens subtended 1.4° × 3.9°. Each target letter subtended 1.4° × 3.9° and was superimposed on a placeholder making up one of the two screens. The target letters were drawn in black (IBM standard palette designation 0), the color of the background, to encourage the participants to attend to each screen.

The cues consisted of one of the two screens changing to bright white (IBM standard palette designation 15) for 50 msec.

Design
A single factor (validity: valid, invalid) within-subjects design was used. In each block, 80% of the trials were valid trials in which the target appeared on the screen that was cued, and 20% were invalid trials in which the target appeared on the screen that was not cued. The order of these trials was randomly determined for each block.

Each block consisted of 80 trials from which data were collected; the participants completed four blocks. This resulted in 256 and 64 observations in the valid and invalid conditions, respectively.

Finally, each block began with three practice trials that were randomly selected from the possible set of trial types (i.e., valid or invalid) for that block. As before, if the participants responded incorrectly to any trial, responded in less than 150 msec, or did not respond within 1,500 msec, the trial was scored as an error. Each error trial was followed by a "recovery" trial that was chosen randomly from the set of possible trial types. The data from practice and recovery trials were not included in the analyses.

Procedure
Screening procedure. This procedure was the same as that used in the previous experiments with one exception: The displays used in the screening task matched the displays used in the experiment; they were made up of two 4 × 2 screens of figure-eight placeholders. The target dot never occluded any of the screen lines when it

Table 2
Error Rates (in Percentages) for Experiment 2

Display Type	Display Length	Validity	
		Valid	Invalid
2-D	Short	3.37	4.93
	Long	2.14	20.39
3-D	Short	6.58	6.91
	Long	5.84	19.08
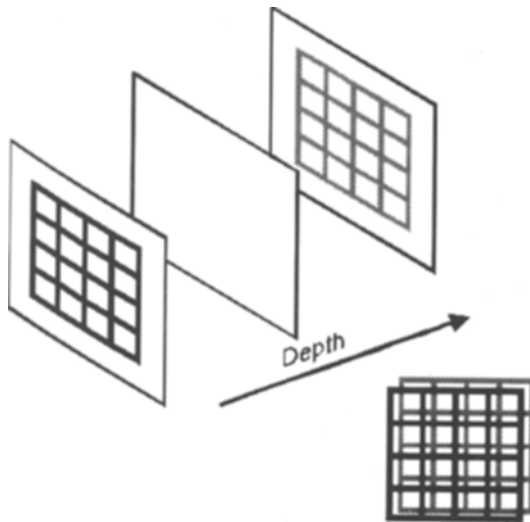
Figure 7. Displays from Experiment 3. Each screen appeared at a different depth. The lower right inset of the figure depicts how the displays would look if viewed from straight on. The cells of each screen were open in this and the following experiments. This enabled the placeholders from the back screen to be seen through the cells of the front screen.

appeared in the front depth plane, nor was it ever occluded by the screen lines when it appeared in the back depth plane.

Experimental procedure. The experiment proper began with a set of written instructions that described the task and told the participants that the target would appear on the cued screen on most of the trials. These instructions emphasized the importance of trying to use the cue to allocate attention to the cued screen.

After reading the instructions, the participants completed two 20-trial practice blocks. After practice, the participants completed four blocks of 80 trials each. They were instructed to respond as quickly as possible while maintaining at least 95% correct accuracy. If the participants' mean accuracy for a given block dropped below this criterion, the following message appeared on the screen: "You are making too many errors. Slow down and increase accuracy." A break was imposed midway through each block and was followed by three practice trials that were randomly selected from all possible trial types.

The series of events for a given trial was identical to that for the long condition in Experiment 2. The fixation display consisted of the two screens appearing for 2,500 msec. The cue then flashed for 100 msec. After the offset of the cue, the fixation display reappeared for 50 msec. A target letter then appeared on one of the screens for 50 msec. After the offset of the target, the screens remained visible either until a response was made or for 1,500 msec. An intertrial interval of 1,000 msec, during which the display was blank, followed each participant's response.

Results

Screening Procedure

Twenty of the 21 participants met criterion on the screening procedure. The mean accuracy score for these 20 participants was 100%.

Experimental Procedure

Response time. Table 3 gives the mean correct RTs for each level of validity. The overall mean RT was 566 msec. A one-way repeated measures ANOVA with validity (valid, invalid) as the factor was conducted on the mean

RTs. This revealed a significant main effect of validity [mean difference = 42 msec; $F(1,19) = 87.95$, $MS_e = 205.15, p < .001$].

Error rate. Table 3 gives the mean ERs for both levels of validity in Experiment 3. The overall mean ER was 2.69. There was no reliable main effect of validity in the ERs [$F(1,19) = 1.73$, n.s.].

Discussion

A reliable cuing effect in depth was observed. Because the participants had no way of knowing at the time of the cue which of the placeholders might turn into the target, the cuing effect cannot be attributed to the participants' having allocated their attention to a particular placeholder that eventually became the target.

Although the specific object token hypothesis can be ruled out, it is still possible that allocating attention in depth requires an object to which attention can be allocated. This follows because the screens of placeholders that were used in Experiment 3 were essentially large objects. When a screen was cued, the participants may have allocated their attention to that screen and then, consistent with previous research (e.g., Egly, Driver, & Rafal, 1994; Moore, Yantis, & Vaughan, 1998), any stimulus that appeared on that object may have been processed more efficiently than the stimuli that appeared on the uncued object (i.e., the other screen). In Experiment 4, we tested the hypothesis that in order to allocate attention in depth, one must allocate attention to one spatially contiguous object.[2]

EXPERIMENT 4

In Experiment 4, we used stimuli like those shown in Figure 8. These were like the stimuli in Experiment 3, except that gaps were introduced between the figure-eight placeholders that made up the screens. One of the two sets of placeholders could then be cued, and, as before, one of the placeholders would eventually turn into the target. Insofar as an object must be spatially contiguous, these stimuli eliminated the opportunity for the participants to allocate their attention to the cued set of placeholders as a single object. Thus, if allocating attention in depth requires that it be allocated to an object, a reliable cuing effect should not be observed.
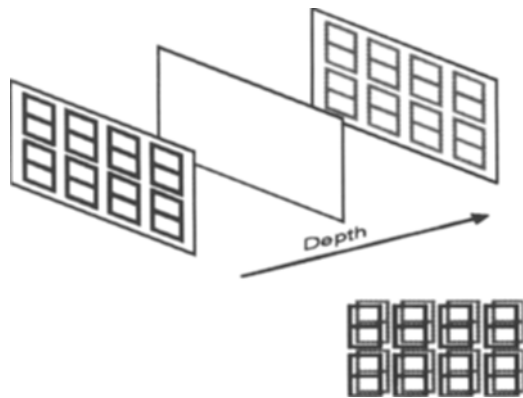
Method

Participants

Nineteen new participants from the same pool as that in Experiments 1–3 were tested in Experiment 4.

Table 3
Response Times (in Milliseconds) and Error Rates
(in Percentages) for Experiments 3, 4, and 5

| | Validity | | | |
| | Valid | | Invalid | |
Experiment	RT	%E	RT	%E
3	545	2.71	587	2.66
4	577	3.08	607	2.39
5	578	9.95	650	19.50

Figure 8. Displays from Experiment 4. Each set of placeholders appeared at a different depth plane. The lower right inset of the figure depicts how the displays would look if viewed from straight on.

Task

The task was the same as that in Experiment 3.

Apparatus and Stimuli

The stimuli were identical to those used in Experiment 3 with the following exception (see Figure 8): The placeholders making up each screen were separated laterally and vertically by 0.5°. Each screen subtended 7.1° × 8.3°.

Design

The design and procedure of Experiment 4 were identical to those of Experiment 3.

Results

Screening Procedure

Eighteen of the 19 participants met the criterion for the screening procedure. The data from 1 other participant were eliminated because the mean ER in the main task exceeded 10%. The mean accuracy score on the screening procedure for the 17 remaining participants was 99.41%, with a range of 95% to 100%.

Experimental Procedure

Response time. Table 3 gives the mean correct RTs for both levels of validity. The overall mean RT was 592 msec. A one-way repeated measures ANOVA, with validity (valid, invalid) as the variable, was conducted on the mean RTs. This revealed a significant main effect of validity [mean difference = 30 msec; $F(1,16) = 50.20$, $MS_e = 147.51$, $p < .001$].

Error rate. Table 3 gives the mean ERs for both levels of validity in Experiment 4. The overall mean ER was 2.74. The same analysis was conducted on the ERs. This analysis revealed no reliable main effect of validity [$F(1,16) = 3.67$, n.s.].

Discussion

The present results rule out the hypothesis that the allocation of attention in depth requires allocation to a single spatially contiguous object. In Experiment 4, the set of placeholders that was cued was not spatially contiguous, and as in Experiment 3, the participants had no way of knowing at the time of the cue which particular placeholder might turn into the target. Yet a reliable cuing effect was observed.

Although there is no evidence that it is necessary to allocate attention to a particular spatially contiguous object when one is attending in depth, two remaining possibilities exist, given the results of Experiment 4. First, because the screens in this experiment were constructed from arrays of regularly spaced placeholders, and because each set of placeholders was drawn in a different color, it is possible that they were perceptually grouped to form, if not a single object, at least a related group of items. If so, perhaps this grouping allowed the allocation of attention in depth.

A second possibility is that, because the placeholders belonging to a given screen were coplanar, they might have been interpreted as lying on a coherent surface. If so, it would be possible that allocating attention in depth involves attending to a given surface that is defined in depth. The importance of surface representations with regard to the allocation of attention has been a recent focus in the literature (He & Nakayama, 1995; Nakayama & He, 1996).

In Experiment 5, we sought to test these two alternative possibilities by retaining the coplanarity of the screens, while eliminating the regular spacing and color support for perceptual grouping. If either of the latter two characteristics drove the results of Experiments 3 and 4, these manipulations should eliminate the cuing effect in depth.

EXPERIMENT 5

In Experiment 5, all of the placeholders on a given trial had the same color. To further discourage the grouping of placeholders on the basis of regular spacing, we introduced random motion into the placeholder displays. Specifically, each of the placeholders moved slightly, or jiggled, randomly around its location. This motion resulted in the placeholders within a screen moving independently of one another (i.e., the opposite of common motion), which thereby strongly discouraged their being perceptually grouped on the basis of anything other than their coplanarity.

Method

Participants

Twenty-two new participants from the same pool as that in Experiments 1–4 were tested.

Task

The task was the same as that in Experiments 3 and 4.

Apparatus and Stimuli

The stimuli were the same as those in Experiment 4 with the following exceptions: First, both screens in a display were drawn in the same color. On a random half of the trials within a block, the screens were drawn in the light orange used in the previous experiments, whereas on the remaining trials, they were drawn in the light the blue

used in the previous experiments. Second, during each trial, the placeholders from each screen jiggled randomly in place. This motion, which lasted throughout each trial, was created by shifting the placeholder by 0.4° in a randomly selected direction every 50 msec. To allow adequate space for the motion of each placeholder, the overall size of each screen was increased to 8.1° × 9.3°, with the size of each placeholder remaining the same as in Experiment 4. Each frame of motion lasted approximately 50 msec, and the total time from the onset of the fixation screen to the cue was approximately 2,000 msec.

Design
The design and procedure of Experiment 5 were identical to those of Experiments 3 and 4, except that the participants completed the screening procedure after having completed the experimental task.

Results

Screening Procedure
Twenty of the 22 participants met the criterion for the screening procedure. Because the mean error rates in the main task were higher in this experiment, which was likely to have been caused by the random motion, we increased the mean ER criterion for the main task to 15%. The data from 5 participants were eliminated on the basis of this criterion. The mean accuracy score on the screening procedure for the 15 remaining participants was 98.85%, with a range of 95% to 100%.

Experimental Procedure
Response time. Table 3 gives the mean correct RTs for both levels of validity. The overall mean RT was 614 msec. A one-way repeated measures ANOVA, with validity (valid, invalid) as the variable, was conducted on the mean RTs. This revealed a significant main effect of validity [mean difference = 72 msec; $F(1,14) = 34.49$, $MS_e = 1,155.49$, $p < .001$].

Error rate. Table 3 gives the mean ERs for both levels of validity in Experiment 5. The overall mean ER was 14.73. There was a reliable main effect of validity in the ERs that was in the same direction as that for the RTs [mean difference = 9.55%; $F(1,14) = 17.88$, $MS_e = .00716$, $p < .001$].

Discussion
The results from this experiment rule out the hypothesis that the cuing in depth effects that were observed in the previous experiments were caused by the placeholders from a given screen being grouped on the basis of color or regular spacing. In this experiment, the placeholders were all of one color, and they all moved independently of one another. The one thing that the placeholders from a given screen still had in common was coplanarity. They, therefore, defined a coherent surface. Thus, these and the previous results are all consistent with the hypothesis that the allocation of attention in depth occurs on the basis of allocating attention to a surface and, therefore, requires coherent surface information to occur. In the next experiment, we tested this hypothesis by disrupting the coplanarity of the placeholders belong-

ing to each screen. If the surface hypothesis is correct, this manipulation should eliminate the cuing effect in depth.

EXPERIMENT 6

Experiment 6 was very similar to Experiment 4, except that we included a condition in which the placeholders were distributed across several different depths, which disrupted the coplanarity of the placeholders. If attention must be allocated to a given surface, reliable cuing effects in this condition should not be observed.

Method

Participants
Twenty-one new participants from the same pool as that in Experiments 1–5 were tested in Experiment 6.

Task
The task was the same as that in Experiments 3–5.

Apparatus and Stimuli
The stimuli were the same as those in Experiment 4 with the following exceptions (see Figure 9): First, both screens in a display were drawn in the same color. On a random half of the trials within a block, the screens were drawn in the light orange used in the previous experiments, whereas on the remaining trials, they were drawn in the light blue used in the previous experiments. Second, on a random half of the trials within a block, the placeholders making each "screen" were randomly located across five different depths (see Figure 9). This randomization was constrained in such a way that one placeholder from each "screen" appeared at each depth, and the remaining three placeholders were randomly located at any of the five depths. This resulted in the possibility that placeholders from the front "screen" could appear at depths behind placeholders from the back "screen" and vice versa. Neither "screen" could be interpreted as a single coherent surface. For this reason, we use quotation marks when referring to "screens" in this experiment.

Design
A 2 (surface type: random, nonrandom) × 2 (validity: valid, invalid) within-subjects design was used. Validity was defined as in the previous experiments. Surface type refers to whether the placehold-
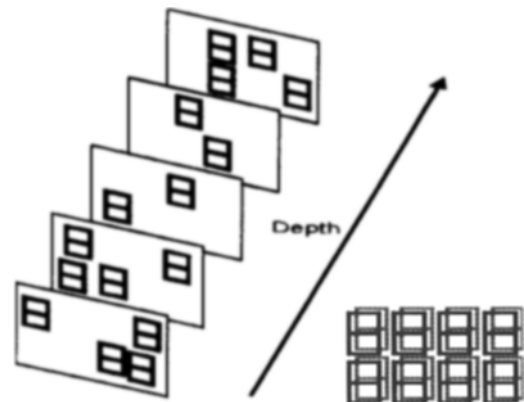
Figure 9. Displays from the random surface condition of Experiment 6. The placeholders were randomly located across five depth planes with the constraint that at least one placeholder appeared at each of the five depth planes.

ers making up a given screen were distributed across different depth planes (random) or appeared at a single depth plane (nonrandom).

Each block consisted of 80 trials from which data were collected, the participants completed four blocks. This resulted in 128 and 32 observations in the valid and invalid conditions, respectively, for each of the two surface types.

Finally, each block began with three practice trials that were randomly selected from the possible set for that block. If participants responded incorrectly to any trial, responded in less than 150 msec, or did not respond within 1,500 msec, the trial was scored as an error. Each error trial was followed by a recovery trial that was chosen randomly from the set of possible trial types. The data from practice and recovery trials were not included in the analyses.

Procedure

The screening and experimental procedures were identical to those in Experiment 5.

Results

Screening Procedure

All 21 participants met criterion for the screening procedure. The data from 2 participants, however, were removed from the analyses because their mean ERs on the main task exceeded the 10% criterion. The mean accuracy score for the remaining 19 participants was 99.21%, with a range of 95% to 100%.

Experimental Procedure

Response time. Figure 10 shows the mean correct RTs for each level of validity and surface type. The overall mean RT was 556 msec. A 2 (surface type: random, nonrandom) × 2 (validity: valid, invalid) repeated measures ANOVA conducted on mean RTs revealed a significant main effect of validity [$F(1,18) = 16.85, MS_e = 298.19, p < .001$], as well as a significant validity × surface type interaction [$F(1,18) = 20.69, MS_e = 287.13, p < .001$].

Planned comparisons between the valid and invalid conditions for both surface types confirmed that responses were faster on valid trials than on invalid trials for nonrandom displays [mean difference = 34 msec; $t(18) = 5.34, p < .001$] but not for the random displays [mean difference = −1 msec; $t(18) = 0.31$, n.s.].

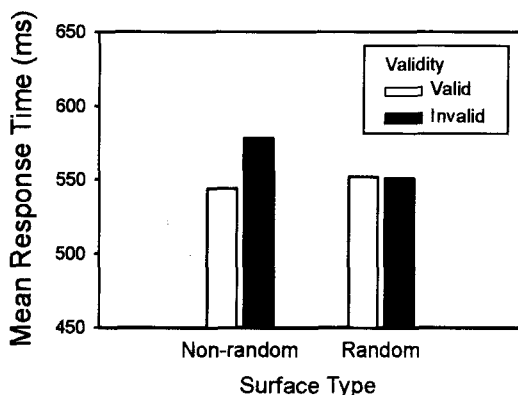**Error rate.** The same analyses were run on the ERs. The overall mean ER was 3.10. Table 4 gives the mean

Table 4
Error Rates (in Percentages) for Experiment 6

Display	Validity	
	Valid	Invalid
Nonrandom	2.38	4.44
Random	3.29	2.30

ERs for both levels of validity and surface type in Experiment 6. No significant effects that were in a different direction from those in the RTs were revealed.

Discussion

As in the previous experiments, a reliable cuing effect was observed when the cued placeholders formed a coherent surface (nonrandom condition). No reliable cuing effect occurred, however, when the cued placeholders were distributed across different depths (random condition). Because coherent surface information for the set of cued placeholders was available in the nonrandom condition, but not in the random condition, these results provide support for the hypothesis that when attention is allocated in depth, it is allocated on the basis of surfaces.[3]

Although the results of this experiment suggest that surface information is sufficient to support the allocation of attention in depth, in the final experiment we tested whether it was necessary. That is, if different organizing information, such as color, was provided in the displays, could attention be allocated to multiple items that failed to fall on a coherent surface in depth?

EXPERIMENT 7

In order to investigate whether organizing information other than surface information can support the allocation of attention to multiple items, we included a condition (like the random condition of Experiment 6) in which the surface information was disrupted. In this experiment, however, the placeholders from each "screen" were of a particular color. If a cuing effect occurred in this condition, it would indicate that, in the absence of surface information, other organizing information was sufficient to support the allocation of attention to multiple items that failed to fall on a coherent surface in depth.

Method

Participants

Seventeen new participants from the same pool as that in Experiments 1–6 were tested.

Task

The task was the same as that in Experiments 3–6.

Apparatus and Stimuli

The stimuli were the same as those in Experiment 6 with the following exception: Both screens were always drawn in a different color. On a random half of the trials within a block, the front screen was drawn in the light orange used in the previous experiments, whereas the rear screen was drawn in the light blue used in the previous experiments. The color of each screen was reversed during the

Figure 10. Mean correct response times for Experiment 6.

remaining trials. As in the previous experiment, on a random half of the trials within a block, the placeholders making each "screen" were randomly located across five different depths. This randomization was constrained in such a way that one placeholder from each "screen" appeared at each depth, and the remaining three placeholders were randomly located at any of the five depths. This resulted in the possibility that placeholders from the front "screen" could appear at depths behind placeholders from the back "screen" and vice versa. Neither "screen" could then be interpreted as a single coherent surface. For this reason, we continue to use quotation marks when referring to "screens" in this experiment.

Design

The design and procedure of Experiment 7 were identical to those of Experiment 6.

Results

Screening Procedure

All 17 participants met criterion for the screening procedure. The data from 1 participant, however, were removed from the analyses because the participant's mean ERs in the main task exceeded the 10% criterion. The mean accuracy score for the remaining 16 participants was 99%, with a range of 95% to 100%.

Experimental Procedure

Response time. Figure 11 shows the mean correct RTs for each level of validity and surface type. The overall mean RT was 533 msec. A 2 (surface type: random, nonrandom) $\times$ 2 (validity: valid, invalid) repeated measures ANOVA conducted on the mean RTs revealed only a significant main effect of validity [$F(1,15) = 18.26, MS_e = 329.28, p < .001$]. Planned comparisons between the valid and invalid conditions for both surface types confirmed that responses were faster on valid trials than on invalid trials for both nonrandom displays [mean difference = 24 msec; $t(15) = 3.79, p < .005$] and random displays [mean difference = 15 msec; $t(15) = 2.51, p < .05$].

Error rate. The same analyses were run on the ERs. The overall mean ER was 3.26. Table 5 gives the mean ERs for both levels of validity and surface type in Experiment 7. No significant effects that were in a different direction from those in the RTs were revealed.
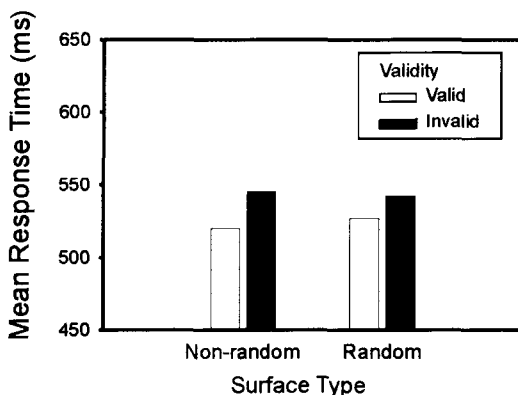
Table 5
Error Rates (in Percentages) for Experiment 7

Display	Validity	
	Valid	Invalid
Nonrandom	3.71	2.73
Random	3.08	3.52

Discussion

Unlike in Experiment 6, a reliable cuing effect was observed in this experiment, whether or not the cued placeholders had formed a coherent surface. Because surface information was available only in the nonrandom condition, these results indicate that, although surface information is sufficient, attention can be allocated to multiple items that fail to fall on a coherent surface in depth, when other organizing information, such as color, is present in the displays. This implies that coherent surface information is not the only organizing principle that determines how attention is allocated to stimuli at different depths.

GENERAL DISCUSSION

The purpose of the present study was to explore the conditions under which visual attention can be allocated in depth. The results from Experiments 1 and 2 indicated that previous failures to observe cuing effects in depth may have been attributable to there having been insufficient time for the resolution of depth in those displays. It need not be concluded, therefore, that attention is allocated on the basis of a two-dimensional representation.

In order to provide the necessary timing to support the resolution of depth, placeholders were added to the display, which raised a new hypothesis. Specifically, the appearance of object tokens (the placeholders) provided the opportunity to attend directly to the object that eventually became the target; this might have been a necessary condition for the allocation of attention in depth. Experiments 3–5 provided evidence against this hypothesis in that cuing effects were observed even when multiple placeholders that could not be grouped on the basis of spatial contiguity, color, or regular spacing were cued. Thus, there is no evidence that to attend in depth, attention must be allocated to a specific object token.

A remaining hypothesis, however, was that the coplanarity of the cued placeholders (i.e., their common depth) provided coherent surface information, and that, consistent with recent claims (e.g., Nakayama, He, & Shimojo, 1996) attention is allocated to surfaces. Experiment 6 provided support for this hypothesis; that is, when the surfaces were disrupted, the cuing effect failed to occur.

The idea that visual attention is allocated to surfaces within a scene (e.g., He & Nakayama, 1995; Nakayama & He, 1996; Nakayama, Shimojo, & Silverman, 1989; see Nakayama et al., 1996, for a review) is based historically on the central role that Gibson attributed to surfaces in providing affordances for action (e.g., Gibson, 1950). The allocation of attention can be thought of as a

Figure 11. Mean correct response times for Experiment 7.

covert action. It is a compelling view, both computationally and functionally. Moreover, when one considers past studies of attention in depth, as well as our Experiment 6, there is substantial support for the idea that surface information plays an important role in allocating attention in depth. Every study that has shown evidence of attention allocation in depth has used displays in which surface information was available. For example, Hoffman and Mueller (1997) used figure-eight placeholders like those used in the present experiments and obtained reliable cuing effects in depth. These authors, like us, discussed the availability of surface information in their displays as one possible explanation of their results. In another recent study, Theeuwes, Atchley, and Kramer (1998) used multiple visual-search displays, each presented at a different depth plane. They showed that search within a depth plane was possible (see also He & Nakayama, 1992). Because the items within a given search display were coplanar, they formed coherent surfaces. Thus, again, the data are consistent with the idea that attention was allocated on the basis of surface information that was available in the display. Finally, in another series of recent studies, Atchley and colleagues (Atchley & Kramer, 1998; Atchley, Kramer, Andersen, & Theeuwes, 1997; Atchley, Kramer, & Theeuwes, 1997) used displays in which objects (pipes in a pictured industrial site) extended across multiple depth planes. The task was to make judgments concerning multiple stimuli that appeared on these objects. An advantage was observed for conditions in which the stimuli appeared on one object compared with conditions in which stimuli appeared on two different objects. This was true despite the fact that the objects extended across multiple depth planes. These results are consistent with the idea that surface information is sufficient to allow attention to spread across depth planes.

Though all of these results suggest that attention is allocated to stimuli at different depths on the basis of surfaces, Experiment 7 demonstrated that surface information does not limit this allocation of attention. Instead, attention was successfully allocated to a subset of stimuli that shared a common color and that appeared at different depths and failed to form a coherent surface. The entire pattern of results suggests that attending in depth, like attending within a 2-D display, is influenced in general by perceptual organizing principles such as common color, common shape, and closure, rather than exclusively on the surface organization of the display.

REFERENCES

ATCHLEY, P., & KRAMER, A. F. (1998, November). *Spatial and object-based attentional selection in 3-D space.* Paper presented at the Annual Meeting of the Psychonomic Society, Dallas.

ATCHLEY, P., KRAMER, A. F., ANDERSEN, G. J., & THEEUWES, J. (1997). Spatial cuing in a stereoscopic display: Evidence for a "depth-aware" attentional focus. *Psychonomic Bulletin & Review, 4*, 524-529.

ATCHLEY, P., KRAMER, A. F., & THEEUWES, J. (1997, November). *Object-*

based attention in depth. Poster presented at the Annual Meeting of the Psychonomic Society, Philadelphia.

BOFF, K. R., & LINCOLN, J. E. (EDS.) (1988). *Engineering data compendium: Human perception and performance.* Wright-Patterson Air Force Base, OH: Armstrong Aerospace Medical Research Laboratory.

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

DOWNING, C., & PINKER, S. (1985). The spatial structure of visual attention. In M. I. Posner & O. S. M. Martin (Eds.), *Attention and performance XI* (pp. 171-187). Hillsdale, NJ: Erlbaum.

EGLY, R., DRIVER, J., & RAFAL, R. D. (1994). Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General, 123*, 161-177.

ENNS, J. T., & RENSINK, R. A. (1990). Influence of scene-based properties on visual search. *Science, 247*, 721-723.

ENNS, J. T., & RENSINK, R. A. (1991). Preattentive recovery of three-dimensional orientation from line drawings. *Psychological Review, 98*, 335-351.

ERIKSEN, C. W., & ST. JAMES, J. D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception & Psychophysics, 40*, 225-240.

GAWRYSZEWSKI, L. D. G., RIGGIO, L., RIZZOLATTI, G., & UMILTÀ, C. (1987). Movements of attention in three spatial dimensions and the meaning of "neutral" cues. *Neuropsychologia, 25*, 19-29.

GHIRARDELLI, T. G., & FOLK, C. L. (1996). Spatial cuing in a stereoscopic display: Evidence for a "depth-blind" attentional spotlight. *Psychonomic Bulletin & Review, 3*, 81-86.

GIBSON, J. J. (1950). *The perception of the visual world.* Boston: Houghton Mifflin.

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

HE, Z. J., & NAKAYAMA, K. (1995). Visual attention to surfaces in three-dimensional space. *Proceedings of the National Academy of Sciences, 92*, 11155-11159.

HOFFMAN, J. E., & MUELLER, S. (1997). *Visuo-spatial attention in three-dimensional space.* Manuscript submitted for publication.

HUMPHREYS, G. W., KEULERS, N., & DONNELLY, N. (1994). Parallel visual coding in three dimensions. *Perception, 23*, 453-470.

JONIDES, J., & MACK, R. (1984). On the cost and benefit of cost and benefit. *Psychological Bulletin, 96*, 29-44.

KLEFFNER, D. A., & RAMACHANDRAN, V. S. (1992). On the perception of shape from shading. *Perception & Psychophysics, 52*, 18-36.

LABERGE, D., & BROWN, V. (1989). Theory of attentional operations in shape identification. *Psychological Review, 96*, 101-124.

MOORE, C. M., YANTIS, S., & VAUGHAN, B. (1998). Object-based visual selection: Evidence from perceptual completion. *Journal of Experimental Psychology: Human Perception & Performance, ??*, ???-???.

NAKAYAMA, K., & HE, Z. J. (1996). Attention to surfaces: Beyond a Cartesian understanding of focal attention. In T. V. Papathomas, C. Chubb, A. Gorea, & E. Kowler (Eds.), *Early vision and beyond* (pp. 181-188). Cambridge, MA: MIT Press.

NAKAYAMA, K., HE, Z. J., & SHIMOJO, S. (1996). Visual surface representation: A critical link between lower-level and higher-level vision. In S. M. Kosslyn & D. N. Osherson (Eds.), *Visual cognition: An invitation to cognitive science* (Vol. 2, pp. 1-70). Cambridge, MA: MIT Press.

NAKAYAMA, K., & MACKEBEN, M. (1989). Sustained and transient components of focal visual attention. *Vision Research, 29*, 1631-1647.

NAKAYAMA, K., SHIMOJO, S., & SILVERMAN, G. H. (1989). Stereoscopic depth: Its relation to image segmentation, grouping and the recognition of occluded objects. *Perception, 19*, 55-68.

POSNER, M. I., NISSEN, M. J., & OGDEN, W. C. (1978). Attended and unattended processing modes: The role of set for spatial location. In H. J. Pick & I. J. Saltzman (Eds.), *Modes of perceiving and processing information* (pp. 137-157). Hillsdale, NJ: Erlbaum.

POSNER, M. I., SNYDER, C. R. R., & DAVIDSON, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General, 109*, 160-174.

RENSINK, R. A., & ENNS, J. T. (1998). Early completion of occluded objects. *Vision Research, 38*, 2489-2505.

THEEUWES, J., ATCHLEY, P., & KRAMER, A. F. (1998). Attentional control within three-dimensional space. *Journal of Experimental Psychology: Human Perception & Performance*, **24**, 1476-1485.

NOTES

1. Because the design of Experiment 1 was based on the Ghirardelli and Folk (1996) experiment, we included a neutral condition. When the data were analyzed in terms of costs and benefits, however, a complicated picture emerged. In the 2-D condition, valid trials were faster than neutral trials [mean difference = 11 msec, $t(14) = 2.28, p < .025$], but there was no reliable cost on invalid trials [mean difference = 14 msec, $t(14) = 1.63$, n.s.]. In the 3-D condition, RTs on *both* valid and invalid trials were slower than on neutral trials [mean valid difference = 15 msec, $t(14) = 3.90, p < .001$; mean invalid difference = 14 msec, $t(14) = 2.55, p < .025$]. Because of the notable difficulties in interpreting neutral trials (Jonides & Mack, 1984), we do not make much of this pattern and do not include the neutral condition in the other experiments.

2. There is another possibility that we address in the other experiments. Like Ghirardelli and Folk (1996) and Hoffman and Mueller (1997), we used stimuli that were different colors at different depths. This opens up the possibility that attention was allocated on the basis of color rather than depth. In Experiments 5–6 we show that this is not necessary for obtaining reliable cuing effects in depth.

3. There is a confound in this experiment between surface type—random versus nonrandom—and number of planes on which stimuli appeared—five versus two. Using only two depth planes for the random case, however, introduces a problem. Specifically, it is fairly easy for participants to attend to one of the two depth planes across which placeholders are distributed in the random condition. They will then enjoy the benefit of the cue on half of the valid trials, even if they are unable to attend in depth without coherent surface information. If this is what the participants were doing, we would expect the cuing effect in the random condition to be approximately half that in the nonrandom condition. That is just what we found in a pilot experiment in which we used two planes in the random condition instead of five [mean difference = 35 msec, $t(8) = 3.6, p < .01$, for nonrandom; mean difference = 13 msec, $t(8) = 2.15, p < .05$, for random]. Moreover, when the data were further analyzed, it was found that there was no cuing effect at all in the random condition until the second half of the experiment, as though the participants had learned through experience to attend to one of the two surfaces. This kind of difference across the duration of the experiment was not found in any of the other experiments. Furthermore, as will be demonstrated in Experiment 7, reliable cuing effects can be demonstrated in depth when the placeholders are located across five depth planes. This suggests that any noise that was added by presenting information across multiple depths does not disrupt the allocation of attention in depth.