Attentional Effects on Preattentive Vision: Spatial Precues Affect the Detection of Simple Features

Jan Theeuwes
TNO Human Factors Research Institute

Arthur F. Kramer and Paul Atchley
Beckman Institute
and University of Illinois at Urbana-Champaign

Most accounts of visual perception hold that the detection of primitive features occurs preattentively, in parallel across the visual field. Evidence that preattentive vision operates without attentional limitations comes from visual search tasks in which the detection of the presence or absence of a primitive feature is independent of the number of stimuli in a display. If the detection of primitive features occurs preattentively, in parallel and without capacity limitations, then it should not matter where attention is located in the visual field. The present study shows that even though the detection of a red element in an array of gray elements occurred in parallel without capacity limitations, the allocation of attention did have a large effect on search performance. If attention was directed to a particular region of the display and the target feature was presented elsewhere, response latencies increased. Results indicate that the classic view of preattentive vision requires revision.

Most current accounts of human vision suggest that there are two functionally distinct forms of visual information processing (e.g., Bergen & Julesz, 1983; Koch & Ullman, 1985; Neisser, 1967; Treisman & Gelade, 1980; Wolfe, 1994). One form is typically characterized as being preattentive, which implies an unlimited-capacity system capable of spatial parallelism in information processing. The other form is termed attentive (focal) processing because it requires the allocation of attentional resources to a limited extent of the visual field. This latter system is limited in capacity and processes information serially.

Visual attributes such as color, orientation, size, and direction of movement (see Wolfe, 1994, for a review) are detected preattentively through early spatially parallel and automatic coding in specialized feature modules (Treisman & Gelade, 1980). During search for an element defined by a single distinctive feature, a response can be made on the basis of the unique activity the element generates in the relevant feature map. If activity is detected in the relevant map, the target must be present; if not, a negative response is made (e.g., Treisman, 1988). Because the detection of these distinctive features can be based on activity registered in feature maps, it is commonly assumed that no attention is necessary for the perception of such features. Evidence that preattentive vision operates without attentional bottlenecks comes from visual search tasks in which the time to detect the presence or absence of an element defined by a distinctive feature is independent of the number of elements in the display (Treisman & Gelade, 1980). Finding no effect of the number of elements in the display indicates that there are basically no capacity limitations in performing the task. If attention is not involved in the preattentive detection of primitive features, then it is expected that detection performance will be unaffected when attention is engaged elsewhere in the visual field. Because detection is assumed to be preattentive, that is, occurring before attention operates (Neisser, 1967), it should not matter where spatial attention is allocated in the visual field.

Manipulation of spatial attention is typically accomplished by precuing the location where the target is likely to appear. Performance on valid trials, in which the precue correctly indicates the target location, is typically better than performance on invalid trials, in which the precue indicates an incorrect location (e.g., Posner, 1978). The allocation of attention in visual space has been described in terms of a “spotlight” (Broadbent, 1982; Posner, Snyder, & Davidson, 1980) and a “zoom lens” (Eriksen & St. James, 1986; Eriksen & Yeh, 1985), metaphors suggesting that attention can be narrowly or widely focused depending on the task demands. The notion underlying these metaphors is that spatial precues in a cuing paradigm exert their influence on visual processing by directing the limited-capacity attentional system to the appropriate location or object in the visual field (e.g., Kramer & Jacobsen, 1991; LaBerge & Brown, 1989). According to these models, performance can be maximized by directing attention to the location where the target is likely to appear. Because attention is limited, the allocation of attention to a cued location is accompanied by a withdrawal of attentional resources from uncued locations. If the target happens to be presented at an uncued location,
performance suffers. By directing the limited-capacity system to an area in space where the target is located, visual analysis of that stimulus is enhanced (Hawkins et al., 1990).

However, in the detection of a highly discriminable primitive feature (i.e., a "pop-out" target), it is expected that precuing the location of the feature target will have virtually no effect on performance because the detection of this feature is assumed to occur in parallel without capacity limitations (Treisman & Gelade, 1980). In other words, when a target "pops out" from the display, it should not matter where attention is allocated in the display.

Treisman (1988) confirmed this notion when exploring the effect of spatial precues on the detection of feature and conjunction targets. In Treisman’s experiments, a flashing pointer was presented 100 ms before the display, indicating with a 75% validity the location where the feature or conjunction target would appear. The results showed that for conjunction targets there was a substantial benefit from a valid cue, whereas for feature targets the cue had little effect. Treisman concluded that the direction of spatial attention is irrelevant when a target is defined by a single discriminable feature but that it has a large effect when the target is defined by a conjunction of discriminable features (Treisman, 1988).

Similar findings were obtained by Prinzel et al. (1986), who measured the proportion of false alarms in a task in which observers had to determine whether a particular colored target letter was present (e.g., a pink X) among three other colored letters. In line with the notion that the direction of spatial attention is irrelevant for detecting highly discriminable features, Prinzel et al. reported no cue validity effect on feature errors when the target shape was present (the X between Os) and the target color was absent. This implies that when the target color was absent in the display (e.g., the color pink was not present in the display), independent of where attention was directed, observers rarely reported that the target (i.e., the pink X) was present (the overall false alarms in the condition were below 2%). Yet, when the target color was present (the color pink) and the letter shape (the X) was absent, there was a small but significant effect of cue validity on feature errors. Although this latter finding suggests that primitive feature encoding does benefit from attention, as claimed by Prinzel et al. (1986), one might argue that the shape difference defining the target was not large enough to render the target as easily discriminable. In fact, to minimize false alarms in the feature condition in which the target color was absent, observers had to resolve the actual shape of the letters. Therefore, given that the shape is not a primitive feature (i.e., a pop-out target), it is not surprising that under these conditions there was a small effect of the direction of attention. However, when the feature defining the target was easily discriminable as the feature “color” in the Prinzel et al. experiment, performance was insensitive to the direction of spatial attention. In line with the work of Treisman (1988; Treisman & Gormican, 1988), when a target pops out from the display, it is detected through early, spatially parallel, and automatic coding; therefore, search performance is independent of the direction of spatial attention.

On the other hand, several researchers have measured simple reaction times (RTs) to the onset of a stimulus in an empty field (Posner, 1980; Posner et al., 1980) and found faster RTs when the target location was validity cued. These results suggest that the allocation of attention does have an effect on the detection of an onset of a stimulus. If one considers the onset of a stimulus (i.e., an abrupt onset) in an empty field as an example of a pop-out target (see, e.g., Theeuwes, 1994), then it can be argued that the allocation of attention in visual space may play a role in the detection of primitive features. However, it has been argued that faster RTs to stimuli appearing at a cued location do not necessarily imply that attention was involved in detecting the stimulus (e.g., Shaw, 1984; Shiu & Pashler, 1994). Observers may actively adopt a more liberal decision criterion for evidence accumulated in channels corresponding to the cued location. In fact, given that Posner used simple RTs to stimulus onsets, it is quite feasible that criterion shifts occurred (e.g., Shaw, 1984). When choice RT was used, spatial precues produced little if any effect on target processing when the target appeared in an empty field (e.g., Grindley & Townsend, 1968; Posner, 1980).

In the current research, we examined whether the simple detection of a primitive feature (a red line segment among gray lines) could occur in parallel (i.e., without capacity limitations) while simultaneously being affected by the allocation of attention in the visual field. Rather than cuing the exact location of the target, we cued a particular area (the left or right side of the stimulus field) within which the target was likely to appear. In other words, unlike in other cuing experiments, in our study observers still had to search for the target among a large set of nontarget elements. This task enabled us to manipulate the display size, making it possible to determine whether search was indeed performed in parallel without capacity limitations. If the detection of a primitive feature occurs without capacity limitations on the basis of the registration of activity in feature maps, then it should not matter where attention is oriented in the visual field. If, however, the detection of a primitive feature (a pop-out target) is affected by the allocation of attention in visual space, it would suggest that even when there are basically no capacity limitations in detecting the target, the allocation of attention in space may serve other purposes than simply locally enhancing the processing capacity of the limited-capacity attentional system.

Experiment 1

Observers viewed displays consisting of two stimulus arrays that were presented to the left and right sides of fixation. Display size was varied by varying the number of line elements in each of the arrays (15 or 25 line elements in each array). Observers had to determine whether a red line was present. A gray outline box served as a cue indicating with 80% validity that the red line would be presented in the stimulus field that was encompassed by the gray outline box. In the neutral condition, there were two outline boxes, each encompassing a stimulus array; this suggested that the target was to appear with an equal probability in either the left or the right stimulus field.
indicating that the target was to appear with an equal probability in either the left or right stimulus array. Display size (30 and 50 elements) and type of cue (valid, invalid, and neutral) were varied within blocks of trials. Half the observers pressed the z key when the target was present and the slash key when it was absent. This assignment was reversed for the other half of the observers. They were told that the outline box indicated with a high probability the side on which the red target line would be presented. The fixation cross remained on throughout a trial. Observers were told to keep their eyes fixated on the fixation dot. Participants received 140 practice trials before the experimental trials. Observers received feedback about their performance (in terms of RT and error rates) after each block of 70 trials.

Results

RTs longer than 900 ms were discarded, which led to a loss of 0.21% of the trials. Mean RTs on target-present and target-absent trials were submitted to separate analyses of variance (ANOVAs) with display size (30 vs. 50) and cue validity (valid, invalid, and neutral) as factors. For target-present trials, there was a main effect only of cue validity, $F(2, 14) = 21.14, p < .001$. The absence of a display size effect ($F < 1$) indicates that search was performed in parallel, suggesting that the target element popped out from the display. As is clear from Figure 2, responses were significantly slower when the cue was invalid than when it was valid, $F(1, 7) = 58.4, p < .001$. Although there was a trend, the small difference in the RT of 8 ms between the neutral and the valid conditions failed to reach significance ($F = 2.9, p = .13$).

For target-absent responses, none of the factors reached significance. Overall, there was no difference ($F < 1$) be-

---

**Method**

**Participants.** Eight 18- to 30-year-olds participated as paid volunteers. All had self-reported normal or corrected-to-normal vision and reported having no color vision defects.

**Apparatus.** A 486 PC with a super video graphics array color monitor controlled the timing of the events, generated stimuli, and recorded RTs. The forward slash key (/) and the z key of the computer keyboard were used as response buttons. Each participant was tested in a sound-attenuated, dimly lit room, with his or her head resting on a chin rest. The monitor was located at eye level, 95 cm from the chin rest.

**Stimuli.** The visual field consisted of two stimulus arrays (4.5° × 7.5°) presented to the left and right sides of fixation (see Figure 1) at an eccentricity of 4.5° (fixation point to center of the stimulus array). Each stimulus array consisted of either 15 or 25 slightly tilted line elements (0.6° long and tilted 23° randomly to either left or right) that were presented at random locations on a 4 × 8 grid. The first display contained either 30 (15 on the left side and 15 on the right side) or 50 (25 on the left side and 25 on the right side) gray line elements. After 1,000 ms, a gray outline box (5.9° × 9.9°) serving as a cue was presented encompassing either the left or the right stimulus array. One hundred milliseconds after the onset of the outline box, on half the trials, the color of one randomly chosen line element changed from gray to an equiluminant red (5.5 cd/m²). We chose equiluminant color changes because previous research has shown that equiluminant color changes do not capture attention even when observers are set to look for them (Theeuwes, 1995). The use of equiluminant color changes ensured that the feature detection consisted of the detection of a unique color (i.e., a red element among gray elements) instead of the detection of a luminance change. The display was extinguished after another 100 ms. The total time of cue and stimulus field was 200 ms, a duration too short to make directed eye movements.

**Design and procedure.** In total, there were 280 target-present and 280 target-absent trials. On 200 target-present trials, the cue encompassing one of the arrays indicated with 80% validity that the red line would be presented within the encompassed stimulus array. On the remaining 160 neutral trials (i.e., 80 target-present and 80 target-absent trials), two outline boxes encompassed both arrays,
between target-present (mean RT = 437 ms) and target-absent (mean RT = 438 ms) responses, which provides additional evidence that search was performed in parallel (see Treisman & Sato, 1990).

ANOVA performed on error rates showed no statistically reliable effects. Overall, more errors were made when the target was present (8.2%) than when it was absent (3.5%), suggesting that observers wait for the presence of activity of the target (i.e., they wait for the element to pop out). If there is no activity, they make a negative response.

Discussion

The results indicate that even when a target popped out from the display, as evidenced by the absence of a display size effect, there was still a large effect of cuing the approximate area in which the pop-out target appeared. The task used consisted of a present-absent detection judgment and could be accomplished by merely noting that something unique was either present or not. In other words, it was not necessary to actually identify the target; any activity registered in any population of feature detectors should be enough to generate a present-absent response. The results demonstrate that the allocation of attention to an approximate area in the visual field does have a large effect on feature detection.

Our results, however, do not necessarily show that the effect of cuing spatial attention operated at an early perceptual level. It is feasible that the cue simply reduced the decision time after the preattentive parallel detection of the unique feature. The cue that was used in Experiment 1 had a validity of 80%, suggesting that observers may have been actively biased toward selecting information from the cued side (e.g., Grindley & Townsend, 1968; Shaw, 1984; Shiu & Pashler, 1994). In other words, because the cue had a validity of 80%, observers might have been able to actively disregard the possibly misleading information from the noise items at the uncued side. Note, however, that even though this hypothesis is feasible, it is somewhat unlikely given the current design because, according to this hypothesis, a spatial precue can have an effect only when the target appears in the presence of potentially confusable visual noise (e.g., Henderson, 1986). Only sources of information that potentially may contribute to the decision (e.g., Is this a target or not?) should affect performance. The absence of a display size effect in the present experiment in fact suggests that the nontarget elements are not confusable with the target and do not contribute to the decision process (i.e., otherwise there would have been a display size effect).

Experiment 2

Experiment 2 was designed to determine whether cuing would affect feature detection at an early perceptual level or whether cuing the approximate area would affect decision criteria at later stages of processing. Because Experiment 1 used a cue that was valid in 80% of the trials, it was informative, enabling observers to actively expect the target to appear within the cued area (e.g., Posner, 1978). This type of endogenous cuing allows observers to intentionally direct attention to a location or an area in visual space. Because this type of cuing allows observers to actively expect the target to appear at a cued location, it may result in the aforementioned bias toward selecting information from the cued side, thereby reducing the number of sources of information that actually contribute to the decision (e.g., Grindley & Townsend, 1968; Shaw, 1984; Shiu & Pashler, 1994).

To examine whether the cuing effect observed in Experiment 1 operated at an early perceptual level, we used an exogenous cue in Experiment 2. This type of cue directs (or captures) attention to a particular location in space automatically without any intention on the part of the observer (e.g., Jonides, 1981; Theeuwes, 1991). In Experiment 2, the cue (the gray outline box) provided no predictable information about the likely target location. Observers were told that the cue was uninformative. Because the cue was completely uninformative, it should be insensitive to expectancy matching; this suggests that there should be no bias for selecting information from the cued side. If the exogenous cue used in this study still produces typical costs and benefits for detecting the pop-out target, then we can conclude that the observed cuing effects are due to attentional effects on early perceptual processing, not to a reduced uncertainty about which sources of information are relevant for the decision.

Method

Participants. Eight 18- to 29-year-olds participated as paid volunteers. All had normal or corrected-to-normal vision and reported having no color vision defects.

Stimuli and procedure. The task was identical to that used in Experiment 1 except that the cue (the outline box) encompassing one of the arrays did not predict where the line would be presented. The cue had a validity of 50%. Observers were told that the cue was uninformative.

Results

RTs longer than 1,100 ms were discarded, which led to a loss of 0.52% of the trials. For target-present trials, there was a main effect only of cue validity, $F(2, 14) = 22.96, p < .001$. Again, the absence of a display size effect ($F < 1$) showed that search was performed in parallel. Figure 3 indicates that responses were slower when the cue was invalid then when it was valid, $F(1, 7) = 27.9, p < .001$. Also, responses in the invalid condition were slower than those in the neutral condition, $F(1, 7) = 30.1, p < .001$. The small difference of 4 ms between the neutral and the valid conditions failed to reach significance ($F = 2.7, p = .14$).

Again, as in the first experiment for target-absent responses, none of the factors reached significance, and there was no reliable difference between target-present and target-absent responses. This provides additional evidence that search was performed in parallel (see Treisman & Sato, 1990).

ANOVA performed on error rates showed no statistically reliable effects. Again, more errors were made when the target was present (8.6%) than when it was absent (3.5%), suggesting that observers waited for the red element to pop.
performed in parallel across the visual display. Even though there was a large effect of the allocation of attention in the task was accomplished by spatially parallel encoding, the absence of a display size effect indicates that search was noting that something unique was or was not present. Any activity registered in any population of feature detectors should cause a pop-out signaling the presence of the target. Detection judgment and could be accomplished by merely pointing at the target feature. Overall, RTs in Experiment 2 were somewhat slower than in Experiment 1, yet the overall pattern of results was basically the same. These findings suggest that even though in Experiment 1 the cue was informative, it still may have worked exogenously in the sense that attention was pulled automatically to the cued area. The fact that there was a clear cuing effect on feature detection even when an uninformative cue was used suggests that expectancy matching—in the sense that observers were biased to select information for the cued side—cannot account for these results.

Discussion

The results of this experiment are clear: Even when an exogenous cue was used that provided no predictive information about the likely target location, cuing the approximate area still had a large effect on the detection of a primitive feature. Overall, RTs in Experiment 2 were somewhat slower than in Experiment 1, yet the overall pattern of results was basically the same. These findings suggest that even though in Experiment 1 the cue was informative, it still may have worked exogenously in the sense that attention was pulled automatically to the cued area. The fact that there was a clear cuing effect on feature detection even when an uninformative cue was used suggests that expectancy matching—in the sense that observers were biased to select information for the cued side—cannot account for these results.

General Discussion

The task used in our studies consisted of a present-absent detection judgment and could be accomplished by merely noting that something unique was or was not present. Any activity registered in any population of feature detectors should cause a pop-out signaling the presence of the target. The absence of a display size effect indicates that search was performed in parallel across the visual display. Even though the task was accomplished by spatially parallel encoding, there was a large effect of the allocation of attention in the visual field. When attention was directed to a region in space opposite to that in which the target was located, RTs were about 30 ms slower than when it was allocated to the region that contained the target. Note that the neutral condition in which both regions were cued produced results that were similar to the valid cuing condition, suggesting that there were small, if any, costs for processing both regions at the same time, a finding that is consistent with the notion of preattentive parallel processing across the visual field. These results suggest that directing attention to a region in space may result in costs when the target does not appear within that region, whereas there are minimal benefits of focusing attention to a region when the target does appear within that region.

Because the allocation of attention in the visual field did have a large effect on feature detection, it would appear that "preattentive" coding occurs within an attended area. Attention can be divided over the visual field (as in our neutral condition), allowing the detection of a feature across the entire visual field. If attention is focused on an area in visual space, feature coding occurs only within the attended area. If the feature is not present within the attended area, in order to allow the detection of the feature, the attentional window has to be reoriented to the other region. Reorienting the attentional window takes time, resulting in an increase in the amount of time necessary to detect a target. A similar notion of "preattentive vision" within an attentional window was suggested by Treisman and Gormican (1988; Treisman & Sato, 1990) to account for grouping in conjunction search. The attentional window may be directed toward one set of distractors in order to exclude distractors from another set, allowing feature search within the selected set. For example, Theeuwes (1996) found that a conjunction of color and shape can pop out because grouping on the basis of color enabled observers to direct their attentional window to a smaller subset of elements, allowing parallel search within the smaller subset of elements. Also, Nakayama and Silverman (1986) found that when observers can direct their attention to a particular plane in depth, a target defined by a conjunction of features becomes, within that particular depth plane, a target defined by a single primitive feature (see also Theeuwes, Atchley, & Kramer, 1998).

Our findings also shed light on theories that conceive of attention as a spotlight or a zoom lens (e.g., Posner, 1980; Eriksen & St. James, 1986). Spatial precues exert their influence on visual processing by directing a limited-capacity attentional system to the appropriate area within visual space. In other words, there is supposed to be an influence of the precue because there are capacity limitations. Yet, our data show that there are basically no capacity limitations in performing this task (i.e., search is performed in parallel) and that there is a simultaneous large effect of the allocation of attention in visual space. Note, however, that in cuing experiments, there are typically both costs and benefits relative to the neutral condition (see, e.g., Posner et al., 1980). In our experiments, we found that there were large costs when the target appeared in an area where attention was not directed but small, if any, benefits when attention
was directed at the area where the target appeared. These findings suggest that the function of attention is not so much to enhance processing in one region by directing the limited-capacity attentional system but rather to inhibit processing in another region (e.g., Green, 1991). For example, on the basis of single-cell recordings in the visual cortex of monkeys, Moran and Desimone (1985) suggested that attention is better conceived of as a mask than as a beam. Alternatively, it may be argued that spatial cuing affects the priority by which information is read out from the first parallel stage (e.g., Duncan & Humphreys, 1989). Along these lines, information at the cued location will have a higher probability of access to visual short-term memory than information at the uncued location. Because information is read out in parallel, in the current experiments facilitation may have been about equal for wide (neutral) and narrower (valid trials) attentional control settings.

The present results provide support for the argument that visual operations that are thought to be preattentive can be influenced by attentional allocation strategies. Joseph, Chun, and Nakayama (1997) came to a similar conclusion in a dual-task study. They found that preattentive feature detection was impaired when observers had to perform an attentionally demanding letter identification task (a rapid serial visual presentation task) in the center of the visual field. Our findings and those of Joseph et al. seem to rule out an architecture for the visual system in which all feature differences are processed along a pathway that operates without having to pass through an attentional bottleneck (e.g., Braun & Sagi, 1990, 1991).

The present results indicate that the classic notion of preattentive processing—processing before attention operates—is incorrect. The extent to which preattentive encoding occurs does depend on the allocation of attention in visual space. The results indicate that even though there were no capacity limitations, there was a large effect of the allocation of attention. Although it has been demonstrated that several stimulus properties can be coded in parallel in the primary or striate visual cortex (e.g., Hubel & Wiesel, 1977; Thorell, De Valois, & Albrecht, 1984), our results indicate that spatial attention modulates these neuronal responses, possibly by inhibiting unattended areas in visual space (Motter, 1994; Schneider, 1995).

References


Received July 1, 1997
Revision received December 22, 1997
Accepted February 23, 1998