Psychonomic Bulletin & Review
2007, 14 (5), 934-938

The size of an attentional window modulates attentional capture by color singletons

Artem V. Belopolsky, Laura Zwaan, and Jan Theeuwes
Vrije Universiteit, Amsterdam, The Netherlands

and

Arthur F. Kramer
University of Illinois at Urbana-Champaign, Urbana, Illinois

Researchers have proposed that during visual search, a color singleton cannot capture attention in a bottom-up fashion (Jonides & Yantis, 1988). In the present study, we manipulated the size of the attentional window by asking participants to detect either a global (diffuse attention) or a local (focused attention) shape before starting the search for a nonsingleton target. We demonstrate that increasing the size of attentional window causes the observers to frequently orient to an irrelevant color singleton. We conclude that although the size of attentional window might be under a top-down control, within the attentional window, an irrelevant salient singleton can capture attention in a bottom-up fashion.

Our visual system provides us with information about what is where in the environment. We often use this information to set and achieve our goals, but since we can only process a subset of the incoming visual information, we need to select the information that is relevant. An important question is to what extent one can exert top-down or endogenous control over what visual information gets selected, and to what degree visual selection is driven by the physical properties of the stimuli.

According to one view, during initial visual processing, the control of visual selection is driven entirely by the physical properties of the stimuli, or is "bottom-up" (Theeuwes, 1991, 1992, 1994). Such a conclusion was derived on the basis of a series of studies using a visual search task in which participants had to search for a unique object (a unique shape, color, or an onset element) and respond to whether the line inside of it was horizontal or vertical. This type of search—also known as the feature search—is very efficient, and search times do not depend on the number of distractors present in the display. Researchers have suggested that the target can be detected preattentively when low-level information about the scene—such as object boundaries, areas of highest local contrast, and other salient features—is extracted in parallel across the visual field (Treisman & Gelade, 1980). In addition to the unique target object, on some trials, a salient distractor object unique in a different dimension was presented.

Although the target of the search was clearly defined, the presence of a salient distractor triggered a shift of attention to its location before attention was allocated to the target. Researchers concluded that the most salient object in the visual field would capture attention irrespective of the top-down goal. Indeed, the relative salience of the target and distractor objects and not the top-down goal was shown to be the critical factor (Theeuwes, 1992): Attention always first shifted to the location of the most salient feature. The role of cross-dimensional salience in attentional allocation is an integral part of several influential models of visual search (Cave & Wolfe, 1990; Nothdurft, 1993; Wolfe, Cave, & Franzel, 1989).

According to another view (Jonides & Yantis, 1988; Yantis & Egeth, 1999), not all salient features are equally capable of attracting attention in a bottom-up fashion. Yantis and colleagues also used a visual search task, but the target of search was a nonsingleton letter. This type of search is not efficient in revealing search times that increase with the number of elements that are present in the display. In each search display, there was always one salient element, and the question addressed was whether search would automatically start at the salient element. With N as the number of elements in the display, the salient element was the target on $1/N$ of the trials, indicating that the chance that the salient element was the target was the same as that for any other letter. Since the salient element was the target at a chance level, there was no incentive to deliberately start searching at the salient singleton. Theeuwes (1991, 1992, 1994) argued that salient elements would capture attention, suggesting that observers would always start searching at the salient element in the display. However, Jonides and Yantis (1988) showed that subjects did not start searching at the salient element in the display. When the unique element happened to be the target (e.g.,

A. V. Belopolsky, a.belopolsky@psy.vu.nl

an element with a unique color or unique luminance), the search slopes were basically the same as those for the condition in which a nonunique element was the target. Contrary to Theeuwes (1991, 1992, 1994), researchers concluded that salient static singletons are treated in the same way as other nonsalient elements in the visual field. Uniqueness in color or luminance is not sufficient to capture attention when it is irrelevant to the top-down goal (for a different conclusion, see Turatto & Galfano, 2001, using a "distance method"). Jonides and Yantis showed that only elements appearing with an abrupt onset (or new objects) have a special status in capturing attention, irrespective of the top-down settings.

Recently, Theeuwes (2004) suggested that the size of the "attentional window" of observers could be one of the factors explaining why salient color singletons fail to capture attention in some studies using a visual search task (Bacon & Egeth, 1994; Folk & Annett, 1994; Jonides & Yantis, 1988). In studies that do not find capture by a color singleton, visual search often occurs in a serial or partly serial fashion, such that the search elements are examined individually or in small clusters. According to this hypothesis, the expectation of a search task causes the attention window of the observer to be adjusted accordingly (see also Gibson & Peterson, 2001), and in the case of a serial search task, the window does not encompass the whole display. This result increases the chance that the unique element is not included in the salience computations and does not capture attention. However, when the target is a unique object—as was the case in the task used by Theeuwes (1992, 1994)—the optimal strategy is to attend to the whole display at once to find the target. As a consequence, the uniquely colored item falls inside the attentional window, is processed preattentively, and captures attention. This idea is supported by a well-known finding that when a target location is known in advance, even abrupt onset does not capture attention (Theeuwes, 1991; Yantis & Jonides, 1990).

In the present study, we used a design similar to Jonides and Yantis (1988), but we also manipulated the size of the attentional window of observers. To ensure that observers spread their attention across the whole display, they had to start searching only when all the elements constructed an upward-pointing triangle. To ensure that observers focused their attention, they had to start searching only when the fixation point was a circle. If our hypothesis about the size of attentional window is correct, the color singleton should capture attention only when attention is diffuse. Therefore, we expected—in the diffuse, but not in the focused, attention condition—the search slope to decrease when the target was uniquely colored. Alternatively, if the size of the attentional window has no effect on attentional capture, the search slopes should be similar, irrespective of whether the target has a unique color, as was the case in Jonides and Yantis.

METHOD

Participants

Fourteen volunteers from the Vrije Universiteit, Amsterdam were paid to participate in a 1-h session. Their ages varied between 19 and 28 years, with a mean age of 22. They all had normal or corrected-to-normal visual acuity and normal color vision.

Stimuli

E-Prime software was used to create and present stimuli. Display elements were the letters *E*, *H*, *A*, *C*, *O*, *F*, *P*, *N*, *U*, and *S* in Times New Roman font, size 32 pt. They were 1.3º in height, 0.8º in width, and they were positioned in a shape of an upward- or downward-pointing triangle (8º base, 8.7º height) around the fixation point (Figure 1). The vertical position of the stimuli was adjusted for the fixation point (width 0.4º, height 0.4º) to be in the perceptual center of the triangle. When display size was 3 elements, the letters were presented at the vertices of the triangle; when display size was 9 elements, two extra positions on each side of the triangle were also occupied. One letter in the display was always red, and the rest of the letters were green. The green (CIE: 0.275/0.612) and red (CIE: 0.593/0.349) colors were matched on luminance (1.5 cd/m^2). The fixation point was gray (1.5 cd/m^2) and could be a plus sign, a circle, or a square. All stimuli were presented on a black background, and the viewing distance from the screen was approximately 75 cm.

Design

Two main conditions—the diffuse attention and the focused attention condition—were varied within subjects (see Figure 1). The design was identical for each condition. Display sizes of 3 (21.7% of the trials) and 9 (78.3%) elements were used. The uniquely colored letter was the target on 1/3 of the trials for the display size of 3 elements and on 1/9 of the trials for the display size of 9 elements. In the diffuse attention condition, participants had to start searching only when the letters made up an upward-pointing triangle (go trials), and they had to withhold their responses when the letters made a downward-pointing triangle (no-go trials). In the focused attention condition, participants had to start searching only when the fixation point was a circle (go trials), and they had to withhold their responses when the fixation point was a square (no-go trials). The no-go trials occurred on 21.7% of all trials. In both conditions, the same exact displays were used, but the irrelevant dimension (circle or square fixation point for the diffuse attention condition and upward or downward pointing triangle for the focused attention

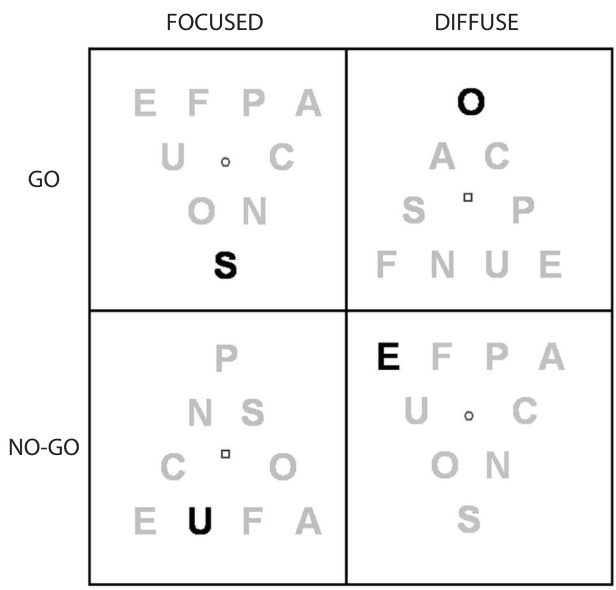

Figure 1. Examples of the displays used in the experiment. For purposes of illustration, the green letters are drawn in light gray and the red stimuli are drawn in black. In the diffuse attention condition, participants had to start searching only when the global shape was an upward-pointing triangle. In the focused attention condition, participants had to start searching only when the fixation point was a circle.

condition) was chosen randomly. Half of the participants performed the diffuse attention condition first, and the other half performed the focused attention condition first. For each condition, all types of trials were mixed. Distractor letters and the target identity (*E* or *H*) were chosen randomly for each trial. There were 414 trials in each task, out of which 324 were go trials.

Procedure

The task was to decide whether there was an *E* or an *H* present in the display and to press, respectively, the "z" or "/" key. The participants were instructed to start searching only when a go signal (which was different for each condition) was present. They were also told that the uniquely colored letter was just as likely to be the target as was any other letter in the display. The trial started with a fixation cross, which stayed on the screen for 1,000 msec. It was followed by a display of letters. If no response was detected, the search display disappeared after 2,000 msec. Participants received feedback about their accuracy and reaction time (RT) after every block of 69 trials. They were asked to respond quickly and accurately. Before the start of each condition, the participants received a sample block of practice trials.

RESULTS

Three participants were replaced because of their high error rates in the no-go trials (>30%). Trials with errors—as well as no-go trials—were excluded from the analysis. RTs that were greater than 2.5 *SD*s from the mean were discarded. This action led to the loss of less than 2% of trials.

The RTs for the diffuse and focused conditions are presented in Figure 2. A three-way repeated measures ANOVA with condition (diffuse or focused), target uniqueness (unique or nonunique), and display size (3 or 9 elements) as factors was performed on correct RTs. There was a main effect of condition [$F(1,13) = 10.74, p < .01$], with participants responding faster when their attention was diffused than when it was focused. There was also a significant effect of display size [$F(1,13) = 74.01, p < .001$], with RTs increasing when more distractors were presented. The main effect of target uniqueness was marginally significant [$F(1,13) = 4.55, p = .05$], suggesting that there was a trend to respond faster when the target had a unique color. There was also a significant interaction between condition and target uniqueness [$F(1,13) = 5.23, p < .05$], with participants responding faster when the target was unique in the diffuse attention condition, but not in the focused attention condition. A significant interaction between condition and display size [$F(1,13) = 6.58, p < .05$] indicated that increase in the search times with the display size was greater in the focused condition than in the diffuse condition. Importantly, the three-way interaction between condition, target uniqueness, and display size was significant [$F(1,13) = 6.64, p < .05$]. This interaction indicates that the difference between the search slopes in the target unique and target nonunique conditions was greater in the diffuse attention condition. No other interactions were significant.

The order in which the diffuse and focused conditions were performed did not interact significantly with other factors when it was included in the ANOVA as a between-subjects factor, suggesting that there was no carryover effect of condition.

A post hoc analysis (separate two-way repeated measures ANOVAs on the diffuse and focused conditions) showed that in the diffuse attention condition, responses were faster when the target had a unique color [$F(1,13) = 9.29, p < .05$]. There was also a significant interaction between the target uniqueness and display size [$F(1,13) = 8.58, p < .05$]. In the focused attention condition, the effect of target uniqueness and the interaction between the target uniqueness and display size were not significant [$F(1,13) = 0.98, p = .34$, and $F(1,13) = 0.02, p = .88$, respectively]. The slopes in the diffuse attention condition were 16.3 msec/item for the unique target and 26.4 msec/item for the nonunique target. In the focused attention condition, the search slopes were 27.9 msec/item for the target unique and 27.4 msec/item for the target nonunique condition. This result suggests that when the attentional window was wide, the color singleton was, on average, examined earlier in the sequence than with a narrow attentional window.

Overall, participants made 7.8% of errors in the go trials and 9.7% of errors on the no-go trials. The mean

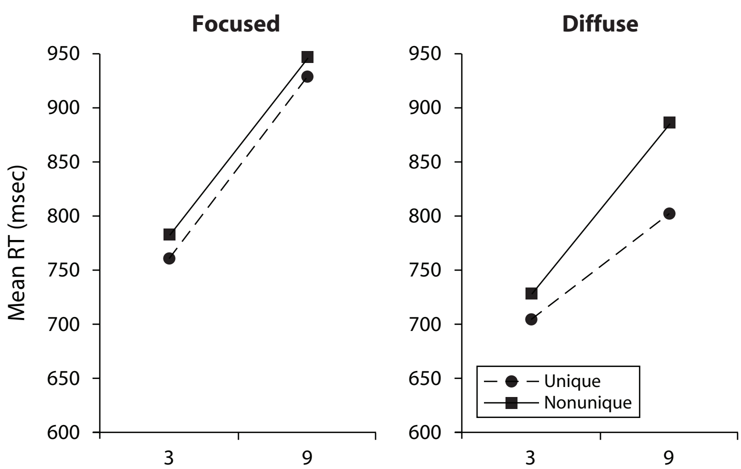

Figure 2. Mean correct reaction times (RTs) in the diffuse and focused attention conditions as a function of target uniqueness and display size.

error rates for each condition are presented in Table 1. The ANOVA on error rates showed no significant effects or interactions (condition, $F < 0.01$; interaction of condition, target uniqueness, and display size, $F < 0.1$).

Table 1
Error Rates (Percent) by Condition, Target Uniqueness, and Display Size in Go Trials

	Unique Target		Nonunique Target	
Condition	3 Elements	9 Elements	3 Elements	9 Elements
Focused	5.0	7.9	8.2	8.6
Diffuse	6.3	9.5	6.1	7.5

DISCUSSION

The present findings suggest that the size of the attentional window of the observer plays an important role in attentional capture by a task-irrelevant color singleton. In our experiment, when attention was initially focused in the center (focused attention condition), the salient color singleton was examined just as frequently as the other elements in the display. This result is similar to the "classic" finding of Jonides and Yantis (1988; also Folk & Annett, 1994; Franconeri & Simons, 2003). However, when attention was initially diffused over the global stimulus arrangement (diffuse attention condition), attention more frequently went to the location of the color distractor, which was evidenced by faster responses and a significantly reduced search slope when the colored element happened to be the target.

Since, in the diffuse attention condition, the search slope for a singleton target was 16.3 msec/item, it appears that attention was not captured by the color singleton on every trial. In our experiment, modeled after that of Jonides and Yantis (1988), manipulation of attention across the visual display resulted in a reduction of target-unique search slope from 27.9 msec/item in the focused attention condition to 16.3 msec/item in the diffuse attention condition. If attention went to the location of the color singleton on every trial, a much shallower target-unique search slope would be expected. In comparison, a search for a relevant color singleton typically produces search slopes that are less than 10 msec/item (Treisman & Gelade, 1980; but see Jonides & Yantis, 1988, Experiment 3). Although diffusing attention was not sufficient to induce capture on every trial, the significant reduction in the search slope suggests that on average, observers' attentions were captured more by the irrelevant color singleton. On some trials, observers might not have been able to maintain the diffuse attention after the presentation of a go signal, and they reverted to the "typical" serial strategy.

Our results suggest that the size of the attentional window is an important factor in determining whether an irrelevant color singleton will capture attention. When visual search is parallel (as was the case in Theeuwes, 1991, 1992, 1994), the attentional window is wide and the differences in features within and across dimensions are calculated in parallel across the whole visual field. When visual search is serial (as was the case in Jonides & Yantis, 1988), the attentional window is adjusted in anticipation of the search display (Gibson & Peterson, 2001). We demonstrated that the size of the attentional window in visual search can also be modulated by instructional manipulations. Additional support for this claim comes from the finding that even feature search for a target with a unique orientation, which presumably does not require focal attention, is affected when participants have to detect a target in a centrally presented RSVP stream (Joseph, Chun, & Nakayama, 1997). It could be that mere focusing of attention in the center prevents salience computations across the visual field and disrupts efficient detection of a salient target in the periphery.

The present findings have implications for the notion that participants can choose to adopt what is called a *feature* or a *singleton detection* search mode, as was suggested by Bacon and Egeth (1994). According to this view, when participants engage in a singleton detection mode, they choose to direct attention to the location having the largest feature contrast (highest salience). When engaged in this mode, the most salient singleton will capture attention regardless of whether it is the target or not. When participants engage in the feature search mode, they choose to direct their attention to a particular feature, and when choosing this mode, there should be no attentional capture by a salient singleton. This view is a part of the contingent capture hypothesis (Folk, Remington, & Johnston, 1992), which states that only salient features that are relevant to the attentional set of the observer (i.e., onset, color) capture attention. In the present experiment, we showed that attentional capture may not be so much under control of these presumed search modes. Instead, the occurrence of capture appears to depend on the attentional window that is applied. Indeed, it is hard to imagine that in the present experiment, the diffuse task that resulted in attentional capture was done in a search mode that Bacon and Egeth would have labeled *singleton detection*. The present findings are in line with previous claims of Theeuwes (2004) indicating that the distinction between these search modes may not be very useful. Without claiming the existence of these search modes, one can simply argue that when search becomes focused and serial (e.g., in Jonides & Yantis, 1988, and in Bacon & Egeth, 1994, Experiments 2 and 3), salient singletons may not capture attention. When attention is spread across the visual field allowing parallel search (as was the case in Theeuwes, 1992, 1994, 2004), attention is captured by salient singletons. Therefore, the present study provides additional evidence that there is no top-down control within the attended window. As was argued before, when searching in parallel for a feature singleton, irrelevant salient singletons capture attention in a purely bottom-up, exogenous way (Theeuwes, 1991, 1992, 1994; Theeuwes & Godijn, 2002).

The question that remains is, why do abrupt onsets capture attention in a serial search task, whereas other salient singletons do not capture attention under very similar conditions? To answer this question, the fundamental differences in neurophysiological organization between systems specialized for detecting transient and static signals must be considered. There are two types of ganglion cells in the retina that have been shown to differentiate between transient (Y-cells, or transient cells) and static (X-cells, or sustained cells) visual signals. Importantly, Y-cells' receptive fields have larger surrounds and are more or less distributed across the retina, whereas the X-cells' receptive fields have smaller surrounds and are mostly restricted to the fovea (Cleland, Levick, & Sanderson, 1973; Fukada & Stone, 1974). If attention acts as sensory gain control of neural responses (Hillyard, Mangun, Woldorff, & Luck, 1995), this difference in distribution of receptive fields might explain why diffusing attention across the visual field is especially helpful in detecting an irrelevant color singleton presented in the periphery, but is not as crucial for detecting an irrelevant onset. Diffusing attention might amplify the weak static signals coming from the periphery, allowing them to have a stronger representation in the saliency map.

As was mentioned earlier, the ability of salient singletons to capture attention in a bottom-up fashion is an important part of many models of visual search (Cave & Wolfe, 1990; Nothdurft, 1993; Wolfe et al., 1989). The size of an attentional window is an additional variable that needs to be considered when attentional capture by a salient singleton is investigated. We propose that salience computations are more or less restricted to the attentional window of the observer, which is under a top-down control. Within the attentional window, however, no top-down control is possible, and attention is first shifted to the location of the most salient feature.

AUTHOR NOTE

This work was based on the master's thesis of L.Z., who is now affiliated with the EMGO Institute, Department of Social Medicine, Vrije Universiteit Medical Center, Amsterdam. We thank Jay Pratt, Kyle Cave, and two anonymous reviewers for their valuable comments on an earlier version of the manuscript. We thank Bas Bastiaanse and Maarten Geers for their help in data collection. Correspondence concerning this article should be addressed to A. V. Belopolsky, Department of Cognitive Psychology, Vrije Universiteit, De Boelelaan 1111, 1081HV Amsterdam, The Netherlands (e-mail: a.belopolsky@psy.vu.nl).

REFERENCES

Bacon, W. F., & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. *Perception & Psychophysics*, **55**, 485-496.

Cave, K. R., & Wolfe, J. M. (1990). Modeling the role of parallel processing in visual search. *Cognitive Psychology*, **22**, 225-271.

Cleland, B. G., Levick, W. R., & Sanderson, K. J. (1973). Properties of sustained and transient cells in the cat retina. *Journal of Physiology*, **228**, 649-680.

Folk, C. L., & Annett, S. (1994). Do locally defined feature discontinuities capture attention? *Perception & Psychophysics*, **56**, 277-287.

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception & Performance*, **18**, 1030-1044.

Franconeri, S. L., & Simons, D. J. (2003). Moving and looming stimuli capture attention. *Perception & Psychophysics*, **65**, 999-1010.

Fukada, Y., & Stone, J. (1974). Retinal distribution and central projections of Y-, X- and W-cells of the cat's retina. *Journal of Neurophysiology*, **37**, 749-772.

Gibson, B. S., & Peterson, M. A. (2001). Inattentional blindness and attentional capture: Evidence for attention-based theories of visual salience. In C. L. Folk & B. S. Gibson (Eds.), *Attraction, distraction and action: Multiple perspectives on attentional capture* (pp. 51-76). New York: Elsevier.

Hillyard, S. A., Mangun, G. R., Woldorff, M. G., & Luck, S. J. (1995). Neural systems mediating selective attention. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 665-681). Cambridge, MA: MIT Press.

Jonides, J., & Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Perception & Psychophysics*, **43**, 346-354.

Joseph, J. S., Chun, M. M., & Nakayama, K. (1997). Attentional requirements in a "preattentive" feature search task. *Nature*, **387**, 805-807.

Nothdurft, H. C. (1993). Saliency effects across dimensions in visual search. *Vision Research*, **33**, 839-844.

Theeuwes, J. (1991). Exogenous and endogenous control of attention: The effect of visual onsets and offsets. *Perception & Psychophysics*, **49**, 83-90.

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*, **51**, 599-606.

Theeuwes, J. (1994). Stimulus-driven capture and attentional set: Selective search for color and visual abrupt onsets. *Journal of Experimental Psychology: Human Perception & Performance*, **20**, 799-806.

Theeuwes, J. (2004). Top-down search strategies cannot override attentional capture. *Psychonomic Bulletin & Review*, **11**, 65-70.

Theeuwes, J., & Godijn, R. (2002). Irrelevant singletons capture attention: Evidence from inhibition of return. *Perception & Psychophysics*, **64**, 764-770.

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

Turatto, M., & Galfano, G. (2001). Attentional capture by color without any relevant attentional set. *Perception & Psychophysics*, **63**, 286-297.

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

Yantis, S., & Egeth, H. E. (1999). On the distinction between visual salience and stimulus-driven attentional capture. *Journal of Experimental Psychology: Human Perception & Performance*, **25**, 661-676.

Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention: Voluntary versus automatic allocation. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 121-134.

(Manuscript received June 27, 2006;
revision accepted for publication September 22, 2006.)