Top-down search strategies cannot override attentional capture

JAN THEEUWES

Vrije Universiteit, Amsterdam, The Netherlands

Bacon and Egeth (1994) have claimed that color singletons do not interfere with search for a shape singleton when, instead of using a *singleton detection* mode, participants are forced to use a *feature search* mode. Bacon and Egeth induced a feature search mode by adding different shape singletons to the display so that observers could not simply respond to *uniqueness* to find the target. We did exactly the same but used larger display sizes to ensure that the target and distractor singletons remained salient. The results show that under these conditions, an irrelevant color singleton interferes with search for a shape singleton. It is argued that the notion of differential search modes may be incorrect and that the results can be explained in terms of bottom-up salience signals.

When one reads a book in the library, a particular event (e.g., someone walking in) may grab one's attention. Does this event grab attention because one is set for this event (i.e., keeping track of people walking in), or does this event grab attention automatically, even though one's top-down goal is to keep reading? This example illustrates one of the most debated issues in the study of visual attention: Can visual stimuli capture attention automatically, independently of the observer's goals, beliefs, or intentions (for recent reviews, see, e.g., Egeth & Yantis, 1997; Ruz & Lupiáñez, 2002; Theeuwes & Godijn, 2001; Yantis, 2000)? Goal-directed, or top-down, control of selection refers to the ability to select those areas, objects, feature attributes, and events needed for our current tasks. Stimulus-driven, or bottom-up, selection refers to the capacity of certain stimulus attributes to attract our attention, irrespective of our goals and beliefs.

To study the characteristics of attentional capture, *feature singletons* typically are used—stimuli that are highly salient and *pop out* from the display (such as a red element surrounded by green elements). When one is instructed to look for such a feature singleton, search time is independent of the number of elements in the display, which is generally considered to reflect *preattentive* processing (see, e.g., Treisman & Gelade, 1980). The question is whether these singletons capture attention even when one is not looking for them.

At one end of the continuum, there are accounts that assume that there is never bottom-up attentional capture. According to the contingent capture account, as proposed by Folk and colleagues (e.g., Folk, Remington, & Johnston, 1992), attentional capture by singletons is completely under top-down control. The ability of a stimulus to capture attention is contingent on whether an attentional-capturing stimulus is consistent with top-down settings, which are established *off line* on the basis of current attentional goals. According to the contingent capture model, only stimuli that match the top-down control settings will capture attention; stimuli that do not match the top-down settings will be ignored. Folk et al. used a spatial-cuing paradigm in which participants had to ignore a "cue" that appeared 150 msec prior to the presentation of the target display (see, e.g., Folk et al., 1992). Only when the search display was preceded by a to-be-ignored featural singleton (the "cue") that matched the singleton for which the observers were searching did the cue capture attention. Thus, when a red target singleton was searched for, an irrelevant red cue that preceded the search display captured attention, whereas an irrelevant onset had no effect on performance.

At the other end of the continuum, there is the notion that attentional capture is basically bottom-up and not subject to top-down control (e.g., Theeuwes 1991a, 1992, 1994b). According to this view, processing in early vision is driven exclusively by bottom-up factors, such as salience, and only later in processing may top-down factors play a role (see also Itti & Koch, 2000; Kim & Cave, 1999; Nothdurft, 1993; Theeuwes, Atchley, & Kramer, 2000). Theeuwes (1991a, 1992, 1994b) came to his conclusions on the basis of visual search tasks (known as the additional singleton task) in which two salient singletons were simultaneously present. The logic underlying the additional singleton task is simple. Participants perform a visual search task, and one item in the display is the target singleton, whereas another singleton is completely unrelated and irrelevant to the task. This condition is compared with a condition in which such an irrelevant singleton is not present. For ex-

ample, Theeuwes (1992) presented participants with circular displays consisting of colored circles or diamonds. Line segments of different orientations appeared in the circles and diamonds, and participants had to determine the orientation of the line segment appearing in the target shape. The target shape that the participants searched for was a singleton, because it was the only diamond present in the display. In the distractor condition, an irrelevant color singleton was also present in the display. Time to find the shape singleton increased when an irrelevant color singleton was present (i.e., one of the circles was red). Even though the participants had a clear top-down set to search for the shape singleton (i.e., the single green diamond), the presence of an irrelevant singleton (i.e., the single red circle) caused interference. It was shown that selectivity depended on the relative salience of the stimulus attributes: When the color singleton was made less salient (by reducing the color difference between the target and the nontarget elements), the color singleton no longer interfered with the search for the shape singleton. Theeuwes concluded that attention is captured automatically and involuntarily by the most salient singleton in the display, regardless of any top-down control settings.

To reconcile these differences in viewpoints, Bacon and Egeth (1994) conducted a crucial and now classical study. Bacon and Egeth first replicated Theeuwes's (1992) experiment, described above, in which a color singleton interfered with search for a shape singleton. In the subsequent experiment, they added different shapes (i.e., squares and triangles) to the display so that the shape singleton was no longer unique. In this condition, the color singleton did not interfere anymore. Bacon and Egeth suggested that under these conditions, observers cannot simply respond to *uniqueness* to find the target. They argued that if, when different shapes are added, participants can no longer rely on a difference signal detection (referred to as *singleton detection mode*), they switch strategies and rely on the so-called *feature search mode*. In a feature search mode, observers are able to direct their attention exclusively to the relevant feature, and irrelevant singletons no longer interfere. Note that within a block of trials, Bacon and Egeth also had displays in which only one shape singleton was present (as in Theeuwes, 1992), and even in these displays, there was no interference. These results suggest that once the feature search mode was set, it was used throughout a whole block of trials. Bacon and Egeth concluded that "goal directed selection of a specific known featural singleton identity may override stimulus-driven capture by salient singletons" (p. 493). These results suggest that when observers "choose" a feature search mode, attentional capture by irrelevant singletons can be eliminated. The notion that choosing a search strategy allows attentional control suggests that attentional capture is under top-down control.

The experiments by Bacon and Egeth (1994) and the interpretation in terms of top-down search strategies (singleton detection vs. feature search mode) are now considered the generally accepted way by which to explain differences in viewpoints on attentional capture. All recent literature reviews (see Egeth & Yantis, 1997; Ruz & Lupiáñez, 2002; Yantis, 1996, 2000) have embraced the concept of search strategies. Despite its impact on the field, it should be realized that there may be alternative and more parsimonious explanations for the absence of attentional capture when shapes are added to the display. By adding different shapes (i.e., squares and triangles), the stimulus field became less homogenous (i.e., more noisy), and this may have rendered the irrelevant color singleton less salient. Indeed, theories such as those of Duncan and Humphreys (1989) recognize that search performance depends to a large extent on how similar distractors are to each other and how dissimilar they are from the target (see also Nothdurft, 1993). Therefore, in Bacon and Egeth's displays, the irrelevant color singleton may not have captured attention not because of some search strategy chosen by the observer but because the color singleton was simply not salient enough to capture attention. As Theeuwes (1992) demonstrated, color singletons may fail to capture attention when they become less salient.

By adding different shapes and rendering the display less homogenous, not only may the distractor singleton become less salient, but also the target singleton may become less salient. The consequence is that it may become impossible to detect the target singleton by parallel preattentive search. Observers must engage in serial search to find the target singleton, and as has been pointed out, it is known that the distraction effect of a singleton is strongly attenuated during partly serial search (see Gibson & Peterson, 2001; Theeuwes, 1991a, 1992; Theeuwes & Burger, 1998). Indeed, even though, in Bacon and Egeth (1994), slopes were relatively flat (up to 11.5 msec/item), they always differed significantly from zero, suggesting that search may have been partly serial.

Experiment 1 tested the hypothesis that the absence of attentional capture in Bacon and Egeth (1994) was due to a reduction in target and distractor salience. Experiment 1 was basically the same as Bacon and Egeth's with several different shape singletons. One may expect that under these conditions, observers engage in a feature search mode. At the same time, we increased the salience of target and distractor singletons by adding more nontarget elements. Because there were many more display elements, the shape target singleton and the color distractor singleton stood out more strongly against the background (see, e.g., Green, 1991; Nothdurft, 1993). Observers searched for a shape singleton (a diamond between circles) while other unique shape singletons (a square and a triangle) were also present. Exactly as in Bacon and Egeth, by having these unique shapes added, the observers could not simply respond to *uniqueness* to find the target. Bacon and Egeth argued that under these conditions, observers will switch to a feature search mode that allows them to direct their attention exclusively to the relevant feature. If a feature search mode exists and observers were to engage in such a mode, one would expect no interference by the salient color singleton. If, however, observers do not have

these search modes at their disposal, one would expect that the color distractor singleton would capture attention, just as in Theeuwes (1991a, 1992).

EXPERIMENT 1

In Experiment 1, observers searched among either 12 or 20 display elements for a shape singleton (a diamond among circles) while other unique shape singletons (a square and a triangle) were also present.

Method

Eight participants from the Vrije Universiteit Amsterdam, ranging in age between 18 and 30 years, participated as volunteers. Display elements were equally spaced around the fixation point on an imaginary circle. For display size 20, there were 12 elements on the outer imaginary circle (5.9° radius) and 8 elements on the inner imaginary (3.7° radius) circle. For display size 12, all the elements were presented on the outer circle. The display consisted of outline circles (1.4° in diameter), one diamond (1.4° on a side), one square (1.4° on a side), and one upward-pointing triangle (2.3° on a side), each containing a line segment (0.6°). The line segment in the diamond (which was the target singleton) was vertical or horizontal, its orientation, determining the appropriate response key (vertical line segment, press the "/" key; horizontal line segment, press the "z" key). In all the other display elements, the line segments were tilted 22.5° to either side of the horizontal or vertical plane. These orientations were randomly distributed in the display. In the distractor condition, one of the green circles was replaced by a red circle. No-distractor and distractor conditions were presented in separate blocks. Display size (12 vs. 20) was varied within blocks. Figure 1 gives an example of the displays.

After presenting a fixation cross for 550 msec, the display was presented, which remained visible for 2,000 msec or until a response was made. Each participant performed 128 practice and 128 experimental trials. Half of the participants started with the no-distractor block; the other half began with a distractor block. The participants were told to respond to the orientation of the line segment inside the diamond shape.

Results

All response times (RTs) lasting longer than 1,700 msec were counted as errors, which led to a loss of less than 1% of the trials. The individual mean RTs were submitted to an analysis of variance (ANOVA) with display size (12 vs. 20 elements) and distraction (no distractor vs. distractor) as within-subjects factors. Only distraction was significant [$F(1,7) = 37.1, p < .001$]. Display size was not reliable ($F < 1$), nor did display size interact with distraction ($F < 1$). As is clear from Figure 2, the presence of a color distractor slowed search by about 65 msec (885 vs. 950 msec). The mean slopes were -0.55 msec/item for the no-distractor condition and 0.38 msec/item for the distractor condition. The absence of a display size effect suggests that search was performed in parallel. Error rates were not systematically related to any of the variables manipulated (mean error rate of 9.8%).

Discussion

Experiment 1 showed that the presence of a color singleton interferes with search for a shape singleton in conditions that are supposed to induce a feature search mode. By adding different shapes, the display became more homogenous, making both target and color distractor singletons more salient. Because the target singleton became more salient, the observers were able to find the singleton by means of parallel preattentive search, as evidenced by the absence of a display size effect.

EXPERIMENT 2

In Experiment 1, the observers were able to find the salient target singleton by means of parallel search. In order to do so, one assumes that the attentional window is set wide to encompass the whole stimulus display, and therefore, any salient element within that window, relevant

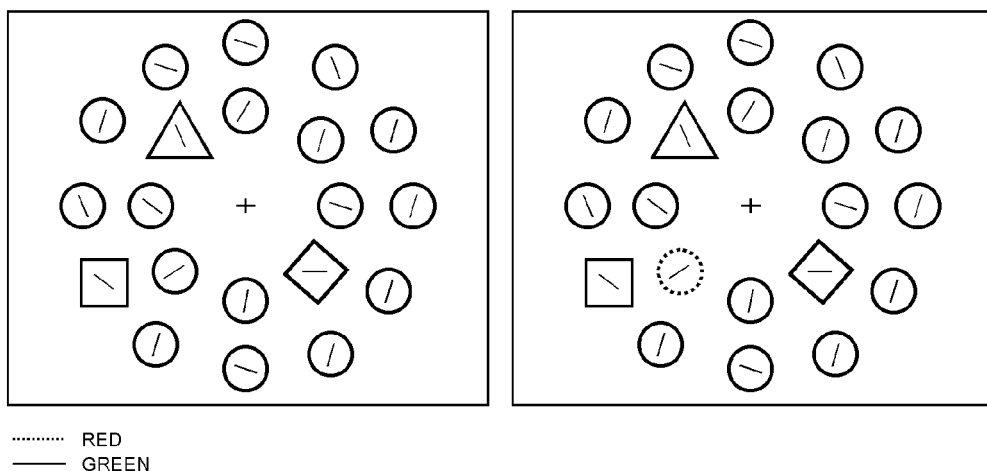
·········· RED
———— GREEN

Figure 1. Examples of the stimulus displays (with display size 20). Participants search for a shape singleton. The left panel presents the condition in which there was no distractor. The right panel presents the condition in which a color distractor singleton was present.
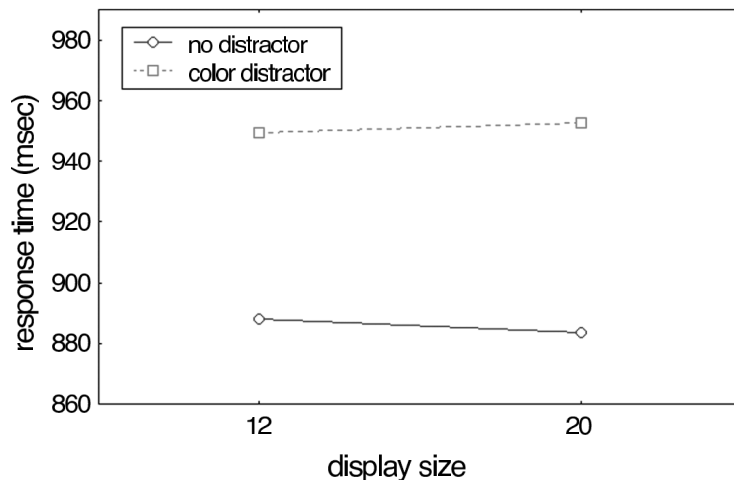
Figure 2. Experiment 1: mean response times as a function of display size for no-distractor and distractor conditions.

or irrelevant, will capture attention. If, however, the target of search is not salient enough, as may have been the case in Bacon and Egeth (1994), then to find the target, observers will reduce the size of the attentional window. As a consequence, potentially distracting singletons outside the attentional window will not compete for attention anymore and will not capture attention. The data in Bacon and Egeth in fact suggest that this may have happened. In the experiments in which there was no interference from the distractor singleton, search slopes were small, but all were significantly different from zero. In their Experiment 1 (a replication of Theeuwes, 1992), in which there was interference from the distractor singleton, search slopes did not differ from zero, suggesting parallel search across the display. It is feasible that when no interference was found, search for the target singleton became partially serial through the display (i.e., small positive search functions), and this clumpwise serial search may have attenuated the distracting effect of the irrelevant singleton.

To test the idea that reducing the salience of target and distractor singletons may induce partly serial search, thereby attenuating the effect of the color distractor, we ran an experiment that was the same as Experiment 1 (i.e., with three different shape singletons), except that we used display sizes of five and nine. These display sizes were the same as those used by Bacon and Egeth (1994).

Method

Eight participants from the University of Illinois at Urbana-Champaign, ranging in age between 18 and 26 years, participated as volunteers. The experiment was exactly the same as Experiment 1, except that we used display sizes of five and nine. Display elements were equally spaced around the fixation point on an imaginary circle with a 5.9° radius.

Results

All RTs lasting longer than 1,800 msec were counted as errors, which led to a loss of less than 1% of the trials. An ANOVA with display size (five vs. nine elements) and dis-

traction (no distractor vs. distractor) as factors showed only an effect of display size [$F(1,7) = 8.2, p < .05$]. Distraction was not reliable ($F < 1$), nor did distraction interact with display size ($F < 1$). As is clear from Figure 3, in Experiment 2, there was no distraction effect, whereas there was a reliable effect of display size. The mean slopes were 13.1 msec/item for the no-distractor condition and 11.1 msec/item for the distractor condition. Even though the display effect was relatively small, it was reliable, suggesting that search was performed serially or at least partly serially. Error rates were not systematically related to any of the variables manipulated (mean error rate of 7.4%).

Discussion

Experiment 2 showed that reducing the number of elements in a display makes both target and distractor singletons less salient. As a consequence, the target singleton cannot be found by parallel search anymore, and observers have to reduce the size of the attentional window to find the target. A more narrow focus of attention results in a search that is partly serial, as evidenced by our search slopes of 11–13 msec/item. The consequence of reduced attentional window is that a color singleton outside the attentional window may not capture attention anymore.

GENERAL DISCUSSION

The key point of these findings is that attentional capture cannot be eliminated without producing a search slope. In Experiment 1, we had parallel search (a slope of 0 msec/item) and a large interference effect of the color singleton. In Experiment 2, we had serial search (a slope of about 12 msec/item) and no interference from the color singleton. Our small positive search slopes are virtually identical to those reported by Bacon and Egeth (1994, Experiments 2 and 3), and it is likely that the absence of attentional capture in their experiments was also due to the fact that observers searched serially through the display.
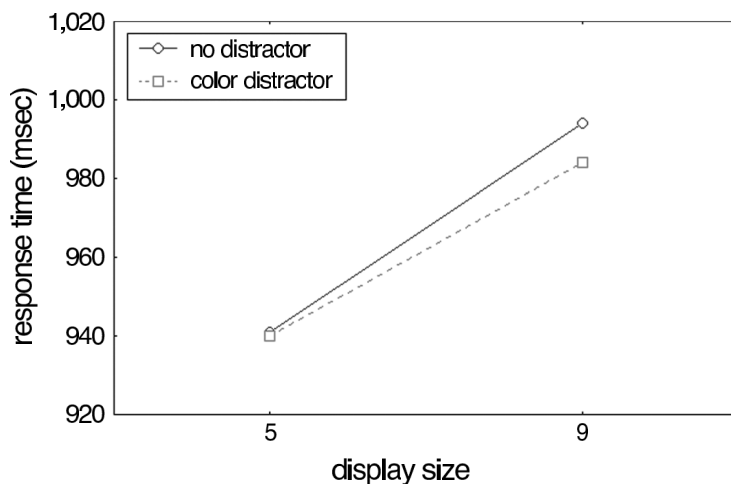
Figure 3. Experiment 2: mean response times as a function of display size for no-distractor and distractor conditions.

The present findings do not need to be explained in terms of search modes, such as feature and singleton search. We created conditions that, according to Bacon and Egeth (1994), should have induced a feature search allowing full top-down control. Even though it may appear that top-down control was possible in Experiment 2, in Experiment 1 there was a large interference effect of the color singleton. To rescue the search mode account one may argue that observers "decided" to engage in the feature search mode in Experiment 2, and for whatever reason, not to engage in this mode in Experiment 1. The reasoning may be that whenever there was interference, observers must have engaged in a singleton detection mode and whenever there was no interference, observers must have been in a feature search mode. This type of reasoning is also used to explain why there were interference effects in Theeuwes (1992) even though, in those experiments, observers had ample opportunity to engage in the feature search mode, because they searched for more than 1,800 trials for the very same target singleton (a green circle). It is clear that this reasoning is circular and does not add anything above and beyond the data: If there is interference, it is assumed that a singleton detection mode is used; if there is no interference, it is assumed a feature search mode is used.

A more parsimonious explanation is to assume that these search modes do not exist and that, in Bacon and Egeth (1994), no interference was found because, in their displays, the target and distractor singletons were not salient enough. It is not the assumed search modes that are under top-down control but the size of the attentional window. Indeed, focusing attention on a location in space prevents attentional capture by onset singletons presented at other locations (Theeuwes, 1991b; Yantis & Jonides, 1990).

Without claiming the existence of two search modes, one can simply argue that when search becomes serial (as in Jonides & Yantis, 1988, and Theeuwes, 1990) or partly serial (as in our Experiment 2 and in Bacon & Egeth, 1994, Experiments 2 and 3), distracting effects are attenuated or may even be absent. Recently, Gibson and Peterson (2001) argued along similar lines and claimed that "the expectation that a particular visual environment will require only highly focused, 'serial', forms of search in order to find the desired objects may come to preclude the processing of more global aspects of the environment that require attention to be divided, such as the relative visual salience associated with objects in the scene" (p. 55). Also, Nakayama and Joseph (1998) argued that the comparison of visual salience requires that attention be divided across the visual display so that all elements can be compared simultaneously.

So what should be concluded regarding the search modes? If one assumes that a singleton detection mode represents a condition in which the target singleton is salient enough to allow parallel search and a feature search mode represents a condition in which search is (partly) serial, there is nothing wrong with these concepts. Indeed, the attentional window may be under top-down control (Theeuwes, 1991a), and setting a smaller attentional window can attenuate distraction. However, the present study provides additional evidence that there is no top-down control within an attended window (but see Folk et al., 1992). 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, 1991a, 1992, 1994a; Theeuwes & Godijn, 2002).

REFERENCES

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

DUNCAN, J., & HUMPHREYS, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, **96**, 433-458.

EGETH, H. E., & YANTIS, S. (1997). Visual attention: Control, representation and time course. *Annual Review of Psychology*, **48**, 269-297.

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.

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.

GREEN, M. (1991) Visual search, visual streams, and visual architectures. *Perception & Psychophysics*, **50**, 388-403.

ITTI, L., & KOCH, C. (2000). Saliency based search mechanism for overt and covert shifts of visual attention. *Vision Research*, **40**, 1489-1506.

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

KIM, M. S., & CAVE, K. R. (1999). Top-down and bottom-up attentional control: On the nature of interference from a salient distractor. *Perception & Psychophysics*, **61**, 1009-1023.

NAKAYAMA, K., & JOSEPH, J. S. (1998). Attention, pattern recognition and popout in visual search. In R. Parasuraman (Ed.), *The attentive brain* (pp. 279-298). Cambridge, MA: MIT Press.

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

RUZ, M., & LUPIÁÑEZ, J. (2002). Attentional capture and exogenous orienting: Upon their automaticity and sensitivity to endogenous control. *Psicológica*, **23**, 283-309.

THEEUWES, J. (1990). Perceptual selectivity is task dependent: Evidence from selective search. *Acta Psychologica*, **74**, 81-99.

THEEUWES, J. (1991a). Cross-dimensional perceptual selectivity. *Perception & Psychophysics*, **50**, 184-193.

THEEUWES, J. (1991b). 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 shape. *Perception & Psychophysics*, **51**, 599-606.

THEEUWES, J. (1994a). Endogenous and exogenous control of visual selection. *Perception*, **23**, 429-440.

THEEUWES, J. (1994b). 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., ATCHLEY, P., & KRAMER, A. F. (2000). On the time course of top-down and bottom-up control of visual attention. In S. Monsell & J. Driver (Eds.). *Attention & performance XVIII* (pp. 105-124). Cambridge, MA: MIT Press.

THEEUWES, J., & BURGER, R. (1998). Attentional control during visual search: The effect of irrelevant singletons. *Journal of Experimental Psychology: Human Perception & Performance*, **24**, 1342-1353.

THEEUWES, J., & GODIJN, R. (2001). Attention and oculomotor capture. In C. L. Folk & B. S. Gibson (Eds.), *Attraction, distraction and action: Multiple perspectives on attentional capture* (pp. 121-149). New York: Elsevier.

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.

YANTIS, S. (1996). Attentional capture in vision. In A. F. Kramer, M. G. H. Coles, & G. D. Logan (Eds.), *Converging operations in the study of visual attention* (pp. 45-76). Washington, DC: American Psychological Association.

YANTIS, S. (2000). Goal directed and stimulus driven determinants of attentional control. In S. Monsell & J. Driver (Eds.), *Attention & performance XVIII* (pp. 73-104). Cambridge, MA: MIT Press.

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.