Psychonomic Bulletin & Review
2006, 13 (3), 524-529

Target–nontarget similarity modulates stimulus-driven control in visual search

MICHAEL J. PROULX and HOWARD E. EGETH
Johns Hopkins University, Baltimore, Maryland

The literature contains conflicting results concerning whether an irrelevant featural singleton (an item unique with respect to a feature such as color or brightness) can control attention in a stimulus-driven manner. The present study explores whether target–nontarget similarity influences stimulus-driven shifts of attention to a distractor. An experiment evaluated whether manipulating target–nontarget similarity by varying orientation would modulate distraction by an irrelevant feature (a bright singleton). We found that increasing target–nontarget similarity resulted in a decreased impact of a uniquely bright object on visual search. This method of manipulating the target–nontarget similarity independent of the salience of a distracting feature suggests that the extent to which visual attention is stimulus-driven depends on the target–nontarget similarity.

Attentional control arises as a combination of goal-directed and stimulus-driven influences (Wolfe, 1994). Goal-directed shifts of attention (e.g., to a spatial location; Yantis & Jonides, 1990) are well known. However, under what visual search conditions are stimulus-driven shifts of attention observed? Consider a task where one is searching for a green circle among green squares. The target can be defined as the unique shape, which can be detected quickly and in parallel (i.e., with response time [RT] independent of the number of elements in the display). Clearly one can adopt a search strategy of looking for a singleton—that is, a unique feature that could be detected solely on the output of a feature-contrast processing map. This search strategy has been called *singleton detection mode* (Bacon & Egeth, 1994; Pashler, 1988). If one is searching for a singleton rather than a circle, then a unique color added to the display—say, a red square—would disrupt search for the circle target. In this *additional singleton paradigm* (Theeuwes, 1991, 1992), the visual system depends on bottom-up processing of the scene to reveal locations of feature contrast, which results in stimulus-driven shifts of attention to salient locations.

Previous research has discovered that the relative discriminability of the target and distractor can influence whether a distractor captures attention. For example, Theeuwes (1992) had all observers search for a green circle. Half of the observers looked for the green circle among green squares; if one of the squares was red (a unique color), then the RT to detect the target was slowed. The other half looked for the green circle among red circles; if one of the red circles was a square (a unique shape), then the RT to detect the target was not slowed. A subsequent experiment demonstrated that this asymmetry was due to the difference in target salience between the two conditions. Thus, although an observer could look for the target-defining feature, it appears that they were searching for a singleton, and the most salient item captured attention whether it was a distractor or the target. It is important to note that the slope of the function relating RT to the number of items in the display was near zero; this flat slope indicates that visual search was efficient, with parallel processing in a large attentional window encompassing all or most of the items in the display (cf. Theeuwes, 2004).

Does a distracting singleton influence attention in search tasks where the target cannot be found efficiently? Unfortunately, there is no clear answer. Consider the case of a bright singleton. Such a stimulus has captured attention in some studies (Todd & Kramer, 1994; Yantis & Egeth, 1999, Experiment 8), but not in others (Folk & Annett, 1994; Jonides & Yantis, 1988).

Why have some studies found that a uniquely bright object influences performance but other studies have not? Several task parameters were varied among these five studies (Folk & Annett, 1994; Jonides & Yantis, 1988; Proulx & Egeth, 2006; Todd & Kramer, 1994; Yantis & Egeth, 1999) that could have been responsible for the range of results reported. One effect of the manipulation of these parameters could be that the target–nontarget similarity (independent of the salience of the distractor) in the search task might vary from study to study. This potential disparity in target–nontarget similarity can be illustrated with the studies that used letter stimuli—namely, those done by Todd and Kramer (1994), Folk and Annett (1994), and

This material is based upon work supported under a National Science Foundation Graduate Research Fellowship to M.J.P. and grants from the Binational Science Foundation and the Transportation Security Administration to H.E.E. We thank Mieke Donk, Jan Theeuwes, Steven Todd, and Cathleen Moore for insightful comments on a previous draft of the manuscript. Correspondence concerning this article should be addressed to M. J. Proulx, Institute of Experimental Psychology II, Bldg. 23.03, 00.48, Universitätstrasse 1, D-40225 Düsseldorf, Germany (e-mail: michael.proulx@uni.duesseldorf.de).

Jonides and Yantis (1988). Although Todd and Kramer used a custom sans serif font for their letter stimuli and thus were able to use all 26 letters of the alphabet for their search task, Folk and Annett (1994) and Jonides and Yantis (1988) used figure 8 style stimuli, in which only the letters that could be formed by the segments of a digital clock style 8 were used. (One should keep in mind that we are referring here to target–nontarget similarity, and not to the guidance of attention by a bright or otherwise distinctive stimulus.) The figure 8 stimuli share many components, or line segments, with one another and, depending on the letters used, might be highly similar to one another. Letters made with a sans serif font can be more distinctive in that certain features of the font may be present in only one or a few letters that contain a certain angle or line segment. It is likely that sans serif font letters differ in target–nontarget similarity from letters constructed from the line segments of a figure 8.

Interestingly, although most studies have focused on the salience of the *distractor* (normally rendered salient by being low in distractor–nondistractor similarity), the impact of the salience of the *target* on attentional prioritization has not been directly addressed for inefficient search tasks. The phrase "attentional prioritization" will be used in the Guided Search sense of the term, where priority of processing is assigned to objects as a function of decreasing activation on an activation map (see Wolfe, 1994).

The present study was designed to address the empirical gap in the literature concerning the effect of target–nontarget similarity on stimulus-driven shifts of attention. To this end, we created a manipulation of target–nontarget similarity that might capture the variability among the previously reviewed studies of attentional capture by a bright object. The present paradigm manipulated orientation to vary the similarity between the target and the nontargets (cf. Duncan & Humphreys, 1989; Wolfe, Friedman-Hill, Stewart, & O'Connell, 1992). As can be seen in panels A–C of Figure 1, target–nontarget similarity was manipulated by varying the number of degrees that the nontargets were rotated clockwise and counterclockwise from vertical (the target's orientation); the less the rotation from vertical, the higher the target–nontarget similarity. The search task was also varied in difficulty by manipulating the number of items in the display, which allowed us to assess attentional prioritization by examining the search slopes relating RT to the number of items in the display.

Bar orientation was chosen because the target–nontarget similarity for a bar could be clearly manipulated. Determining the target–nontarget similarity of a letter target and comparing that across font types is not as straightforward and would be a suitable question for future work, given the success of the present manipulation. Note that the difference in orientation between the nontargets and the target never crossed the categorical boundaries found by Wolfe et al. (1992). This helped ensure that the search task would be inefficient and, by convention, that the participants would not obviously be engaged in singleton detection mode to find the target (Turatto & Galfano, 2000; Yantis & Egeth, 1999; note, however, that singleton detection and feature search mode can both be implemented in efficient search tasks—see Leber & Egeth, 2006). If the difference in orientation had crossed the categorical boundary, then, using the logic of Turatto and Galfano and others, subjects could have used singleton detection mode to find the target and the bright singleton would have captured attention (see, e.g., Bacon & Egeth, 1994; Theeuwes, 1991). Fur-

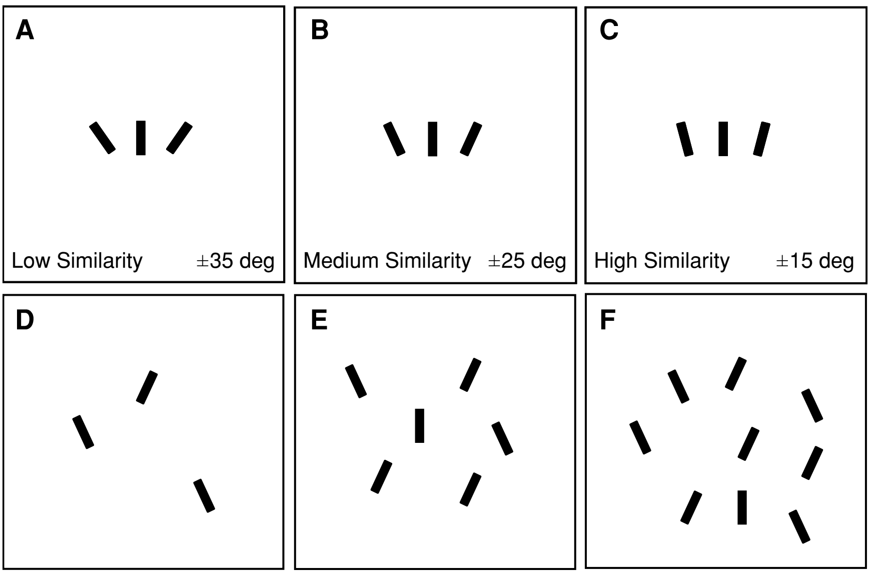

Figure 1. Example of target-present stimuli; the target was the vertical bar. Panels A–C highlight the target–nontarget similarity manipulation, where the distractors were rotated the indicated angle for each between-subjects condition. Panels D–F provide examples of the varied placement of the target and distractors for the medium target–nontarget similarity condition. Panel D shows a target-absent trial for display size 3. Panels E and F show target-present trials for display sizes 6 and 9.

thermore, we used two distinct distractors (left-tilted and right-tilted bars); this should also have prevented the use of singleton detection mode (Turatto & Galfano, 2000; Wolfe et al., 1992).

The salience of the irrelevant singleton distractor was held constant (in terms of distractor–nondistractor similarity) in the present study so that we could focus on the impact of the salience of the target (cf. van Zoest & Donk, 2004). The singleton distractor is made irrelevant by assigning it a $1/d$ probability of being the target vertical bar (where d is the number of elements in the display). Although the manipulation of the number of items in the display likely had an effect on the salience of the distractor and the target, this effect was likely to be the same for all three target–nontarget similarity conditions and should not confound the main manipulation of interest.

The impact of the manipulation of the target–nontarget similarity on stimulus-driven shifts of attention was evaluated using the *irrelevant singleton paradigm* as follows. If an object captures attention, then when that object (i.e., the irrelevant distractor) is at the target location, the search slope relating RT to the number of items in the display (the *display size*) will be near zero because the attention-capturing item will be the first attended on each trial and will therefore have full attentional priority (Yantis & Jonides, 1984). Thus we can index the attentional prioritization of the irrelevant bright singleton by observing the search slope when the singleton target is present as it moves toward (increased prioritization) or away from (decreased prioritization) zero milliseconds per item. For example, if an increase in target–nontarget similarity resulted in the bright singleton's having an increased impact on visual search, then we would expect that observers would be more likely to attend to the bright singleton, and the slope of the function when the target was a singleton would decrease as a function of increased target–nontarget similarity. However, if an increase in target–nontarget similarity resulted in the bright singleton's having a decreased impact on visual search, then we would expect that observers would be less likely to attend to the bright singleton, and the slope of the singleton target function would increase as a function of increased target–nontarget similarity. Of course, if there were to be no effect of target–nontarget similarity, then we would expect that observers would attend to the bright location just as often in the target–nontarget similarity conditions, with no change in singleton target slope.

METHOD

Participants

The participants were 45 students from Johns Hopkins University who received either course credit or payment.

Design and Procedure

The apparatus, stimuli, and procedure largely replicated those of Yantis and Egeth (1999). Three, six, or nine oriented blue bars (0.9º × 0.15º) were presented at a viewing distance of 55 cm on each trial. The background luminance was 0.1 cd/m^2, and most of the bars were 2.0 cd/m^2 (RGB 0, 0, 144); there was one bright bar on each trial, which was 9.2 cd/m^2 (RGB 0, 0, 240). The target was always a vertical bar. The distractors were bars tilted clockwise or counterclockwise a particular number of degrees, depending on the condition. Participants were randomly assigned to one of three conditions (15 per condition): (1) *low* target–nontarget similarity, with distractors that were ±35º from vertical; (2) *medium* target–nontarget similarity, ±25º; or (3) *high* target–nontarget similarity, ±15º. See Figure 1 for examples of the stimuli.

The bright bar that was presented on every trial had a $1/d$ probability of being the target, with d the number of elements in the display: 3, 6, or 9. All participants were informed that the bright singleton was no more likely to be the target than any other bar.

Each observer participated in five blocks of 108 trials per block. The observer pressed the right button on a custom box if the vertical target bar was present and the left button if it was absent. The target was present on half of the trials. Observers were instructed to respond as rapidly as possible while making fewer than 5% errors. Incorrect responses were followed by a 1-kHz feedback tone for 100 msec and a recovery trial. After each response there was a 2-sec interval before the next trial. Feedback was displayed at the end of each block, including RT and accuracy for each block. The participants began with a practice block of 20 trials, and each block began with 3 warm-up trials. Data from the practice, warm-up, incorrect, and recovery trials were not included in the RT analyses; no RTs were trimmed.

RESULTS

The mean correct RTs (see Figure 2) were submitted to an overall ANOVA with display size (3, 6, or 9) and trial type (target absent, nonsingleton target present, singleton target present) as within-subjects variables and similarity condition (low, medium, or high) as a between-subjects variable. There were significant main effects of display size [$F(2,84) = 278.2, p < .05$] and trial type [$F(2,84) = 154.0, p < .05$]. There was also a significant interaction between display size and trial type [$F(4,168) = 99.9, p < .05$]. Display size also interacted with similarity condition [$F(4,84) = 25.3, p < .05$], and trial type interacted with similarity condition as well [$F(4,84) = 6.2, p < .05$]. Finally, the three-way interaction of display size, trial type, and similarity condition was also significant [$F(8,168) = 3.8, p < .05$]. The remaining analyses focus on the target-present data because they are the most relevant for examining the impact of the bright singleton as a function of similarity condition.

First, consider the nonsingleton target data in Table 1 and Figure 2B. Note that the manipulation of target–nontarget similarity was effective: The slope of the function when the target was not the singleton increased dramatically with increasing target–nontarget similarity; this observation was confirmed with a repeated measures ANOVA for the nonsingleton target-present mean RTs, with an interaction of similarity condition and display size [$F(4,84) = 29.8, p < .05$]. Consider next the effect on attentional prioritization by examining the singleton target slopes. Figure 2C provides a depiction of the singleton target data—that is, mean RT as a function of display size when the target was present and happened to be the bright singleton for each target–nontarget similarity condition. It is readily apparent that the slopes of the functions for each target–nontarget similarity condition increased as

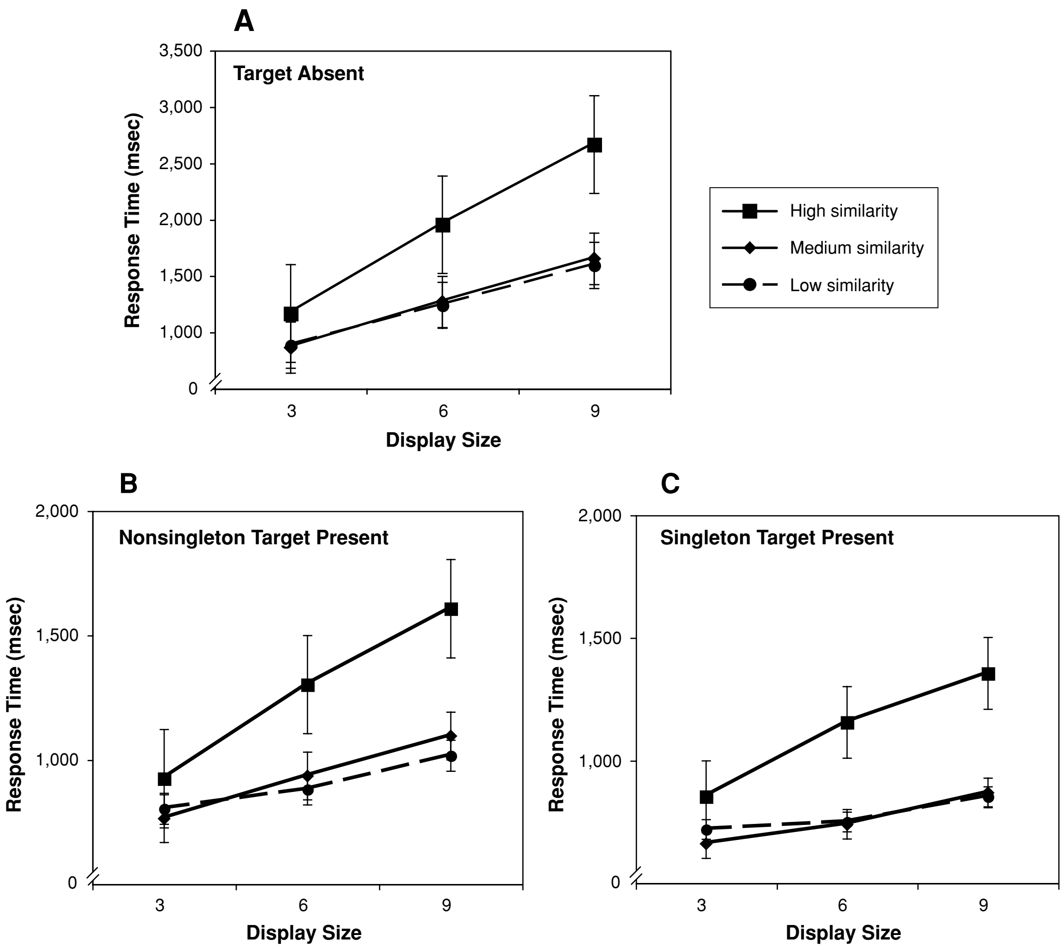

Figure 2. Plot of the mean RTs for each similarity condition as a function of display size for the (A) target-absent trials, (B) nonsingleton-target trials, and (C) singleton-target trials.

similarity increased; this was tested with a repeated measures ANOVA for the singleton target present mean RTs, and the significant interaction of similarity condition and display size confirmed this [$F(4,84) = 11.1$, $p < .05$]. Inspection of Figure 2C reveals that as target–nontarget similarity increased, the impact of the bright singleton on visual search decreased. This is most clearly seen through a comparison of the low and high target–nontarget similarity slopes. If the bright singleton's impact had increased as a function of increasing target–nontarget similarity, then the high target–nontarget similarity slope would have been less than the low target–nontarget similarity slope; in fact, the high target–nontarget similarity slope was much greater than the lower target–nontarget similarity slope (84 vs. 22 msec/item, respectively). The miss rates for the target-present and target-absent trials are shown in Table 2. The error rates generally correspond to the RTs and are therefore not indicative of a speed–accuracy trade-off.

Table 1
Search Slopes (Milliseconds per Item) by Target Type and Condition

Target–Nontarget Similarity	Target Present		Target Absent
	Target Singleton	Target Nonsingleton	
Low	22	36	118
Medium	35	55	131
High	84	114	250

As can be seen in Table 1, the target was found more efficiently when it happened to be the bright singleton than when it was not. Separate ANOVAs evaluated this observation for the mean RTs from each target–nontarget similarity condition by testing for a significant interaction between target type (singleton or nonsingleton) and display size. This interaction test revealed that there was a significant difference between the singleton target (22 msec/item) and the nonsingleton target (36 msec/item) slopes in the low target–nontarget similarity condition [$F(2,28) = 3.9$, $p < .05$]. Similarly, there was a significant difference between the singleton target (35 msec/item) and the nonsingleton target (55 msec/item) slopes in the medium

Table 2
Error Rates (Mean Percentages) by Condition and Target Type

		Display Size		
Target Type	Similarity	3	6	9
Target singleton	High	2.4	4.0	6.0
	Medium	0.4	1.8	0.0
	Low	1.1	0.0	1.3
Target nonsingleton	High	4.7	8.9	11.8
	Medium	2.6	6.4	9.0
	Low	2.7	4.2	4.3
Target absent	High	0.4	0.4	0.4
	Medium	0.1	0.1	0.4
	Low	0.5	0.7	0.6

target–nontarget similarity condition [$F(2,28) = 10.2$, $p < .05$]. However, there was no significant difference between the singleton target (84 msec/item) and the nonsingleton target (114 msec/item) slopes in the high target–nontarget similarity condition [$F(2,28) = 3.1$, $p > .05$]. The failure of the largest numerical difference in slopes (30 msec in the high target–nontarget similarity condition) to reach significance becomes more understandable when the slope ratios in the three conditions are considered (the difference in slopes divided by the nonsingleton slope). From low to high target–nontarget similarity, the slope ratios are, respectively, .39, .36, and .26. Thus, the disparity between the slopes, as a function of the increasing target–nontarget similarity, actually grows smaller; this further buttresses the argument that the most attentional prioritization of the bright singleton took place in the low target–nontarget similarity condition, and the least attentional prioritization of the bright singleton took place in the high target–nontarget similarity condition.

DISCUSSION

We began this study by asking whether target–nontarget similarity affects stimulus-driven shifts of attention to a distractor. The results imply that if target–nontarget similarity is increased, then the visual system relies less on stimulus-driven shifts of attention to find the target; this is demonstrated by the reduced prioritization evoked by a bright singleton. This suggests that as target–nontarget similarity is increased, subjects are able to restrict attention more selectively to the target-defining feature dimension rather than be driven to distraction by irrelevant feature singletons. The implication for the literature, then, is that the studies that did not find attention capture by a luminance singleton perhaps had a higher target–nontarget similarity than did the studies that found attention capture. The generalization of this account is of course limited by the fact that target salience in the present study was manipulated within the confines of one experimental paradigm and stimulus type. Thus, future work examining target salience with letters and other stimuli will provide an important test for the restricted selectivity account.

The literature contains several theoretical accounts concerning the conditions that give rise to stimulus-driven control of attention. Here, we will examine how the present findings can be explained in the context of these accounts.

Theeuwes (2004) noted that if a search task is difficult, then attentional capture by a salient distractor is less likely. The mechanism that gives rise to this effect is an attentional window. Items are analyzed in parallel within this attentional window, and items outside the window do not cause interference. This attentional window can vary in size like the zoom lens proposed by Eriksen and Yeh (1985). When a search task is easy, the window can include the entire display and process all of the items in parallel, with prioritization going to the most salient element in the display. When the search task is difficult, the window would be smaller and would thus be impervious to salient items outside of it. This model can account for the present results by assuming that as target–nontarget similarity is increased, the window gets smaller, and the salient distractor thus has less impact on search.

In addition to Theeuwes's (2004) attentional window account, there is the concept of search modes (Bacon & Egeth, 1994). Bacon and Egeth described two search modes: singleton detection mode and feature search mode. The singleton detection mode relies on local feature contrast detectors to signal discrepancies in the display that would provide stimulus-driven guidance of attention. The feature search mode relies on a top-down attentional set for a particular target-defining feature to guide attention. This search mode model can account for the present results if one assumes that singleton detection and feature search are on a continuum of reliance on bottom-up versus top-down guidance of attention in an amount proportional to the target–nontarget similarity. If the target is highly dissimilar to the nontargets, then singleton detection mode is used to rely on stimulus-driven attention; however, if the target is highly similar to the nontargets, then feature search mode is relied upon more to provide top-down attentional guidance to find the target.

The fact that these two aforementioned (and often competing) models can be extended to the present results may seem surprising, since the attentional window and search mode models were largely devised to explain attentional allocation in parallel search tasks (see Leber & Egeth, 2006). In the present study, we purposefully used an inefficient search task. Further work that extends these models to inefficient search tasks, and additional manipulations of target–nontarget similarity, may assist in further teasing apart their predictions. The attentional misguidance account of attentional capture, however, did arise from an inefficient visual search task (Todd & Kramer, 1994). Todd and Kramer found that observers were more likely to anchor search at the location of a salient feature (say, red) as its salience was increased as a function of display size (when all of the other items homogeneously shared another feature; say, green). The present results suggest that this misguidance to a distracting feature is lessened as target–nontarget similarity is increased.

A temporal view of attentional allocation provides yet another account of the data reported here. Van Zoest, Donk, and Theeuwes (2004) account for eye movement data by suggesting that stimulus-driven control of attention is present only for a short period of time after the onset of a search display as top-down control slowly takes effect. This account would predict that only fast RTs would reflect stimulus-driven control. Thus, as target–nontarget similarity is increased and RT is increased, the effects of stimulus-driven control would be expected to dissipate, as we report here. Although van Zoest et al. focused on eye movement latencies and we report manual RTs, the predictions should remain the same (participants were free to make eye movements in the present study). This account converges with the attentional window account described previously (Theeuwes, 2004) in that the high target–nontarget similarity condition likely required more fixations, thus making it less likely that the distractor would fall within the attentional window (which presumably shifts with each eye movement).

As the target–nontarget similarity is increased along the defining attribute of the task (here, orientation), the prioritization of a salient and irrelevant singleton on another dimension (here, brightness) is reduced. This method of manipulating the salience of the target independently of the salience of a distracting feature suggests that visual attention is stimulus-driven to the degree that salience is useful for finding the target, and this reliance on bottom-up processing does not appear to be restricted to the target's feature dimension. As our review of several theoretical accounts of attentional control suggests, future research will be needed to determine the specific mechanisms that relate target–nontarget similarity to attentional prioritization.

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.

Eriksen, C. W., & Yeh, Y. Y. (1985). Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception & Performance*, **11**, 583-597.

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

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

Leber, A. B., & Egeth, H. E. (2006). It's under control: Top-down search strategies can override attentional capture. *Psychonomic Bulletin & Review*, **13**, 132-138.

Pashler, H. (1988). Cross-dimensional interaction and texture segregation. *Perception & Psychophysics*, **43**, 307-318.

Proulx, M. J., & Egeth, H. E. (2006, August 8). Biased-competition and visual search: The role of luminance and size contrast. *Psychological Research* (DOI 10.1007/s00426-006-0077-z).

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

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

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

Todd, S., & Kramer, A. F. (1994). Attentional misguidance in visual search. *Perception & Psychophysics*, **56**, 198-210.

Turatto, M., & Galfano, G. (2000). Color, form and luminance capture attention in visual search. *Vision Research*, **40**, 1639-1643.

van Zoest, W., & Donk, M. (2004). Bottom-up and top-down control in visual search. *Perception*, **33**, 927-937.

van Zoest, W., Donk, M., & Theeuwes, J. (2004). The role of stimulus-driven and goal-driven control in saccadic visual selection. *Journal of Experimental Psychology: Human Perception & Performance*, **30**, 746-759.

Wolfe, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review*, **1**, 202-238.

Wolfe, J. M., Friedman-Hill, S. R., Stewart, M. I., & O'Connell, K. M. (1992). The role of categorization in visual search for orientation. *Journal of Experimental Psychology: Human Perception & Performance*, **18**, 34-49.

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. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 601-621.

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 March 23, 2005;
revision accepted for publication September 17, 2005.)