Available online at www.sciencedirect.com

SCIENCE DIRECT®

Vision Research 46 (2006) 2217–2235

Vision Research

www.elsevier.com/locate/visres

Searching for camouflaged targets: Effects of target-background similarity on visual search

Mark B. Neider, Gregory J. Zelinsky *

Department of Psychology, State University of New York at Stony Brook, Stony Brook, NY 11794-2500, USA

Received 2 July 2005; received in revised form 8 January 2006

Abstract

Do observers search for camouflaged targets by looking through the distractors or by scrutinizing the target-similar background? In four experiments observers searched for toy targets among distractors under varying set size and target-background similarity (TBS) conditions. Manual errors and RTs increased with TBS, although search slopes did not significantly differ. Eye movement analyses revealed that the majority of fixations fell on discrete distractors rather than on the target-similar background, even under high TBS conditions. These data suggest a biased search process; salient patterns segmented from a background are preferred while more target-similar unsegmented regions of the background are relatively neglected.

Keywords: Complex backgrounds; Segmentation; Object-based attention; Eye movements; Guided search

---

1. Introduction

Visual search, our ability to detect a target among distractors, is one of the most thoroughly researched of human cognitive behaviors. Much of this research has used relatively simple stimuli presented against a uniform background (see Wolfe, 1998, for a review), and from this work we have learned a great deal about the features preattentively available to the visual system (Julesz, 1981; Treisman & Gelade, 1980; Treisman & Gormican, 1988) and the processes that use these features to guide search to a target (Motter & Belky, 1998; Wolfe, 1994a; Wolfe, Cave, & Franzel, 1989; Zelinsky, 1996). However, much less is known about basic search processes in the context of more complex stimuli, stimuli requiring either a high-dimensional representation of the target's features or the enlistment of segmentation processes to separate objects from backgrounds.

Previous attempts to study visual search using more ecologically valid stimuli can be broadly classified into two categories: those that attempt to understand how real-world scenes might impose constraints on search behavior (e.g., Aks & Enns, 1996; Biederman, Glass, & Stacy, 1973; Henderson, Weeks, & Hollingworth, 1999; Neider & Zelinsky, 2006; see Henderson & Hollingworth, 1999, for a review), and those that attempt to understand how basic search processes deal with visually complex stimuli. This latter category can be further divided into studies investigating: the effects of 3D structure on search (e.g., Enns & Rensink, 1990, 1991; He & Nakayama, 1992; Kleffner & Ramachandran, 1992), the effects of 3D object rotation on search guidance (e.g., Newell, Brown, & Findlay, 2004), the use of categorical and functional features in search (e.g., Levin, 1996; Levin, Takarae, Miner, & Keil, 2001), the relationship between search and visual clutter (e.g., Bravo & Farid, 2004; Ho, Scialfa, Caird, & Graw, 2001) and the general characterization of search using real-world objects against a uniform background (e.g., Biederman, Blickle, Teitelbaum, & Klatsky, 1988), simple objects against a complex background (e.g., Wolfe,

---

* Corresponding author. Fax: +1 631 632 7876.
E-mail address: Gregory.Zelinsky@sunysb.edu (G.J. Zelinsky).

0042-6989/$ - see front matter
doi:10.1016/j.visres.2006.01.006

1994b; Wolfe, Oliva, Horowitz, Butcher, & Bompas, 2002), and images of real-world objects against complex backgrounds (e.g., Zelinsky, 1999, 2001; Zelinsky, Rao, Hayhoe, & Ballard, 1997). Other work has gone a step farther, taking search and search-related tasks out of the laboratory and into the real-world (e.g., Land & Hayhoe, 2001; Turano, Geruschat, & Baker, 2003). These studies, in addition to advancing our understanding of ecologically valid search, also inform the recent computational models of search that seek to describe basic processes in the context of fully realistic images of targets and scenes (e.g., Itti & Koch, 2000; Navalpakkam & Itti, 2005; Oliva, Torralba, Castelhano, & Henderson, 2003; Parkhurst, Law, & Niebur, 2002; Rao, Zelinsky, Hayhoe, & Ballard, 2002; Zelinsky, 2005; see Itti & Koch, 2001; for a review). The current study adds to this growing body of work by examining the effects of target-background similarity on the search for real-world objects.

There have been very few experiments directly addressing the effects of background on search. Gould and Carn (1973) monitored the eye movements of observers performing a multi-target search task with and without a complex background (a vertical grating). Although the background manipulation was not the main focus of their study, they did find that adding a background resulted in more eye movements and a constant increase in the manual search times. Gould and Carn suggested that the longer search times stemmed from visual noise generated by the background during the early stages of search processing, and that the additional eye movements were the result of a constriction in the observers' useful field of view. Because of the noise introduced by the background, observers were forced to search a smaller region of the display with each fixation, resulting in longer search times.

More recently, Wolfe et al. (2002) reported a series of experiments examining the effects of background complexity on visual search and object segmentation. They offered four hypotheses for how background complexity might degrade search performance: (1) the preattentive segmentation of search objects might take longer with a background, (2) pieces of the background might be mistakenly segmented as search objects, thereby increasing set size, (3) imperfect segmentation might slow the accumulation of information needed to identify items as targets, and (4) background noise might make it harder to select individual objects in the search display. Wolfe and colleagues had their subjects search for a T among Ls under a variety of background complexity manipulations. Their general pattern of results showed that increasing background complexity produced an increase in search times, but this increase was expressed almost entirely in the intercept of the RT × set size function, not by a change in slope. Moreover, the RT cost associated with a background was larger in the target-absent data compared to the target-present data. Because hypothesis 1 predicts equivalent background costs in the target-present and absent conditions, and hypothesis 4 predicts steeper search slopes with a background, both of these options were ruled out given the obtained data. Based on the results of a multi-target search task designed to distinguish between hypotheses 2 and 3, Wolfe and colleagues ultimately concluded that backgrounds degrade search performance by slowing the rate of information accumulation in a post-attentive stage of item identification.

The conclusion reached by Wolfe et al. (2002) was based on conditions in which targets and distractors could be segmented from a background with relative ease, but what if the background was visually more similar to the search objects—would a different data pattern emerge? To address this question, Wolfe et al. (2002) systematically varied the similarity between the background and the search objects (Experiment 6). Both the background and objects consisted of checkerboard elements. The target and distractors consisted of 2×3 configurations of elements, with the target being vertically oriented and the distractors being horizontally oriented. Similarity was manipulated by varying the element size of the background, expressed as a ratio of background element size to target element size. Contrary to their earlier pattern of results, they found that search slopes increased with the similarity between the background and the search objects, reaching approximately 80 ms/item when the background-to-object element ratio was 1:1. Following their framework of hypotheses, these authors concluded that the item-by-item selection of objects in a display becomes more difficult when these objects are highly similar to a background.

The current study picks up where Wolfe et al.'s Experiment 6 left off, exploring further the effects of target-background similarity (TBS) on search. Our work differs from the Wolfe et al. (2002) study in two key respects. First, whereas Wolfe and colleagues used Ts and Ls or simple checkerboard patterns in their experiments, we used real-world objects. Simple stimuli were used in the earlier study in order to cleanly manipulate visual similarity between the search items and the background. We overcome this problem by constructing backgrounds from an actual patch of the target object, with the size of this patch determining the level of TBS. Small image patches produce backgrounds that are relatively dissimilar to the target; large image patches produce backgrounds that essentially camouflage the target. The primary goal of this study is to better understand this effect of camouflage on search, how does search change when the target blends into a background? Second, in Wolfe et al.'s Experiment 6, targets and distractors were highly similar to each other, and therefore of comparable similarity to the background. This is not the case in true camouflage situations in which distractors are often highly visible against a background and only the target is difficult to detect (King, Stanley, & Burrows, 1984). How does search change as a function of TBS when distractor-background similarity (DBS) is relatively low and held constant? We conducted Experiment 1 to answer these questions.

2. Experiment 1

There are three ways in which TBS might affect the search RT × set size function (Fig. 1). First, it is possible that adding a camouflage background might not change search slopes at all. Although RTs might increase for reasons discussed in Wolfe et al. (2002), the slope of the RT × set size function might not vary with TBS. We would expect this pattern if observers continue to inspect distractors under camouflage conditions, and if the rejection of these distractors can be accomplished as efficiently with a background as without (efficient-rejection hypothesis; Fig. 1A). A second possibility is that search slopes will increase with TBS, as suggested by Wolfe et al. (2002; Experiment 6). Finding this pattern would mean that search is again directed to distractors, but that additional time is needed to select or reject each distractor when it appears on a target-similar background (inefficient-rejection hypothesis; Fig. 1B), possibly due to a form of flanker competition (Eriksen & Yeh, 1985). Third, slopes might flatten under camouflage conditions due to the direction of search away from the distractors and toward the background. By definition, high TBS means that the target and background will share visual features. To the extent that search is guided to target-similar features in a display (Rao et al., 2002; Wolfe, 1994a; Zelinsky, 2005), the distractors should be rendered irrelevant to the task once the level of TBS exceeds the level of target-distractor similarity (distractor-independence hypothesis; Fig. 1C). Extending this logic, given that the addition of each target-dissimilar distractor to a display will cover a portion of the target-similar background, one might even expect negative search slopes under camouflage conditions. In Experiment 1 we determine which of the above three predictions best characterize search performance under conditions of minimal (uniform background) and moderate TBS.

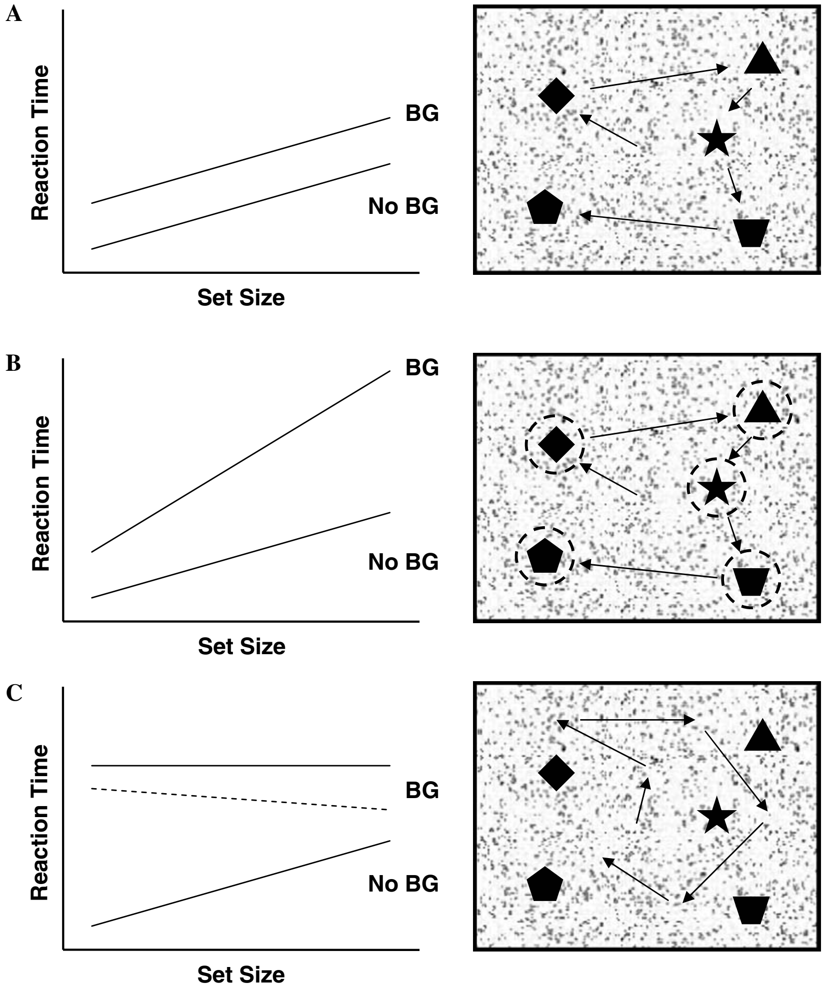

Fig. 1. Three possible relationships between target-background similarity and search slopes. The left panels show hypothetical RT × set size functions; the right panels show hypothetical allocations of attention or gaze with a background. (A) Efficient-rejection hypothesis. Observers inspect and reject distractors in the background (BG) and no-background (No BG) conditions with equal efficiency. (B) Inefficient-rejection hypothesis. Observers require more time to reject each distractor (dotted circles), resulting in steeper background search slopes. (C) Distractor-independence hypothesis. Observers look to the background rather than the distractors, resulting in flat or even negative search slopes.

2.1. Method

2.1.1. Participants

Twenty undergraduate students from Stony Brook University participated for course credit. All had normal, or corrected-to-normal vision, by self-report.

2.1.2. Apparatus and stimuli

Stimuli were displayed in color on a ViewSonic 19″ flat-screen CRT monitor by a Pentium-based computer (300 MHz) running MS-DOS (v. 6.22). Observers registered their responses by pressing one of two buttons on a custom-made button box.

Targets and distractors were selected from a set of 50 images of children's toys taken from the Hemera Photo Objects database. Examples include a rubber ducky, toy keys, and a teddy bear. Although the objects naturally varied in shape, all were scaled to just fit into a 75 × 75 pixel box, or approximately 1.5° × 1.5° of visual angle.

For each of the 50 objects, a corresponding camouflage background was constructed by taking a 20 × 20 pixel square (.4° × .4°) from the center of each object, then using this pattern to tile over an 800 × 600 pixel canvas (Fig. 2). In cases where shape irregularities made it impossible to obtain a 20 × 20 pixel square from the center of an object, the tile was taken from as near to center as possible.

Two types of search displays were used in this experiment. Camouflage-background displays consisted of search objects superimposed over a tiled background. In the case of target-present trials, the target in the display would be the object used to tile the background. Uniform-background displays consisted of search objects on a uniform dark background. For both types of displays, distractors were selected randomly and without replacement from the non-target objects in the 50 item set. To place objects in displays, each 800 × 600 pixel background was divided into a 10 × 7 grid, with objects assigned randomly to these grid locations. The center six grid locations were excluded in order to prevent observers from looking at or near an object following search display onset. Given these placement constraints, the minimum distance between the display's center and the center of the nearest object was 3.2°, the minimum center-to-center distance between objects was 1.6°. The composite search displays (objects plus background) subtended 17.6° × 12.6° when viewed at a distance of 112 cm. Fig. 3A shows a grayscale sample of a camouflage display used in Experiment 1.

2.1.3. Design and procedure

There were 600 experimental trials, evenly divided into two background conditions (camouflage or uniform), two target conditions (present or absent), and five set size conditions (9, 19, 29, 39, and 49 objects), leaving 30 trials per cell of the factorial design. All manipulations were randomly interleaved throughout the experiment.

Each trial began with the observer fixating a central cross and pressing either of the two response buttons to initiate the task. The target object for the trial was then presented centrally on a dark background for 1 s, followed by the onset of the search display. Observers were instructed to indicate the presence or absence of the target by pressing the left or right buttons, respectively, and to make their judgments as quickly as possible while maintaining accuracy. There were 40 practice trials illustrating all of the conditions, but no feedback was provided. The entire experiment lasted approximately 3 h and was completed in one session.

2.2. Results and discussion

Error rates averaged 5.7% in the target-present (TP) data and 1% in the target-absent (TA) data, with no significant differences between background conditions.

Fig. 4 shows the mean RTs for correct trials plotted as a function of background condition and set size. The clearest patterns emerging from these data are the main effects of background in the target-present, $F(1,19) = 72.18$, $p < .001$, and target-absent, $F(1,19) = 61.26$, $p < .001$, conditions, and the highly significant target × background interaction, $F(1,19) = 36.47$, $p < .001$. Observers took longer to register their judgments when there was a camouflage background, and this difference was larger in the TA data (approximately 930 ms) than in the TP data (approximately 330 ms). As for the primary question motivating

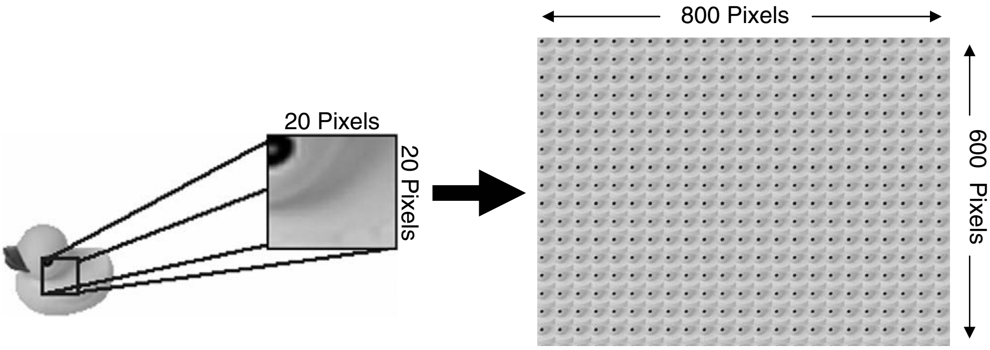

Fig. 2. Target-similar backgrounds were created by taking a square region from the center of each target object and using it to tile an 800 × 600 pixel canvas.

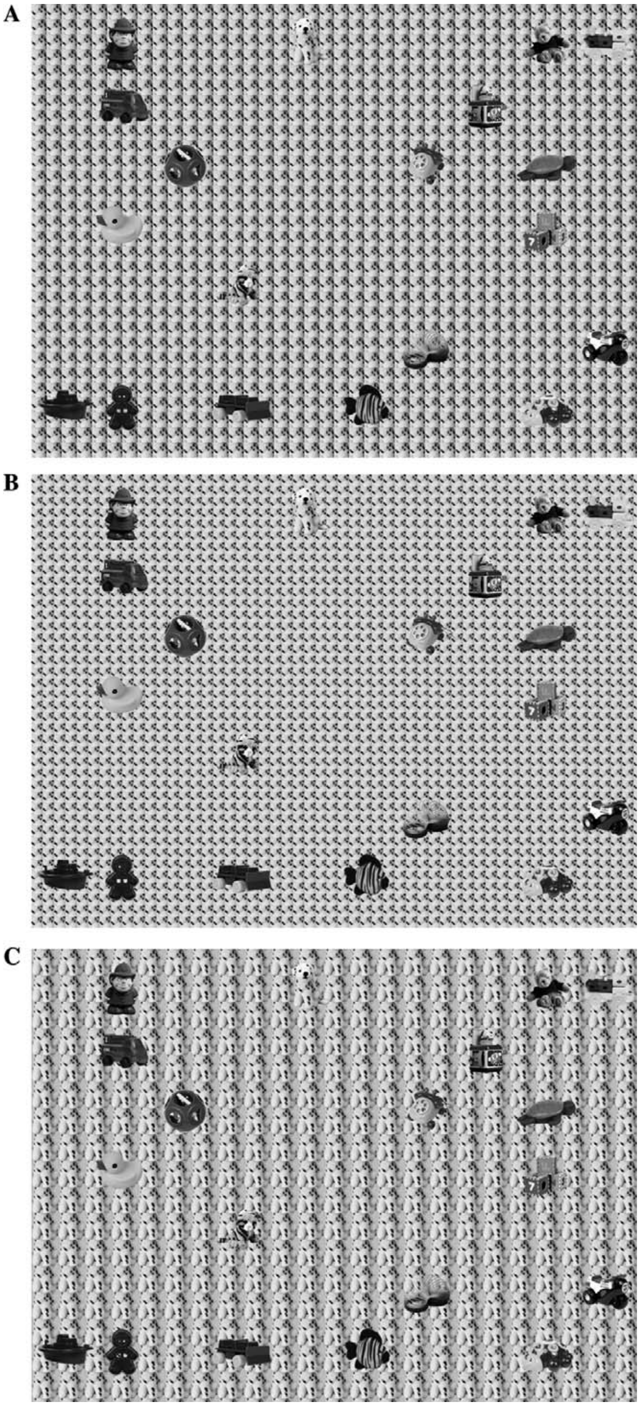

Fig. 3. A sample configuration of search items on three differently tiled camouflage backgrounds. The target in (A–C) is a Dalmatian doll located at the middle-top of the image. (A) A background made from 20-pixel tiles (Experiment 1). (B) A low TBS background made from 15-pixel tiles (Experiments 2 and 3). (C) A high TBS background made from 35-pixel tiles (Experiments 2 and 3). Note that randomized configurations of search items were used in the actual stimuli; identical configurations are shown here so as to highlight the TBS manipulation.

this experiment, whether adding a camouflage background would result in steeper, equivalent, or shallower slopes relative to a uniform background condition, the data are less clear. Least-squares regression lines fit to the TP data had slopes of 23 and 18 ms/item in the camouflage and uniform conditions, respectively. An opposite pattern emerged in the TA data, with a 58 ms/item slope in the camouflage condition and a 64 ms/item slope in the uniform condition. Despite the small sizes of these slope differences (about 5 ms/item in either direction), both background × set size interactions were reliable; target-present, $F(4,76) = 4.63$, $p < .01$, target-absent, $F(4,76) = 3.46$, $p < .05$.

Further complicating the interpretation of these data is the suggestion of a non-linearity in the RT × set size function. Each distractor added to a camouflage display covers a bit more of the target-similar background. Given that false target signals will be less likely to arise from covered background regions, it is reasonable to expect that the search slope might flatten at large set sizes. We conducted trend analyses to explore this possibility and found a significant quadratic trend in the target-present RT × set size function, $F(1,19) = 12.07$, $p < .01$. Repeating this analysis after excluding the 49-item set size condition revealed only a linear trend in the TP data, $F(1,19) = 12.49$, $p < .01$. While these analyses suggest an attenuation of the RT × set size function at large set sizes, two other analyses question the meaningfulness of this effect with regard to backgrounds. First, no corresponding quadratic trend was found in the TA data, $F(1,19) = 3.52$, $p = .08$. If the deceleration in the TP search slope was due to the background becoming covered with distractors, one would also expect a decelerating search slope when targets did not appear in the displays. Second, the effect size for the TP background × set size interaction was quite small, $\hat{\eta}_p^2 = .2$, far smaller than the size of the background effect, $\hat{\eta}_p^2 = .79$, or the set size effect, $\hat{\eta}_p^2 = .90$. If search is guided to a camouflage background, as would be predicted by feature-based theories, covering a portion of this background with distractors has only a negligible effect on this guidance process, and only at very large set sizes.

In summary, search slopes increased slightly with camouflage in the TP data, consistent with the inefficient-rejection hypothesis and Wolfe et al.'s Experiment 6, but decreased slightly with camouflage in the TA data, consistent with the distractor-independence hypothesis. Further supporting distractor-independence is the fact that RTs did not increase between 39 and 49 items in the TP data. However, given the clear set size effects obtained under camouflage conditions, and the fact that slope flattening was limited to only the target-present data at the largest set size, a strong version of the distractor-independence hypothesis can be dismissed. We therefore conclude that the data from Experiment 1 are most consistent with the efficient-rejection hypothesis. Adding distractors to a camouflage background failed to systematically increase or decrease search efficiency relative to a no-camouflage baseline. The only meaningful effect of background was an elevation of slope intercept, with the size of this background cost being larger in the TA data relative to the TP data. Following the framework established by Wolfe et al. (2002), we interpret this pattern as evidence for either the

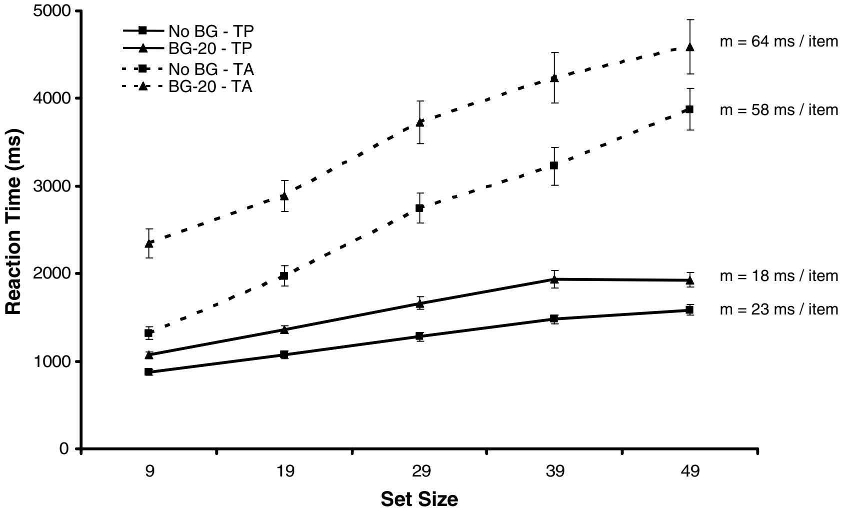

Fig. 4. Mean reaction times from Experiment 1. BG indicates the camouflage background data. No BG indicates the uniform background data. Error bars indicate one standard error of the mean.

creation of false targets in the background (effectively adding a constant number of items to the search set size) or the slowing of the post-attentive identification of search items.

3. Experiment 2

Experiment 1 showed that search is impaired when performed in conjunction with a target-similar background, but questions remain as to how best to characterize this effect of camouflage in terms of search slopes. One reason for our conflicting patterns of results may be that there exist multiple modes of search behavior under camouflage conditions, with the degree of camouflage determining the mode used in a given search task. If the level of TBS used in Experiment 1 did not consistently induce any one of these search modes, then evidence for multiple patterns might appear in the data.

A potential relationship between TBS and search mode might be described in terms of the efficient-rejection, inefficient-rejection, and distractor-independence hypotheses. At relatively low levels of TBS, camouflage backgrounds may be treated simply as complex backgrounds, resulting in the elevated intercepts reported in Experiment 1 and described more thoroughly by Wolfe et al. (2002; Experiments 1–4). The search process would remain dominated by the inspection and efficient-rejection of distractors (efficient-rejection hypothesis). At higher levels of TBS, search slopes might increase as finer discriminations would be needed to segment each distractor from the target-like background (inefficient-rejection hypothesis). Just as search efficiency decreases with increasing target-distractor similarity (Duncan & Humphreys, 1989), so too might be the case for target-background similarity in this mode of search. Finally, at very high levels of TBS, search might shift into a qualitatively different mode, one oriented toward the inspection of the camouflage background rather than the distractors (distractor-independence hypothesis). To better understand the effects of TBS on search, in Experiment 2 we manipulated the degree of camouflage by varying the size of the target region used to tile the search background.

We also monitored eye movements in Experiment 2 in hopes of better characterizing search under camouflage conditions. Eye movements can reveal search processes that are not immediately obvious from a manual RT analysis (Zelinsky et al., 1997), and this is particularly true if the task might engage multiple modes of search behavior. Consider the possibility that observers initially inspect the distractors during each search trial (efficient-rejection or inefficient-rejection hypotheses), but shift to a background search mode (distractor-independence hypothesis) if the distractor search fails to reveal the target. Teasing apart the sequential use of these search modes within a trial would be difficult if limited to a manual RT dependent measure. However, with the higher spatio-temporal resolution provided by oculomotor measures, this problem becomes tractable. If observers were systematically conducting a background search following a distractor search, then a relatively large percentage of the early fixations in a trial should be on distractors, and a relatively large percentage of the later fixations should be on the background. More generally, to the extent that we find fixations devoted to the background regions of the display, these fixations can be interpreted as evidence for the use of a distractor-independent search strategy.

3.1. Method

3.1.1. Participants

Eight undergraduate students from Stony Brook University participated for course credit. All had normal, or corrected-to-normal vision, by self-report, and none were observers in Experiment 1.

3.1.2. Apparatus and stimuli

Eye movement and manual RT data were collected using the EyeLink II eye tracking system (SR Research Ltd.). The spatial resolution of this video-based eye tracker was estimated to be 0.2°, and eye position was sampled at 500 Hz. Search displays subtended $27^\circ \times 20^\circ$ (800×600 pixels) and were presented in color on a ViewSonic 19" flat-screen CRT monitor at a refresh rate of 100 Hz. A custom-made program written in Visual C/C++ (v. 6.0) and running under Microsoft Windows XP was used to control the stimulus presentation. Head position and viewing distance were fixed with a chinrest, and all responses were made with a GamePad controller attached to the computer's USB port. Judgments were made with the left and right index-finger triggers; trials were initiated with a button operated by the left thumb.

The stimuli were the same as those used in Experiment 1, with the following two exceptions. First, rather than having just one level of TBS, two levels were used in Experiment 2. A 15×15 and a 35×35 pixel region ($.3^\circ \times 3^\circ$ and $.9^\circ \times 9^\circ$, respectively) was taken from the center of each object and used to tile an 800×600 pixel display, thereby creating two camouflage backgrounds for use with each target. Note that the level of TBS in these displays varied directly with the size of the tile patch, meaning that a target appearing on a 15-pixel background was relatively distinct (low TBS; Fig. 3B) whereas the same target appearing on a 35-pixel background was highly camouflaged (high TBS; Fig. 3C).[1] Second, because it was not possible to obtain a 35×35 pixel tile from 13 of the 50 objects used in Experiment 1, these 13 objects were replaced with other objects of more regular shape. These new objects, also selected from the Hemera collection, were of the same type (i.e., toys) and size as the other 37 objects in the original stimulus set.

3.1.3. Design and procedure

There were 540 experimental trials, evenly divided into three background conditions (uniform dark, low TBS, and high TBS), two target conditions (present or absent), and three set size conditions (19, 34, and 49 objects), leaving 30 trials per cell of the factorial design. Note that only three set size conditions were used in this experiment so that the running time could be kept under 2 h per observer. All manipulations were again randomly interleaved throughout the experiment.

Prior to data collection, observers participated in 20 practice trials, followed by a 9-target calibration procedure needed to map eye positions to screen locations. The tracker was drift corrected before each trial by having observers press a button when looking at a central fixation target. The remaining procedure was identical to the description provided under Experiment 1, except for the provision of accuracy feedback following a search judgment. On target-present trials, a green box was drawn around the target object informing the observer as to its location in the search display. The words "No target" were drawn to the screen on target-absent trials. Both feedback displays were visible for 1 s, followed immediately by the fixation target signaling the start of the next trial.

3.2. Results and discussion

3.2.1. Manual errors

Miss rates increased with TBS in our task. Misses averaged 9.6%, 9.2%, and 32.8% in the uniform, low, and high TBS conditions, respectively. There was also a modest increase in miss rates with set size; 10.1%, 19.2%, and 22.2% in the 19, 34, and 49-object displays, respectively, resulting in a background × set size interaction, $F(4,28) = 6.00$, $p < .01$. Errors were uniformly infrequent in the TA data, averaging less than 4% in all of the background and set size conditions. Although obtaining a high error rate (33%) is never desirable in a search experiment, in this case it is informative, indicating that our camouflage manipulation was successful. Under high camouflage conditions one would expect observers to frequently fail to locate the target, and that is exactly what happened when targets appeared against a background of 35×35 pixel tiles. The significant background × set size interaction in these error data also provides some initial evidence against the distractor-independence hypothesis. Far from being irrelevant to the task, adding distractors to the display resulted in misses increasing with TBS.

3.2.2. Manual RTs

Fig. 5A shows RT × set size functions for each of the background and target conditions when observers responded correctly. Consistent with Experiment 1, RTs increased dramatically with TBS, and these increases were larger in the TA data (background × target interaction; $F(2,14) = 8.27$, $p < .005$). Observers searching a TP (TA) display took an average of 438 ms (1259 ms) longer to respond in the low TBS condition compared to the uniform background condition, and an additional 1775 ms (3171 ms) in the high TBS condition relative to the low TBS condition. In contrast to these large differences in function intercepts, the effect of TBS on search slopes was relatively minor. Slopes in the TP data were 32, 33, and 37 ms/item in the uniform, low, and high TBS conditions, respectively. Corresponding TA slopes were 86, 70, and 58 ms/item. There were no significant interactions between set size and TBS in either the target-present, $F(4,28) = .66$, $p = .628$, or target-absent, $F(4,28) = 1.99$, $p = .124$, data.

The manual data from Experiment 2 replicate and clarify the patterns found in Experiment 1. The effect of TBS

---

[1] Given our concern that the 20×20 pixel level of TBS used in Experiment 1 did not induce the consistent use of a single search mode, in Experiment 2 we chose higher and lower levels of TBS. However, rather than symmetrically bracketing the 20×20 level, we chose a tile size for the high TBS condition (35×35 pixels) that more closely approximates true camouflage conditions.

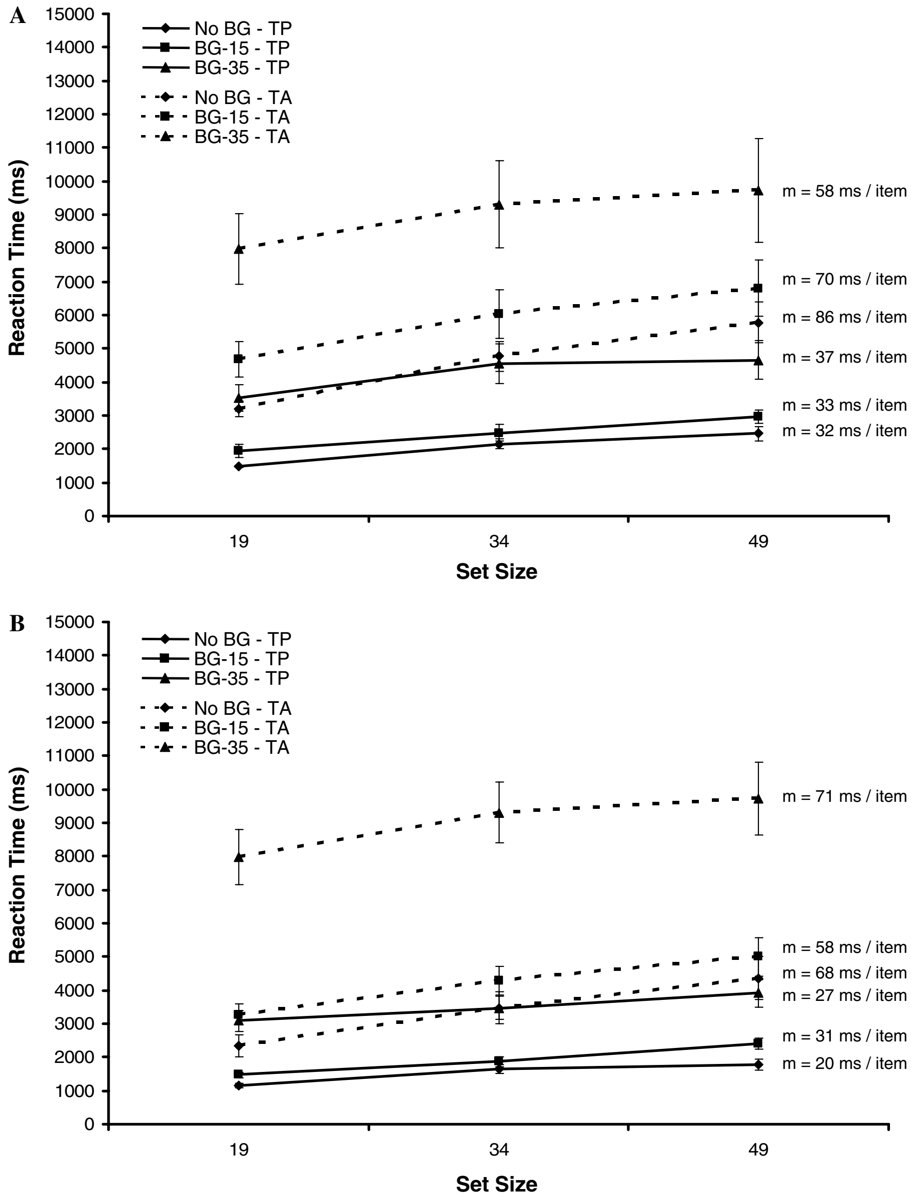

Fig. 5. Mean reaction times for Experiments 2 (A) and 3 (B). BG-15 = low TBS condition with 15-pixel tiles; BG-35 = high TBS condition with 35-pixel tiles. No BG = uniform background condition. Error bars indicate one standard error of the mean.

on search is expressed almost entirely by elevations of the RT × set size functions, not by changes in slope. In terms of our theoretical framework, these data offer compelling support for the efficient-rejection hypothesis, and allow us to reject both the inefficient-rejection hypothesis and a strong version of the distractor-independence hypothesis. Observers were clearly inspecting distractors, and their ability to reject these items was not affected by TBS. The fact that slopes did not increase with TBS in our task also clarifies data reported by Wolfe et al. (2002, Experiment 6). Recall that these authors found that search slopes increased as targets and distractors were made more similar to the background. Our data suggest that this relationship was due to distractor-background similarity, not TBS. Contrary to the inefficient-rejection hypothesis, the time needed to reject background-dissimilar distractors does not depend on the similarity between the background and target.

Left unresolved from our analyses of the manual data is the possibility of a sequential multi-mode search strategy. Our observers may have systematically searched through the distractors, thereby producing a set size effect, then shifted to a search of the background before making their judgments. As a working hypothesis, we believe that this background search may have been responsible for the effect of TBS observed in the intercepts. As TBS increased, observers might have spent more time searching the background, thereby increasing the overall search time without affecting search slopes. We analyzed the gaze behavior of observers in our task to test this hypothesis.

3.2.3. Number of fixations

To quantify the inspection of background regions, we classified each fixation made during search (i.e., before the button press ending the trial) as being either an object or background fixation. A fixation was classified as being on an object if it was located within a 1° (50 pixel) radius of an object's center. Given that each object subtended at most 1.5°, this means that fixations could be slightly off of an object yet still be classified as an object fixation. All fixations not falling within the bounding circle surrounding an object were classified as background fixations.

Fig. 6A shows the number of fixations on the background and Fig. 7A shows the proportion of background fixations, both for correct trials. The first thing to note from these analyses is that background fixations were not uncommon in this task, accounting for approximately 25–50% of all the fixations occurring during search. Although this breakdown means that roughly 50–75% of the fixations were on distractors, the very existence of background fixations supports the contribution of a distractor-independent search process. Also consistent with the distractor-independence hypothesis are the significant increases in the number and proportion of background fixations with TBS in both the TP, $F(2,14) \geqslant 8.80$, $p < .005$, and TA data, $F(2,14) \geqslant 17.78$, $p < .001$. As TBS increased, observers made more fixations and devoted more of their fixations to inspecting the background. Predictably, these effects of TBS were larger in the TA data, as confirmed by significant background × target interactions, $F(2,14) \geqslant 7.10$, $p < .01$. Because RTs were longer in the TA trials, there were more fixations and therefore a greater opportunity for the expression of background differences. Also as expected, the number and percentage of background fixations decreased as set size increased, resulting in significant background × set size interactions in the TA data, $F(4,28) \geqslant 2.87$, $p < .05$. At larger set sizes there was less background to inspect, resulting in fewer background fixations.

Although there are clear effects of TBS on the fixations made during search, the base rate of fixations in the uniform background condition should be considered when interpreting the size of these effects. Observers viewing a uniform background display in the TA 19-object condition made 40% of their fixations to the background, and none of these fixations can be explained in terms of TBS. In fact, none of the hypotheses offered in this study or in the Wolfe et al. (2002) study offer an adequate theoretical accounting of this behavior. Following Zelinsky et al. (1997), we interpret these fixations as evidence for center-of-gravity averaging in oculomotor programming (see also Findlay,

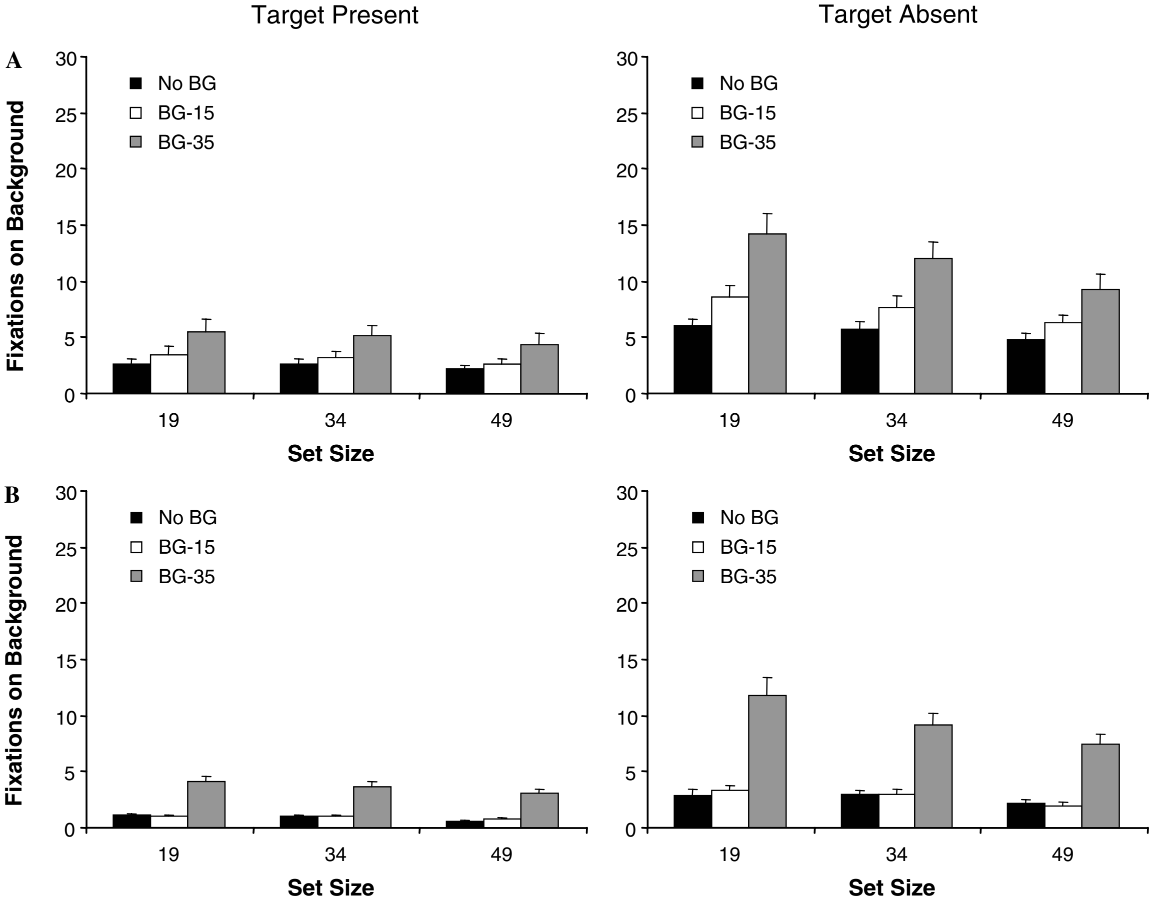

Fig. 6. Mean number of fixations made to the background in Experiments 2 (A) and 3 (B). Error bars indicate one standard error of the mean.

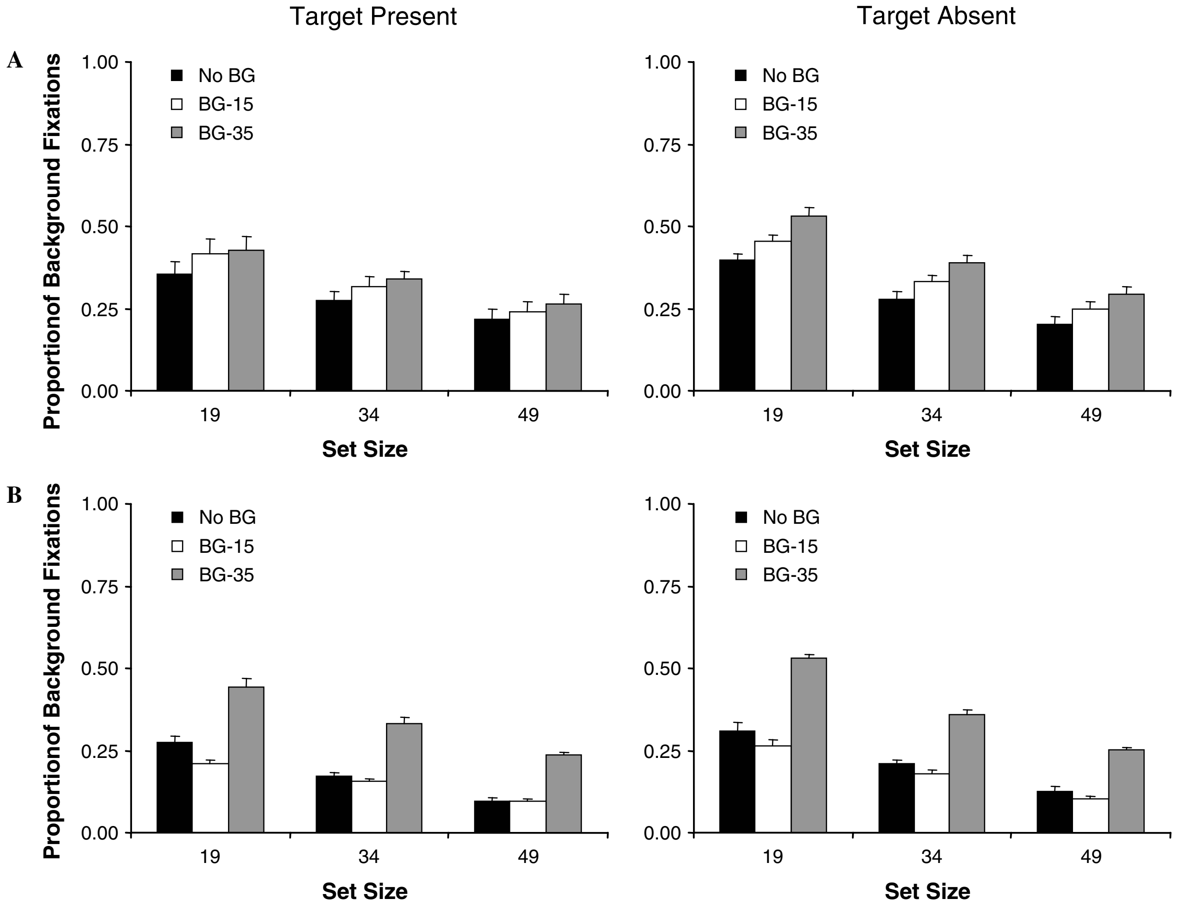

Fig. 7. Mean proportion of fixations per trial on the background in Experiments 2 (A) and 3 (B). Error bars indicate one standard error of the mean.

1982, 1987). When observers are freely moving their eyes and complying with instructions to search as rapidly as possible, many of their fixations fall between adjacent objects rather than directly on objects. Center-of-gravity averaging is potentially useful to observers as it might enable them to reject two (or more) objects with a single fixation, but in the current context such averaging behavior would be counted toward our measure of the background fixation rate. Indeed, observers viewing the 49-object high TBS displays averaged 33 fixations during their TA search, meaning that an estimated 16 distractors per display were rejected without ever being directly fixated. To the extent that the background fixation rate in the uniform condition serves as a control for center-of-gravity averaging during search, then our estimates of true background fixations become more modest. After factoring out the uniform background base rate, the increase in background fixations as a function of TBS ranged from a low of .5 fixations (4%) in the TP low TBS 49-object data, to a high of 8.1 fixations (31%) in the TA high TBS 19-object data.

3.2.4. Analysis of fixation sequence

Background fixations increased with TBS, but we do not yet know when these background fixations occurred during search. One possibility is that observers shifted to a background search mode after determining that the target did not appear among the distractors. Another possibility is that there exists no background search mode and that observers simply looked occasionally to the background as they inspected the display objects. To distinguish between these two possibilities, we analyzed the background fixations as a function of when they occurred relative to the distractor fixations. Fig. 8A shows this analysis for the three background conditions in the TA data, collapsed across set size.[2] If a background fixation was made before the observer fixated 10% of the distractors, that fixation would be assigned to the 0.10 bin. Likewise, if a background fixation occurred after more than 90% of the distractors were fixated (e.g., all but two of the distractors in a 34-object display), that fixation would be assigned to the 1.0 bin. To test for the strong version of the multi-mode hypothesis, a special >1.0 bin was created to hold those background fixations occurring after all of the distractors were visited by gaze.

Unexpectedly, we found that most background fixations occurred very early during search, before 10% of the dis-

---

[2] An analysis of the TP data yielded similar results, but these data were less stable due to the smaller number of fixations in this condition and the greater variability in the number of distractors inspected during a trial.

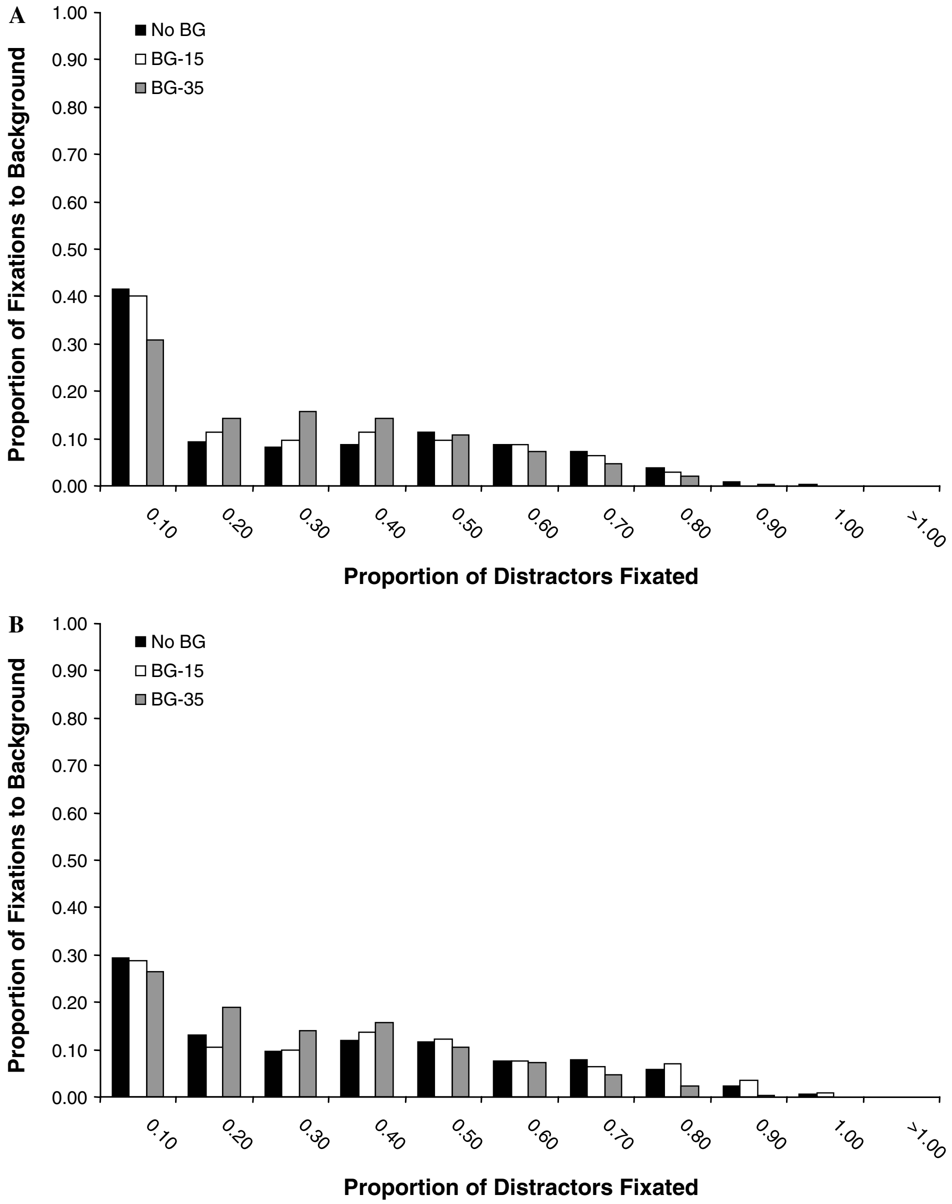

Fig. 8. Mean proportion of fixations on the background as a function of the proportion of distractors fixated during TA trials. (A) Data from Experiment 2; (B) data from Experiment 3. BG-15 = low TBS condition with 15-pixel tiles; BG-35 = high TBS condition with 35-pixel tiles; No BG = uniform background condition.

tractors were fixated. We attribute these fixations to the fact that observers were initially fixated on the background at the start of each trial, and often made two or three additional background fixations while settling on a distractor from which to start their search. Contrary to the multimode hypothesis, the fact that the >1.0 bin is empty means that observers never waited until all of the distractors were inspected before making background fixations. Indeed, once the early background fixations are discounted, this analysis offers no evidence to support the existence of a separate mode of search devoted to the inspection of backgrounds, not even when these backgrounds looked very similar to the target. Instead, observers appear to have occasionally fixated the background while searching through the distractor objects, with the temporal distribution of these background fixations being relatively uniform throughout search.

4. Experiment 3

The data from Experiment 2 were relatively clear; observers frequently looked to target-similar backgrounds during their search, but they did so while inspecting the distractors. However, despite its clarity, this data pattern is also highly counterintuitive. According to traditional conceptions of search (Duncan & Humphreys, 1989; Treisman, 1988; Wolfe et al., 1989), a distractor is distracting in proportion to its likelihood of being confused with a target. In most search tasks, it is therefore reasonable to expect fixations on distractors because these items share more features with the target compared to a uniform background. This was not the case in Experiment 2 in which targets in the high TBS condition were far more similar to the background than the distractors. Indeed, under high camouflage conditions the distractors indicated display locations where the observers *should not look*. Why then did our observers not adopt the more rational strategy of inspecting the background?

One possibility is that our experimental design may have biased observers against adopting a background mode of search. The background conditions in both Experiments 1 and 2 were randomly interleaved throughout the experiments, meaning that observers were not able to anticipate whether each new search display would have a uniform, low TBS, or high TBS background. Because observers were unable to set themselves into an optimal background-specific search mode, they may have adopted a mode that would work best for the majority of the trials. Arguably, this would be a mode emphasizing distractors rather than backgrounds. Assuming that low TBS search is most efficient under a distractor search mode, observers in Experiment 2 might have set themselves for a distractor search as this would work best in two-thirds of the trials (i.e., those with uniform and low TBS backgrounds). To test for this possibility, in Experiment 3 we made the background manipulation a between-subjects variable. If interleaving was discouraging observers from adopting a background search mode, we should now find evidence for this mode in the high TBS condition. Specifically, we would expect to see shallow slopes and a large proportion of background fixations under high camouflage conditions, and steep slopes with few background fixations under uniform and low TBS conditions.

4.1. Method

Twenty-four undergraduate students from Stony Brook University participated in this experiment for course credit. None of these observers participated in Experiments 1 or 2, and all of them had normal or corrected-to-normal vision, by self-report. These 24 observers were randomly assigned to uniform, low TBS (15-pixel tiles), and high TBS (35-pixel tiles) background conditions. Except for this shift from a within-subjects to a between-subjects design, all aspects of the stimuli, apparatus, design, and procedure were identical to Experiment 2.

4.2. Results and discussion

4.2.1. Manual errors

As in Experiment 2, miss rates in Experiment 3 increased with TBS and set size. Misses averaged 13.4%, 14.6%, and 30.1% with uniform, low TBS, and high TBS backgrounds, and 13.8%, 21.3%, and 23% with 19-, 34-, and 49-object displays, resulting in significant main effects of background, $F(2,21) = 12.45$, $p < .001$, and set size, $F(2,42) = 13.3$, $p < .001$. Errors in the TA data also increased with TBS (3.8%, 2.2%, and 9.3% with uniform, low TBS, and high TBS backgrounds), but this trend was not significant, $F(2,21) = 3.21$, $p > .05$. Overall, Experiments 2 and 3 showed the same patterns of errors, but errors were slightly more common in Experiment 3, particularly in the uniform and low TBS conditions.

4.2.2. Manual RTs

Fig. 5B shows the mean RTs for correct trials in the background and target conditions plotted as a function of set size. These data replicate almost perfectly the patterns reported for Experiments 1 and 2. As in the previous experiments, RTs increased with TBS in both the TP, $F(2,21) = 22.44$, $p < .001$, and the TA, $F(2,21) = 12.89$, $p < .001$, data. This effect of TBS on search was again larger in the TA data, resulting in a significant background × target interaction, $F(2,21) = 7.71$, $p < .005$. Also consistent with Experiment 2 is our failure to find significant differences in slope among the background conditions in either the TP, $F(4,42) = 1.18$, $p > .10$, or TA data, $F(4,42) = .56$, $p > .10$. The effect of increasing TBS similarity on the RT × set size functions was expressed almost entirely by higher intercepts rather than changes in slope. However, RTs in Experiment 3 were faster overall compared to Experiment 2. This trend, when considered together with the higher error rates in Experiment 3, suggests that blocking the background conditions may have resulted in observers adopting more liberal response criteria. Aside from this evidence for a speed-accuracy tradeoff, blocking appears not to have otherwise affected the patterns of manual responses.

4.2.3. Eye movement data

Fig. 6B shows the average number of fixations on the background, and Fig. 7B shows the proportion of background fixations, both for correct trials. All of the data patterns reported for these measures in Experiment 2 replicated in Experiment 3. Most notably, as increasing TBS approximated true camouflage conditions, observers made more fixations, and a greater proportion of these fixations, on the background, in both TP, $F(2,21) \geqslant 29.67$, $p < .001$, and TA trials, $F(2,21) \geqslant 30.53$, $p < .001$. However, although the number and proportion of background fixations in the high TBS conditions did not meaningfully differ between Experiments 2 and 3, there were fewer fixations on the uniform and low TBS backgrounds in Experiment 3 relative to Experiment 2. This lower background fixation rate is not surprising given the faster RTs in Experiment 3 and the typically strong correlation between RT and number of fixations (Zelinsky & Sheinberg, 1997). With regard to the hypothesis in question, these findings suggest that blocking the background conditions did not cause observers to increase their scrutiny of high TBS backgrounds. It is perhaps fairer to conclude that interleaving the background conditions in Experiment 2 made it more difficult for observers to avoid looking occasionally to backgrounds during search.

The fixation number and proportion analyses revealed no evidence for a background search mode, but did observers respond to blocking by distributing their background fixations differently during search? If blocking made it easier for observers to adjust their search mode depending on camouflage conditions, we should have found a greater proportion of background fixations occurring before distractor fixations in the high TBS group. However, as the temporal fixation analysis in Fig. 8B shows, background fixations in Experiment 3 were distributed much like those in Experiment 2. Despite knowledge that the target would be highly camouflaged by the background, observers in the high TBS group still chose to look primarily at the target-dissimilar distractors during their search.

5. Experiment 4

What was it about the distractors used in Experiments 2 and 3 that caused observers to direct their gaze to these search items? One possibility is that the attentional processes governing search operate on object-based representations (e.g., Baylis & Driver, 1993; Duncan, 1984; Goldsmith, 1998; Kramer & Jacobson, 1991; Prinzmetal, 1981; Vecera & Farah, 1994). If attention is biased toward objects, then the background regions of the display will be relatively neglected by search. A second possibility is that observers were shifting their gaze to texture discontinuities in the display (e.g., Julesz, 1981; Thielscher & Neumann, 2005). One necessary byproduct of the tiling process used to generate our stimuli is the creation of repeating patterns in the backgrounds (Fig. 3). Observers might therefore have looked to the distractors because these patterns give rise to highly salient discontinuities in the background texture, not because of their special object status.

To tease apart these competing possibilities, in Experiment 4 we modified the search displays to include both object distractors and non-object texture discontinuities. Roughly half of the distractors in each display consisted of the same toy objects used in Experiments 1–3; the other half of the distractors were created by rotating a patch of the background, thereby creating a local texture discontinuity. Importantly, these texture elements lacked the semantic associations and internal part structure that typically define objects (Singh & Hoffman, 2001).[3] If observers were biased against looking to the background due to the uniformity of the background texture, fixations should now be more evenly distributed between objects and texture elements, and the proportion of fixations on both distractor types should be greater than the proportion of background fixations. Indeed, because the texture elements would be more similar to the camouflaged target than the objects, observers may even prefer to look to the texture elements during their search. We will refer to this as the discontinuity-bias hypothesis. However, if observers avoided fixating the background because their attention was biased to objects, then the proportion of object fixations should remain high (as observed in Experiments 2 and 3) and the proportions of fixations on the background and texture elements should not differ. We will refer to this as the object-bias hypothesis.

5.1. Method

Eight undergraduate students from Stony Brook University participated in this experiment for course credit. None of these observers participated in Experiments 1, 2, or 3, and all of them had normal or corrected-to-normal vision, by self-report. As in the previous experiments, these observers indicated the presence or absence of a target under low TBS (15-pixel tiles) and high TBS (35-pixel tiles) camouflage conditions. However, rather than having the distractors consist solely of children's toys, the 19-, 34-, and 49-item displays were now divided evenly (as much as possible) between toys and texture elements (see Fig. 9). A texture element was created by randomly rotating a 75×75 pixel patch of the background (a $1.5° \times 1.5°$ dimension corresponding to the size of the bounding box enclosing each distractor object) by 45°, 135°, 225°, or 315°, thereby creating an orientation discontinuity relative to the background texture. Perceptually, a texture element looked like a salient dimple in the background texture rather than a discrete object. All other aspects of the stimuli, apparatus, design, and procedure were identical to Experiment 2.

5.2. Results and discussion

5.2.1. Manual errors

Miss rates increased with TBS and set size. Misses averaged 3% and 37.8% with low TBS and high TBS backgrounds, $F(1,7) = 99.92$, $p < .001$, and 10.4%, 13.9%, and 16.49% with 19-, 34-, and 49-distractor displays, $F(2,14) = 9.02$, $p < .005$. False positive rates were less than 4% in all conditions. These patterns of errors are generally consistent with the patterns observed in Experiments 2 and 3, although misses were more prevalent under high TBS conditions in Experiment 4. We attribute this higher miss rate to the greater target-distractor similarity in this experiment. The fact that half of the distractors consisted of texture elements that were visually similar to the camouflaged target made this search task very challenging in the high TBS condition (compare Figs. 9A and B). A higher miss rate would also be expected if observers were biased to search the objects in the display and not the texture elements. Such an object bias might cause highly camouflaged targets to be excluded from the search set, and consequently missed.

---

[3] Although defining what is and is not an object is clearly beyond the scope of this study, we do believe that the texture elements used in Experiment 4 were less object-like than children's toys, thereby satisfying the specific requirements of the question under investigation.

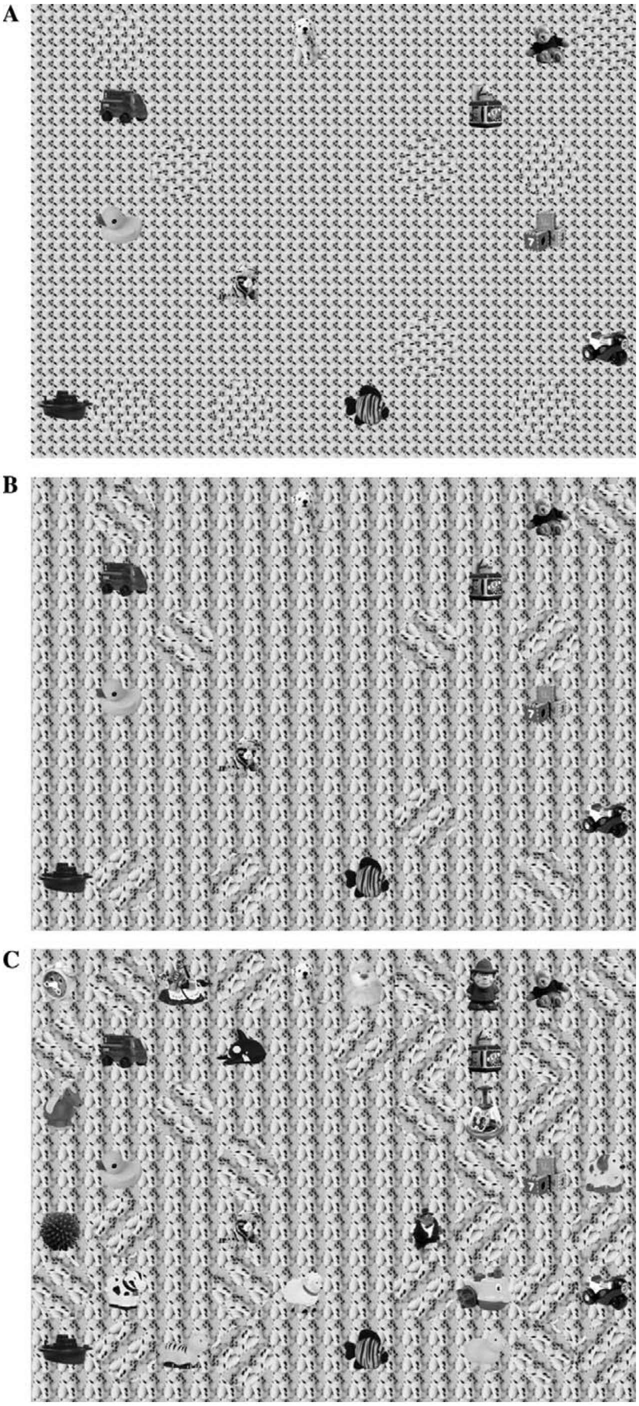

Fig. 9. Sample search displays illustrating two levels of TBS and set size. The target in (A–C) is a Dalmatian doll (middle-top of each image). (A) A low TBS background (15-pixel tiles) in the 19 set size condition. (B) A high TBS background (35-pixel tiles) in the 19 set size condition. (C) A high TBS background (35-pixel tiles) in the 49 set size condition. Note that search items were randomly configured in the actual stimuli; correlated spatial configurations are shown here so as to highlight the effects of set size and TBS on the perception of the texture elements.

5.2.2. Manual RTs

Fig. 10 plots mean RTs for correct trials as a function of TBS and set size. As in Experiments 1–3, RTs increased with TBS in both the TP, $F(1,7) = 33.37$, $p < .005$, and the TA data, $F(1,7) = 16.33$, $p < .01$, and this increase was larger in the TA condition, $F(1,7) = 6.23$, $p < .05$. Slopes did not significantly differ between the background conditions in either the TP, $F(2,14) = 1.79$, $p = .2$, or the TA data, $F(2,14) = 2.58$, $p = .11$, again indicating an effect of TBS on the intercept of the RT × set size function rather than the slope.

5.2.3. Eye movement data

How did observers allocate their gaze to the two distractor types while searching for a target? An object bias predicts fixations on the toy distractors; a discontinuity bias predicts a more equal distribution of fixations between the toy and texture distractors and a higher proportion of fixations on distractors compared to the background. To test these hypotheses we calculated the average number (Fig. 11) and proportion of fixations per trial (Fig. 12) to the background, texture elements, and the search objects, both for correct trials only. Analyses of these data allow us to reject the strong forms of both the object-bias and texture-bias hypotheses. As indicated by the high TBS data in Fig. 12, observers preferred to fixate texture elements over objects in the 19-item displays in both the TP, $t(7) = 2.46$, $p < .05$, and the TA trials, $t(7) = 6.01$, $p < .01$. This preference to look at the texture elements is inconsistent with the object-bias hypothesis as these elements are clearly less object-like than toys. Non-significant trends also exist for a texture preference in the 19-item low TBS condition and in the corresponding number of fixation data shown in Fig. 11. Observers searching the 49-item displays showed the opposite fixation preference, clearly preferring to fixate objects over texture elements. This very pronounced preference for objects, which was significant for all 49-item comparisons in Fig. 11, $t(7) \geqslant 4.7$, $p < .005$, and Fig. 12, $t(7) \geqslant 12.36$, $p < .001$, is inconsistent with the texture-discontinuity hypothesis. Indeed, observers tended to look more to the background than the texture elements when searching for targets at this set size regardless of TBS condition.

Rather than providing unambiguous support for purely texture-based or object-based search processes, these data suggest a more complex relationship between search and distractor type, one that depends on the number of distractors in the display. The proportion of fixations to objects increased with set size, $F(1,7) \geqslant 86.5$, $p < .001$, and the proportion of fixations to texture elements decreased with set size, $F(1,7) \geqslant 22.29$, $p < .001$. This set size dependency is confirmed by highly significant region × set size crossover interactions in both the number of fixations, $F(2,14) \geqslant 49.19$, $p < .001$, and the proportion of fixations, $F(2,14) \geqslant 36.43$, $p < .001$. Observers preferred to look at the texture elements at the low set size and the objects at the higher set sizes.

We interpret this crossover interaction as evidence for a modified version of the discontinuity-bias hypothesis. According to this hypothesis, observers look to discrete discontinuities in the display regardless of whether these

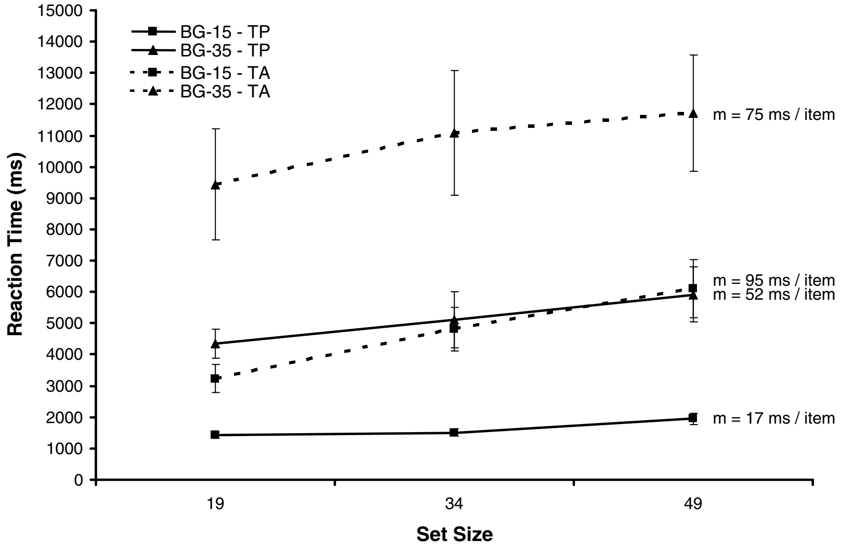

Fig. 10. Mean reaction times for Experiment 4. BG-15 = low TBS condition with 15-pixel tiles; BG-35 = high TBS condition with 35-pixel tiles. Error bars indicate one standard error of the mean.

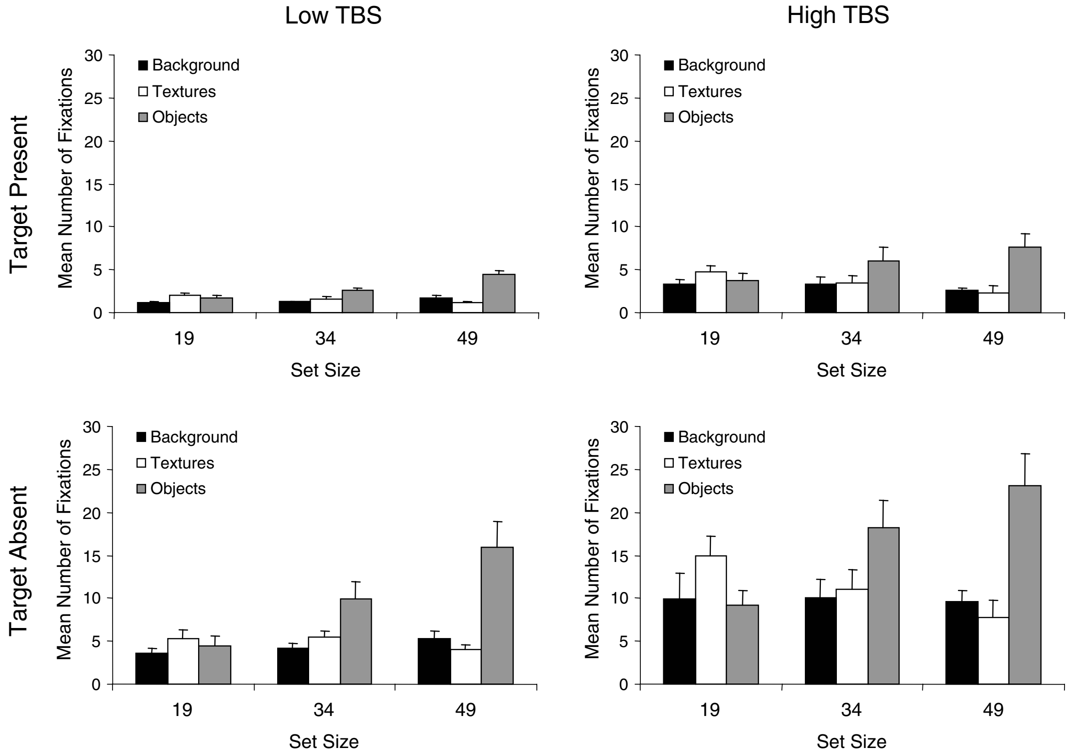

Fig. 11. Mean number of fixations made to the background, texture element, and object regions as a function of set size. Error bars indicate one standard error of the mean.

discontinuities correspond to objects or texture elements. However, we speculate that low-level perceptual grouping mechanisms are instrumental in determining whether search treats these discontinuities as discrete items (which might potentially signal a target) or larger regions of texture within the background pattern. Search theorists have often appealed to the grouping of like distractors into larger perceptual units as an explanation for efficient distractor rejection (e.g., Grossberg, Mingolla, & Ross, 1994; Treisman, 1982), and we believe a similar mechanism might explain our current data. At low set sizes each texture element is likely to be widely spaced from its neighbors and therefore perceived as a non-grouped individuated discontinuity (Fig. 9B). As for observers' preference to fixate these distractors in 19-item displays, we attribute this behavior to the greater visual similarity between the texture

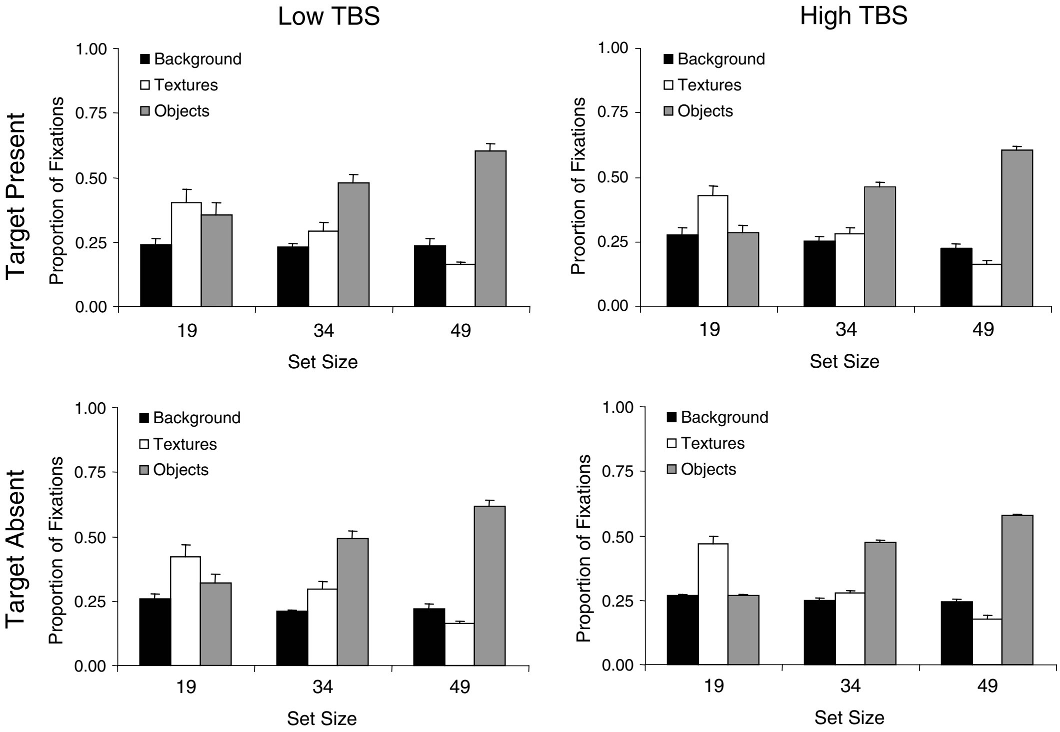

Fig. 12. Mean proportion of fixations made to the background, texture element, and object regions as a function of set size. Error bars indicate one standard error of the mean.

elements and the target under high TBS conditions creating a strong guidance signal for search (Wolfe, 1994a). At higher set sizes, individual texture elements are more likely to become grouped into larger continuous regions of texture discontinuity that can be easily rejected as a potential target. Under this scenario, the visually heterogeneous object distractors would be left as the only individuated discontinuities in the display, making them highly salient and the most likely patterns to attract gaze.

A strength of our proposal is that it does not require the assumption of two qualitatively different search modes, one for textures and another for objects. Instead, the fixation preference indicated by the crossover interaction is explained in terms of a relationship between set size, grouping, and distractor type, within the context of a single discontinuity-based search process. However, it should be noted that this explanation, while offering a plausible interpretation of our data, is also extremely broad in scope. It is not possible to know from this one experiment whether our explanation will generalize to other search tasks, nor was it the goal of this study to provide these sorts of definitive answers. What we did set out to learn from this experiment was whether the fixation preferences observed in Experiments 2 and 3 were a result of the distractors being objects, and in this regard the current data were informative. Although objects are typically correlated with visual discontinuities in an image, we found in Experiment 4 a case where observers preferred to fixate non-object texture discontinuities during their search. We tentatively conclude from this finding that both objects and discontinuities are important determiners of gaze direction during search, and that one underlying cause of an object bias may be the pronounced visual discontinuities that they create in the context of complex backgrounds.

6. General discussion

In a recent and important study documenting the effects of background complexity on search, Wolfe et al. (2002) found that search performance does indeed degrade with increasing background complexity, but that this impairment is expressed almost entirely by an increase in the intercept of the RT × set size function rather than by a change in slope (Experiments 1–4). Search efficiency worsens only under conditions of high similarity between the search items and the background (Experiment 6).

Our work replicates the main data patterns reported by Wolfe et al. (2002; Experiments 1–4), and permits the generalization of their conclusions to a search task using complex real-world objects presented under true camouflage conditions. In four experiments, we found the efficient-selection and rejection of distractors regardless of background presence or absence (Experiment 1), level of TBS (Experiments 2–4), or distractor heterogeneity (e.g., objects and texture elements; Experiment 4). The failure of background complexity to influence search efficiency appears to be a relatively robust effect, observed now in stimuli ranging from Ts and Ls to children's toys. Contrary to Wolfe et al. (2002), none of our manipulations produced the large slope increases that they reported in their Experiment 6, not even when highly target-similar texture elements were used as distractors (Experiment 4). Although we do not know the exact cause of this minor discrepancy, we speculate that Wolfe et al.'s (2002) use of visually simple checkerboard patterns may have resulted in extreme camouflage conditions that we could not recreate using large and unoccluded visually complex objects.

The current study also advances our understanding of the relationship between backgrounds and search in three respects. First, we offer the first clear experimental evidence for an effect of TBS on search. Although Wolfe et al. (2002) also manipulated target-background similarity (Experiments 3 and 4), their manipulations were confounded with changes in distractor-background similarity. We removed this confound by having backgrounds correlate only with the target. We can therefore conclude that the large effects of background on search function intercepts observed in our study, and possibly those reported by Wolfe et al. (2002), were caused by the target being similar to the background, not by the background being similar to the distractors. Future work will attempt to extend this finding to an even wider range of search stimuli and tasks. We are particularly interested in generalizing our work to real-world scenes, which typically have less background regularity and a greater diversity of object scales than the stimuli used in the current study.

Our study's second contribution serves to clarify the underlying cause of the background effect. Wolfe et al. (2002) narrowed the source of the intercept differences to two possibilities: (1) that parts of the background may be mistakenly segmented into objects, thereby creating false distractors, or (2) that backgrounds may slow the process of identifying objects. Based on a failure to find differences between a one-target and a two-target search task (see their Experiment 5 for details), these authors concluded that backgrounds interfere with object identification. The eye movement data from the current experiment are difficult to reconcile with this conclusion. From Figs. 6 and 11 we know that fixations on the background increased with TBS. If backgrounds were slowing the rate of information accumulation about search objects, thereby making them harder to identify, why then would our observers occasionally choose to look away from these objects and fixate on the background? To the extent that observers were making background fixations in our task, we believe these fixations can be better explained by the false distractor hypothesis. While looking through the objects or texture elements, observers occasionally shifted their gaze to a false distractor discontinuity on the background, and these additional fixations contributed to the elevated intercepts and the longer search times.

The third contribution of this study is theoretical and focuses on the slope of the RT × set size function rather than the intercept. In the Wolfe et al. (2002) study, targets and distractors were highly similar, so steep search slopes were to be expected. However, in the current study targets and distractors were visually dissimilar, and the background pattern was correlated only with the target object. Under these conditions, the distractor objects should be functionally irrelevant to the task, leading us to entertain the possibility of finding shallow slopes or even complete distractor-independence. This prediction was not supported by the data, with steep and largely equivalent slopes obtained regardless of background condition. Why did our observers prefer to look at these target-dissimilar distractor objects more than the target-similar background?

We interpret our data as evidence for a modified discontinuity-based search process, one in which search is biased towards objects and other visual discontinuities in a display irrespective of feature guidance signals.[4] The attention and search literatures have long distinguished between processes that are location-based, meaning that attention must be directed to locations before features can be combined into objects (e.g., Itti & Koch, 2001; Koch & Ullman, 1985; Treisman, 1988; Treisman & Gelade, 1980; Wolfe et al., 1989), and processes that are object-based, meaning that visual features are bound preattentively into perceptual objects, which can then be selected for further processing by focused attention (e.g., Baylis & Driver, 1993; Duncan, 1984; Goldsmith, 1998; Kramer & Jacobson, 1991; Prinzmetal, 1981; Vecera & Farah, 1994). Although the relationship between objects and feature locations has made it difficult to cleanly tease apart these theories, our methodology directly pits object-based guidance against feature-based guidance, thereby allowing their dissociation in the context of a search task. The winner of this guidance contest is clear from our manual and eye movement data; when observers were faced with the choice of searching through target-dissimilar objects or target-similar features, they preferred to inspect the objects. Moreover, this seemingly irrational behavior is not likely due to high-level biases causing observers to adopt an object fixation strategy; observers preferred to fixate objects even under high camouflage conditions within a blocked design (Experiment 3).

Consistent with object-based theories of attention and our data from Experiment 4, we believe that preattentive processes segment visual scenes into objects and other salient discontinuities, and that the search process then uses these object-based representations to guide attention and eye movements. Under high camouflage conditions, targets would therefore be frequently missed because they fail to segment from the background and never become candidates for search inspection. Similarly, the background region, although often target-like in appearance, attracted relatively few fixations because it failed the segmentation stage and was not passed to the search process. Occasionally, a minor visual discontinuity on the background might

---

[4] For the purpose of this discussion we will use the term *object* to refer to any visual discontinuity consisting of preattentively bound features. According to this minimalist definition, objects can be thought to exist along a continuum ranging from salient discontinuities in texture to visually complex and semantically meaningful entities (e.g., children's toys).

be segmented by mistake, thereby causing the region to become a false distractor. We believe that the background fixations in our data that cannot be attributed to center-of-gravity averaging were likely due to such segmentation failures. However, despite their differences, we do not believe that the object or discontinuity-based views outlined here are fundamentally incompatible with predominantly feature-based theories of bottom–up (Itti & Koch, 2000; Koch & Ullman, 1985) or top–down (Rao et al., 2002; Zelinsky, 2005) search guidance. Specifically, although saliency-based theories of search deal with points of activation on a saliency map rather than objects, it would be easy to modify these theories to compute salience values for only those scene regions suggested by a preattentive segmentation process. Such a discontinuity-restricted saliency map would combine the strengths of feature and object-based search theories, allowing search to be guided by feature contrast (Itti & Koch, 2000) or the correlation between the target and scene (Zelinsky, 2005), yet still demonstrate a fixation preference for objects and other visual discontinuities as required by the current data.

Our findings also have implications for the understanding, and potential augmentation, of visual search in naturalistic contexts. The clearest extension of our work is to situations in which a searcher is attempting to acquire a camouflaged target. Despite knowledge that the target may be quite similar in appearance to the background, our data suggest that search behavior may nevertheless be guided to the easily segmented patterns in a scene. One may therefore want to resist the urge to look at these patterns, focusing attention instead on the more subtle irregularities in the background that may define a camouflaged target. Shifting roles, if one's goal is to avoid detection by a searcher, it may be wise to distance oneself from highly salient objects that are likely to attract scrutiny, trusting your camouflage to enable you to "hide in plain sight". Future work will use computer generated 3D models to test these search heuristics under more naturalistic camouflage conditions.

Acknowledgments

This work was supported by grants from the Army Research Office (DAAD19-03-1-0039) and the National Institute of Mental Health (R01 MH63748) to G.J.Z. We thank Xin Chen and Christopher Dickinson for their technical assistance and thoughtful comments throughout this project, and Shannon Fitzhugh and Joseph Schimdt for their assistance with data collection.

References

Aks, D., & Enns, J. (1996). Visual search for size is influenced by a background texture gradient. *Journal of Experimental Psychology: Human Perception and Performance, 22*, 1467–1481.

Baylis, G. C., & Driver, J. (1993). Visual attention and objects: Evidence for hierarchical coding of location. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 451–470.

Biederman, I., Blickle, T. W., Teitelbaum, R. C., & Klatsky, G. J. (1988). Object search in nonscene displays. *Journal of Experimental Psychology: Learning, Memory, and cognition, 14*, 456–467.

Biederman, I., Glass, A. L., & Stacy, E. W. (1973). Searching for objects in real-world scenes. *Journal of Experimental Psychology, 97*(1), 22–27.

Bravo, M. J., & Farid, H. (2004). Search for a category target in clutter. *Perception, 33*(6), 643–652.

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General, 113*, 501–517.

Duncan, J., & Humphreys, G. (1989). Visual search and stimulus similarity. *Psychological Review, 96*, 433–458.

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

Enns, J. T., & Rensink, R. A. (1991). Pre-attentive recovery of three-dimensional orientation from line drawings. *Psychological Review, 98*(3), 335–351.

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

Findlay, J. (1982). Global visual processing for saccadic eye movements. *Vision Research, 22*, 1033–1045.

Findlay, J. (1987). Visual computation and saccadic eye movements: A theoretical perspective. *Spatial Vision, 2*, 175–189.

Goldsmith, M. (1998). What's in a location? Comparing object-based and space-based models of feature integration in visual search. *Journal of Experimental Psychology: General, 2*, 189–219.

Gould, J. D., & Carn, R. (1973). Visual search, complex backgrounds, mental counters, and eye movements. *Perception and Psychophysics, 14*(1), 125–132.

Grossberg, S., Mingolla, E., & Ross, W. (1994). A neural theory of attentive visual search: Interactions of boundary, surface, spatial, and object representations. *Psychological Review, 101*, 470–489.

He, Z. J., & Nakayama, K. (1992). Surfaces versus features in visual search. *Nature, 359*(6392), 231–233.

Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology, 50*, 243–271.

Henderson, J. M., Weeks, P., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 210–228.

Ho, G., Scialfa, C. T., Caird, J. K., & Graw, T. (2001). Visual search for traffic signs: The effects of clutter, luminance, and aging. *Human Factors, 43*(2), 194–207.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shift of visual attention. *Vision Research, 40*(10–12), 1489–1506.

Itti, L., & Koch, C. (2001). Computational modeling of visual attention. *Nature Reviews Neuroscience, 2*(3), 194–203.

Julesz, B. (1981). Textons, the elements of texture perception and their interactions. *Nature, 290*, 91–97.

King, M. G., Stanley, G. V., & Burrows, G. D. (1984). Visual search processes in camouflage detection. *Human Factors, 26*(2), 223–234.

Kleffner, D. A., & Ramachandran, V. S. (1992). On the perception of shape from shading. *Perception and Psychophysics, 52*(1), 18–36.

Koch, C., & Ullman, S. (1985). Shifts of selective visual attention: Toward the underlying neural circuitry. *Human Neurobiology, 4*, 219–227.

Kramer, A. F., & Jacobson, A. (1991). Perceptual organization and focused attention: The role of objects and proximity in visual processing. *Perception and Psychophysics, 50*, 267–284.

Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities. *Vision Research, 41*(25–26), 3559–3565.

Levin, D. T. (1996). Classifying faces by race: The structure of face categories. *Journal of Experimental Psychology: Learning, Memory and Cognition, 22*, 1364–1382.

Levin, D. T., Takarae, Y., Miner, A., & Keil, F. C. (2001). Efficient visual search by category: Specifying the features that mark the difference between artifacts and animals in preattentive vision. *Perception and Psychophysics, 63*, 676–697.

Motter, B., & Belky, E. (1998). The guidance of eye movements during active visual search. *Vision Research, 38*, 1805–1815.

Navalpakkam, V., & Itti, L. (2005). Modeling the influence of task on attention. *Vision Research, 45*, 205–231.

Neider, M. B., & Zelinsky, G. J. (2006). Scene context guides eye movements during visual search. *Vision Research, 46*(5), 614–621.

Newell, F. N, Brown, V., & Findlay, J. M. (2004). Is object search mediated by object-based or image-based representations. *Spatial Vision, 17*, 511–541.

Oliva, A., Torralba, A., Castelhano, M., & Henderson, J. M. (2003). Top-down control of visual attention in object detection. *IEEE Proceedings of the International Conference on Image Processing (Vol. I), IEEE*, 253–256.

Parkhurst, D. J., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual selective attention. *Vision Research, 42*, 107–123.

Prinzmetal, W. (1981). Principles of feature integration in visual perception. *Perception and Psychophysics, 30*, 330–340.

Rao, R. P. N., Zelinsky, G., Hayhoe, M. M., & Ballard, D. H. (2002). Eye movements in iconic visual search. *Vision Research, 42*, 1447–1463.

Singh, M., & Hoffman, D. (2001). Part-based representations of visual shape and implications for visual cognition. In T. Shipley & P. Kellman (Eds.). *From fragments to objects: Grouping and segmentation in vision. Advances in Psychology Series* (Vol. 130, pp. 401–459). New York: Elsevier.

Thielscher, A., & Neumann, H. (2005). Neural mechanisms of human texture processing: Texture boundary detection and visual search. *Spatial Vision, 18*(2), 227–257.

Treisman, A. (1982). Perceptual grouping and attention in visual search for features and objects. *Journal of Experimental Psychology: Human Perception and Performance, 8*, 194–214.

Treisman, A. (1988). Features and objects: The Fourteenth Bartlett Memorial Lecture. *Quarterly Journal of Experimental Psychology, 40A*, 201–237.

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

Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review, 95*, 15–48.

Turano, K. A., Geruschat, D. R., & Baker, F. H. (2003). Oculomotor strategies for direction of gaze tested with a real-world activity. *Vision Research, 43*(3), 333–346.

Vecera, S., & Farah, M. J. (1994). Does visual attention select objects or locations? *Journal of Experimental Psychology: General, 123*, 146–160.

Wolfe, J. M. (1994a). Guided Search 2.0: A revised model of guided search. *Psychonomic Bulletin and Review, 1*, 202–238.

Wolfe, J. M. (1994b). Visual search in continuous, naturalistic scenes. *Vision Research, 34*(9), 1187–1195.

Wolfe, J. M. (1998). Visual search. In H. Paschler (Ed.), *Attention* (pp. 13–71). London: University College London Press.

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

Wolfe, J. M., Oliva, A., Horowitz, T. S., Butcher, S. J., & Bompas, A. (2002). Segmentation of objects from backgrounds in visual search tasks. *Vision Research, 42*, 2985–3004.

Zelinsky, G. (1996). Using eye saccades to assess the selectivity of search movements. *Vision Research, 36*, 2177–2187.

Zelinsky, G. (1999). Precueing target location in a variable set size "nonsearch" task: Dissociating search-based and interference-based explanations for set size effects. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 875–903.

Zelinsky, G. (2001). Eye movements during change detection: Implications for search constraints, memory limitations, and scanning strategies. *Perception and Psychophysics, 63*, 209–225.

Zelinsky, G. (2005). Specifying the components of attention in a visual search task. In L. Itti, G. Rees, & J. Tsotsos (Eds.), *Neurobiology of attention* (pp. 395–400). Elsevier.

Zelinsky, G., Rao, R. P. N., Hayhoe, M. M., & Ballard, D. H. (1997). Eye movements reveal the spatiotemporal dynamics of visual search. *Psychological Science, 8*, 448–453.

Zelinsky, G., & Sheinberg, D. (1997). Eye movements during parallel-serial visual search tasks. *Journal of Experimental Psychology: Human Perception and Performance, 23*, 244–262.