
Exploring the perceptual causes of search set-size effects in complex scenes

Mark B Neider[¶], Gregory J Zelinsky

Department of Psychology, Stony Brook University, Stony Brook, NY 11794-2500, USA;
e-mail: mneider@illinois.edu

Received 3 April 2009, in revised form 5 March 2010; published online 14 June 2010

Abstract. We explored search set-size effects in the context of scenes, where objects are difficult to delineate and define. Observers searched for a tank target in landscape scenes populated by boulder, shrub, and tree distractors, which varied in their color and size similarity to the target. Scenes contained either 75 or 125 distractors. In two experiments we found that adding boulder or shrub distractors that were similar to the target in color and in size produced a consistent set-size effect, but adding distractors that were similar to the target only in size did not. Tree distractors, which were larger than the target, also produced a set-size effect regardless of whether they had a target-similar color (experiment 1) or not (experiment 2). In a third experiment we varied target-tree color similarity and tree color heterogeneity and found no change in search efficiency. We interpret these data as evidence for two independent sources of set-size effects in scenes, one subject to target–distractor similarity relationships (shrubs/boulders) and the other reflecting overall visual clutter (trees), with these latter clutter effects definable by edge content.

1 Introduction

The set-size effect—how search varies with the number of objects in a display—has long been a cornerstone measure of search efficiency. Much of what we know about search set-size effects comes from experiments with simple stimuli, such as the search for a T among rotated L distractors. From such work we have learned a great deal about the features preattentively available to the visual system (Julesz 1981; Treisman and Gelade 1980; Treisman and Gormican 1988; see Wolfe 1998 for a review), and the processes that use these features to guide search to a target (Motter and Belky 1998; Wolfe 1994; Wolfe et al 1989; Zelinsky 1996, 2008). Search in the real world, however, often takes place under far messier conditions. Imagine pulling out of your driveway when your spouse asks you to run back inside to fetch keys that were left in the kitchen. Although this search task may be characterized in terms of the same feature-matching processes used to describe a T in L search, the featurally diverse and highly cluttered context of a typical kitchen in some sense makes this task qualitatively different. Do the factors known to affect set-size effects in simpler contexts also apply to more realistic searches through visually complex scenes?

The search literature has focused on two broad causes of set-size effects in the context of scenes. One factor known to affect search efficiency is the semantic content of the search scene, and the knowledge of this content that observers bring to the search task (see Henderson 2003, and Henderson and Hollingworth 1999, for reviews). Scene semantics can affect search processes in a myriad of ways, ranging from constraints imposed on the search space by scene context (eg Castelhana and Henderson 2007; Henderson et al 1999; Torralba et al 2006) to associations learned between targets and particular scene locations (eg Brockmole et al 2006; Eckstein et al 2006; Neider and Zelinsky 2006). Although the difficulties inherent in quantifying the number of objects in scenes have prevented these studies from addressing the question of set-size effects, recently Neider and Zelinsky (2008) used quasi-realistic scenes, each

[¶]Current address: Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.

depicting a countable number of complex objects, to ask how set-size effects might be influenced by semantic factors. Observers searched for a vehicle target hidden among tree distractors, which varied in number and density. They found that search efficiency improved as trees were added to the display; a reverse set-size effect. Eye-movement analyses revealed that this counterintuitive finding depended on the grouping of trees to form scene-specific regions. Observers tended to fixate individual trees when the set size was small, but fixated the open 'field' regions between clumps of trees when the set size was large. Because the number of these emergent field regions was small compared to the number of trees, search efficiency improved despite the addition of tree distractors. They concluded that real-world search scenes cannot be characterized strictly in terms of an objectively countable number of objects, and opted for a more idiosyncratic conception of set size, one in which scene constraints are used to create semantically defined regions (eg fields) through which to search.

Of course, purely perceptual factors might also contribute to search set-size effects in scenes, as they do in simpler search tasks (see Wolfe 1998 for a review). Perhaps the most widely studied of these factors is the relationship between set size and the similarity between targets and distractors. It has long been known that when a target has no features in common with the distractors, search is fast and largely independent of the number of distractors sharing the display—a phenomenon known as 'pop-out' (Treisman and Gelade 1980; Treisman and Gormican 1988). However, when targets cease to be featurally unique from distractors, search efficiency decreases, resulting in the standard set-size effect. Duncan and Humphreys (1989) characterized this effect on set size in terms of the similarity relationships between the search objects; search efficiency decreases as target–distractor similarity increases, and increases as distractors become more similar to each other. This effect of target–distractor similarity on set size is commonly attributed to a change in search guidance. As target–distractor similarity increases, the signal defining the target grows smaller, resulting in the potential for more inadvertent inspections of distractors (eg Wolfe 1994; Wolfe et al 1989; Zelinsky 2008). This relationship between target–distractor similarity and set-size effects also dovetails nicely with signal detection theory (eg Palmer et al 2000; see Verghese 2001 for a review), which assumes a matched filter to obtain a target signal and treats distractors as independent sources of noise. Increasing target–distractor similarity, or set size, creates greater overlap between the target and distractor distributions, and ultimately elevates the probability of confusing a distractor with a target.

The present study explores the role of perceptual factors in determining set-size effects in the context of scenes. Previous work relating search to scene perception has not embraced this question—a neglect that likely stems from the aforementioned difficulty in defining the scene objects relevant to search; set size cannot be manipulated if one cannot objectively count the number of objects in a fully realistic scene. Recent efforts have instead adopted a surrogate measure of set size, one that attempts to characterize perceptual influences in terms of global measures of scene clutter (eg Bravo and Farid 2004; Rosenholtz et al 2005). This approach has produced promising findings that show a general increase in search times with the degree of clutter in a scene. However, in another sense this approach sidesteps the question of how search efficiency relates to the set size of a scene, as well as the effort to understand the relationship between set size and target–distractor similarity in the context of scenes. If there is no longer the concept of an individuated distractor object, it becomes impossible to quantify how their number or similarity to the target affects search efficiency. Given that the search literature has benefited greatly from the constructs of a distractor object and a set size, we believe that it is important to attempt some understanding of scene search in terms of these basic characterizations, lest they be prematurely abandoned in favor of newer and less proven constructs.

Building on the Neider and Zelinsky (2008) study, in our approach we attempt a compromise between fully realistic scenes, which have an uncountable number of objects with unknown similarity relationships, and simpler search stimuli, which have known set-size and similarity relationships but an unclear relevance to scenes. Specifically, we use quasi-realistic search scenes created from computer-generated 3-D models. By doing this we capture some of the visual and semantic complexity characteristic of realistic scenes while preserving a degree of control over the individual search objects. We are therefore able to quantify the exact number of objects in each scene, creating a window into set-size effects as they might exist in the real world. We are also able to manipulate, at least to a first approximation, the similarity relationships between these objects. Is it the case that search in this more realistic context will show the same effects of perceptual factors identified using simpler patterns (eg Duncan and Humphreys 1989), or might the influences of individual distractors and their similarity relationships to targets be overshadowed by the high levels of clutter associated with larger set sizes?

2 Experiment 1

To determine whether set-size effects in quasi-realistic scenes depend on the visual similarity between a target and distractors, we had observers search for a vehicle target (a tank) amidst shrub, boulder, and tree distractors, all embedded in a rendered wooded landscape. Our set-size manipulation consisted of the explicit addition of objects to the scene—something that would be difficult to achieve in fully realistic scenes without altering the overall context. We manipulated target–distractor similarity by having the target share different features with different distractors. Trees were similar to the target in color, but not in size; boulders were similar to the target in size, but not in color; shrubs were similar to the target in both color and size. We also manipulated the proportion of each distractor type in a scene, so as to study the interaction between set size and target–distractor similarity (also see Shen et al 2000). If search is guided to target-like objects in a scene (eg Wolfe 1994; Zelinsky 2008), then set-size effects should be steepest when the majority of distractors are shrubs, and relatively shallow when the majority of distractors are trees or boulders. Alternatively, if perceptual similarity effects between the target and distractors are trumped by the overall noise introduced by growing clutter, then no set-size differences should be found between these three conditions.

2.1 Methods

2.1.1 Participants. Twelve 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. Search displays subtended $27 \text{ deg} \times 20 \text{ deg}$ (1280×960 pixels) and were presented in color on a ViewSonic 19 inch 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 controlled the stimulus presentation. Head position and viewing distance were fixed with a chin-rest, and all responses were made with a gamepad controller attached to the USB port of the computer. Judgments were made with the left and right index-finger triggers; trials were initiated with a button operated by the left thumb.

Search displays were created with Autodesk's 3D Studio Max software and depicted a quasi-realistic landscape scene under normal daytime lighting, as viewed from a slightly elevated perspective. The target in all trials was a modern military tank (shown in close-up in figure 1a), which subtended approximately 0.65 deg visual angle. Distractors were a combination of shrubs ($\sim 0.55 \text{ deg}$), boulders ($\sim 0.55 \text{ deg}$), and trees ($\sim 2 \text{ deg} \times 4.4 \text{ deg}$). Shrubs and boulders were similar to the target in size; shrubs and



Figure 1. [In colour online, see <http://dx.doi.org/10.1068/p6428>] Sample target images from (a) experiment 1 and (b) experiment 2. Images are shown enlarged relative to their appearance in the search scenes.

trees were similar to the target in color (both were shades of green with patches of gray). The target never appeared within 2 deg of the center of the display, was never occluded, and was oriented so as to appear from a side view in the search image. The target was also rotated around the y axis on a trial-by-trial basis ($\pm \sim 0-20^\circ$) in order to account for changes in ground slope and to keep scenes looking as realistic as possible, although this rotation was kept to a minimum. Except for these constraints, targets and distractors were randomly placed on the ground surface. Sample search scenes are shown in figure 2.

2.1.3 Design and procedure. There were a total of eighty experimental trials, distributed equally over all conditions. Set size was varied across two levels: 75 distractors and 125 distractors. In the 75 set-size scenes, shrub, tree, and boulder distractors appeared in equal proportions (25/25/25). We refer to this as the EQ condition. In the 125 set-size scenes, one distractor type appeared 75 times in each scene, with the other distractor types each appearing 25 times (75/25/25). We refer to these as the high shrub (HS), high boulder (HB), and high tree (HT) conditions, with the ‘high’ designation referring to the favored distractor. The target was present on half of the trials; the other half were target-absent. All manipulations were randomly interleaved throughout the experiment.

To designate the target object, participants were shown an image of the tank target (~ 8.34 deg, side view, black background) for 10 s at the start of the experiment. There was no target preview prior to each trial. Trials began with the observer fixating a central fixation target and pressing the ‘X’ button on the gamepad to initiate the onset of the search scene. This scene remained visible until the observer’s response. Their task was to indicate the presence or absence of the target by pressing the left or right triggers, respectively, and to do so as quickly as possible while maintaining accuracy. Accuracy feedback was provided after each judgment. On target-present (TP) trials, a green box was drawn around the target object indicating its location in the search display. The words “No Target” were drawn to the screen on target-absent (TA) trials. Both feedback displays were visible for 1 s, followed immediately by the fixation target signaling the start of the next trial. There were eight practice trials, and the entire experiment lasted approximately 45 min.

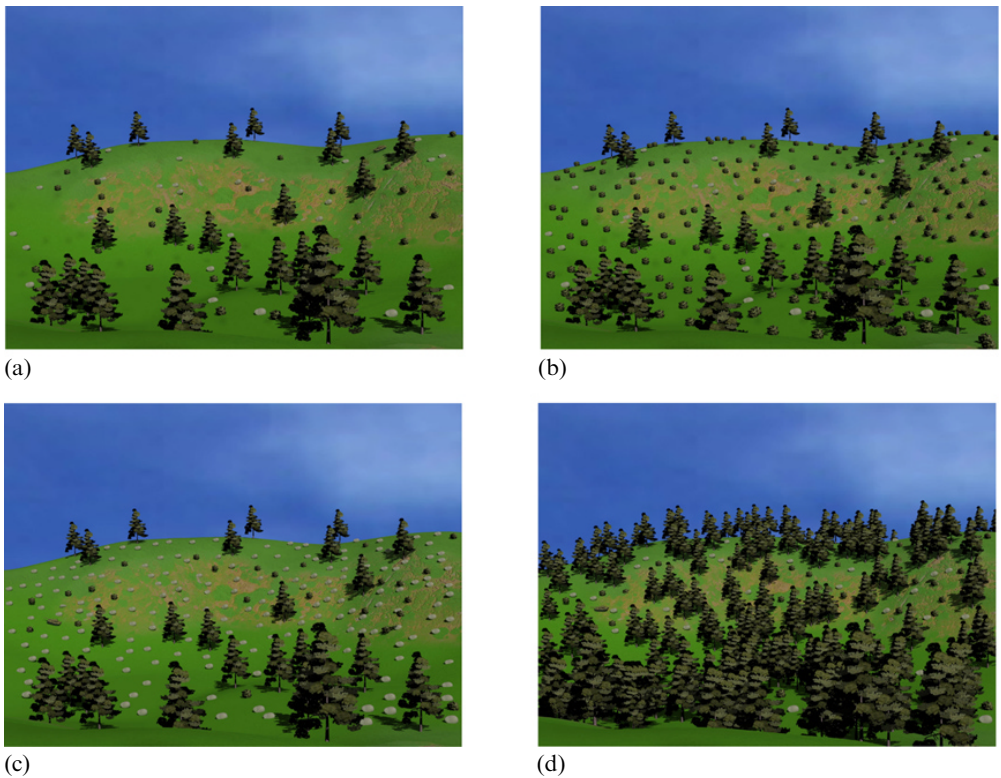


Figure 2. [In colour online.] Sample images from experiment 1 in the (a) EQ, (b) HS, (c) HB, and (d) HT conditions. Note: EQ = equal; HS = high shrub; HB = high boulder; HT = high tree.

2.2 Results and discussion

Error rates were generally low, but were higher in TP (4.4%) than in TA (1%) ($F_{1,11} = 11.31$, $p < 0.01$) trials. No significant effects of set size or distractor-group composition were found ($F_{3,33} \leq 0.72$, $p > 0.3$). If observers found some distractors to be more target-like than others, this did not manifest itself in increased error rates.

Figure 3 shows manual reaction times (RTs) for correct trials, plotted as a function of distractor-group composition and target presence. If search is guided to target-like distractors, then search efficiency should decrease as more of those items are added to the scene. Consistent with this prediction, we found a significant main effect of distractor-group composition in both TP ($F_{3,33} = 8.85$, $p < 0.001$) and TA ($F_{3,33} = 10.06$, $p < 0.001$) trials. This effect was largely driven by differences between the EQ condition and the HS and HT conditions. Increasing the number of shrub (HS) or tree (HT) distractors resulted in longer search times than in the EQ condition in both TP (615 ms) ($t_{11} > 4.44$, $p < 0.005$) and TA (1325 ms) ($t_{11} > 3.34$, $p < 0.01$) trials. These are clear expressions of a set-size effect; adding distractors to the scene degraded search performance. The evidence for a set-size effect in the HB condition was mixed. Although adding boulders to TP scenes increased RTs (by 290 ms) relative to the EQ condition ($t_{11} = 5.71$, $p < 0.001$), no set-size effect was found in the TA trials. Indeed, adding boulders to TA scenes actually produced a small RT benefit (87 ms) compared to the EQ condition, although this difference was not significant ($t_{11} = 0.32$, $p = 0.75$). Overall, increasing the number of objects in our scenes degraded search efficiency when the distractors were shrubs or trees; increasing the number of boulders produced a smaller set-size effect, or none at all.

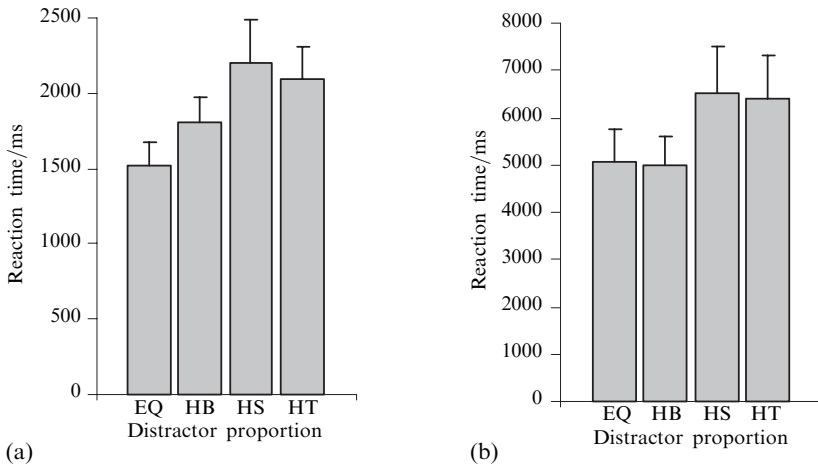


Figure 3. Mean reaction times as a function of distractor type proportion and target presence for experiment 1. (a) Target present; (b) target absent. EQ = equal; HB = high boulder; HS = high shrub; HT = high tree. Error bars indicate 1 SEM.

Our finding of set-size effects in the HS and HT conditions, and the absence of a robust set-size effect in the HB condition, generally supports a mechanism that guides search to target-similar distractors (Wolfe 1994; Zelinsky 2008). In the context of scenes, in which there is no expectation that all of the objects will be inspected, the functional consequence of such a mechanism is the perceptual segregation of objects into search-relevant and search-irrelevant groups (Neider and Zelinsky 2008). Boulders were similar to the target in size but highly dissimilar in color, resulting in their relative omission from the relevant search set and the near absence of a set-size effect. Shrubs, however, were similar to the target in terms of both size and color. This relatively high level of target–distractor similarity made these objects very relevant to the search task, as evidenced by a large set-size effect.

Inconsistent with this straightforward relationship between target–distractor similarity and search efficiency is the set-size effect observed for tree distractors. Given that trees were similar to the target in color but not size, the tree set-size effect should have been comparable to the one observed in the HB condition, and smaller than the set-size effect observed in the HS condition. This clearly was not the case. Comparable set-size effects were found regardless of whether shrubs or trees were added to the scene (TP condition: $t_{11} = 0.57$, $p = 0.58$; TA condition: $t_{11} = 0.35$, $p = 0.73$). This suggests that trees and shrubs were equally distracting in this task, despite their dramatically different visual similarity relationships to the target.

Might the source of the tree set-size effect be due to increasing clutter rather than to target–distractor similarity? To explore this possibility we analyzed visual clutter for each of our scene types. Clutter was estimated by counting the edges in each image with the Canny edge-detection method (Canny 1986). Edge content was demonstrated in recent work to be a relatively simple and reliable predictor of clutter effects in a search task (Henderson et al 2009). Performing this analysis for the EQ condition yielded 19 944 edges. Edge counts were higher but roughly the same in the HS (26 601) and HB (24 915) conditions—a finding that we anticipated because of the comparable size and shape of the shrub and boulder objects. However, edge count was substantially higher in the HT condition (46 932). The trees were much larger than either the shrubs or boulders, and had many jagged edges. Both of these factors contributed to the higher edge estimates for this distractor type. This difference in edge count, and by extension visual clutter, suggests that the comparable set-size effects observed for shrubs and trees

may have been produced by different underlying processes. In the case of shrub distractors, steep set-size effects may have been driven by target–distractor similarity, and not visual clutter; in the case of tree distractors, the set-size effect is better explained by clutter than target–distractor similarity.

3 Experiment 2

The data from experiment 1 suggest that both color similarity to the target and visual clutter may be important factors in determining the objects that are included in the relevant search set, and the consequent expression of a set-size effect. However, the relative contributions of these two factors remain unclear. Adding trees to the search scene increased clutter and decreased search efficiency, but each of these tree distractors also shared the target's color. Although shrubs were also similar to the target in size and shape, if target–distractor color similarity trumps these other dimensions, the comparable tree and shrub set-size effects may have been caused by color similarity alone.

We conducted experiment 2 to clarify the role of target–distractor color similarity in producing the shrub and tree set-size effects from experiment 1. We did this by changing the color of the tank target to match the boulder distractors rather than the shrubs and trees. If color similarity to the target was responsible for both the shrub and the tree set-size effects from experiment 1, this manipulation should produce a relatively steep set-size effect for boulders (HB condition), and a shallower set-size effect for both shrubs and trees (HS and HT conditions). However, if clutter, and not target–distractor color similarity, was responsible for the previously observed tree set-size effect, then we would expect no effect of this manipulation—the set-size effect for trees should remain steep, perhaps as steep as the set-size effect that we now expect to find for boulders.

3.1 Methods

There were twelve observers, none of whom participated in experiment 1. The stimuli and methodology were identical to the descriptions provided in experiment 1, except for the color of the target in the preview and search scenes, which was now gray to match the boulder distractors (figure 1b).

3.2 Results and discussion

Error rates averaged < 3% in the TP trials and < 2% in the TA trials, with no significant differences. Changing the color of the target did not affect the pattern of errors.

RTs for correct trials are shown in figure 4. As in experiment 1, set-size effects were found regardless of how the distractors are apportioned (HB, HS, or HT); adding objects to the scene decreased search efficiency relative to the EQ condition (TP condition: $t_{11} > 2.88$, $p < 0.05$; TA condition: $t_{11} > 2.34$, $p < 0.05$). Also, as in experiment 1, these set-size effects depended on target–distractor color similarity, although the specific effects were now in an opposite direction. There was now a small set-size effect for shrubs in the TP data, and a significantly larger set-size effect for boulders in the HB condition ($t_{11} = 2.99$, $p < 0.05$). A similar, albeit noisier, pattern appeared in the TA data. The set-size effect in the HB condition was significant ($t_{11} = 2.34$, $p < 0.05$); the set-size effect in the HS condition was not ($t_{11} = 1.67$, $p = 0.12$).⁽¹⁾ For both target conditions, increasing the number of target-similar boulders in these scenes was more disruptive to search than increasing the number of target-dissimilar shrubs.

Critically, we also found a pronounced set-size effect for trees ($t_{11} = 2.51$, $p < 0.05$), which was comparable in size to the set-size effect from the target-similar distractor type, in this case boulders ($t_{11} = 1.85$, $p > 0.05$). This pattern lends support to the suggestion

⁽¹⁾ Although from figure 4 the contrast between EQ and HS appears significant, the mean shrub RT was inflated by the data from one of the twelve observers.

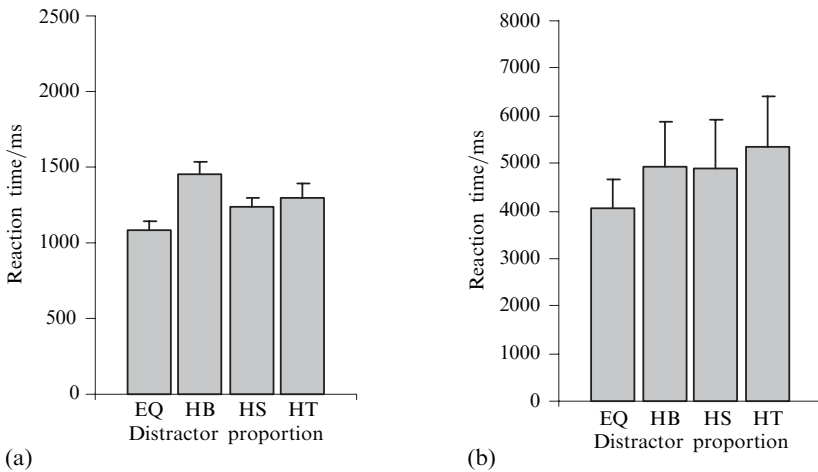


Figure 4. Mean reaction times as a function of distractor type proportion and target presence for experiment 2. (a) Target present; (b) target absent. Note: EQ = equal; HB = high boulder; HS = high shrub; HT = high tree. Error bars indicate 1 SEM.

that the tree set-size effects observed in experiments 1 and 2 were due to the large difference in visual clutter between the EQ and HT conditions. Adding tree distractors to the scene, regardless of their color similarity to the target, increased visual clutter and in turn produced a large set-size effect.

Taken together, the findings from experiments 1 and 2 are consistent in suggesting that target–distractor color similarity indeed contributes strongly to set-size effects in scenes, probably by ensuring that target-similar objects are well represented in the relevant search set. The shrubs in experiment 1 were perceptually most similar to the target, so the targeted inspection of these objects produced large set-size effects (393 ms in TP trials and 1442 ms in TA trials). Likewise, large set-size effects for boulders were found in experiment 2 (365 ms in TP trials and 855 ms in TA trials) as a result of these objects now being most similar in color to the target. The fact that these dramatically different data patterns stemmed from such a minor manipulation (the handful of pixels determining the color of the target) highlights the clear importance of target–distractor color similarity in producing set-size effects in scenes. Equally clear, however, is the fact that the tree set-size effects were not modulated by target–distractor color similarity; search efficiency decreased with the addition of green-colored trees regardless of the color of the target. Our analysis of edge count suggests that visual clutter may instead be the causal factor in producing these three set-size effects.

Also interesting is the fact that set-size effects were generally smaller, and search times faster overall, in experiment 2 than experiment 1. We speculate that this too may be related to target–distractor color similarity, as there were more target-colored objects in experiment 1 (trees and shrubs) than in experiment 2 (only boulders). If true, this may suggest cross-talk between the tree and shrub distractors in experiment 1; although color similarity to the target cannot explain the set-size effects observed for trees, the green color of the trees may have introduced noise into the search process, making it harder to confine search to the shrub distractors in experiment 1.

4 Experiment 3

The finding that tree distractors in our task were largely immune from effects of target color similarity is surprising. Not only has the search literature been consistent in showing that color is one of the most influential features in determining target–distractor similarity effects on search (eg Motter and Belky 1998; Rutishauser and Koch 2007;

Williams 1967), our own experiments 1 and 2 also showed clear relationships between target color and set-size effects for shrubs and boulders in the context of scenes. But is it really the case that target–distractor color similarity is completely ineffective in modulating these tree set-size effects? Alternatively, perhaps the large differences in tree clutter between the EQ and HT conditions were simply swamping any smaller effects of color similarity between the trees and the target. In this experiment we address this possibility by manipulating target–tree color similarity, and tree color heterogeneity, while holding set size (and overall visual clutter) constant at 75 distractors—the same number of trees used in the HT conditions from experiments 1 and 2. In one condition, all of the trees were the target color; in a second condition the trees were a non-target color; and in a third condition the trees were a mixture of colors. If search efficiency is affected by target color similarity in the absence of clutter differences, search times should be longest when all the trees are the same color as the target, and shortest when the trees are all a non-target color. An intermediate result would be expected when trees are mixed in color, although the greater feature variability accompanying this condition might also serve to decrease such efficiency (Duncan and Humphreys 1989). Of course, if target color similarity is indeed ineffective in the case of the tree distractors, no differences would be expected between these color conditions.

4.1 Methods

Twelve observers, none of whom participated in experiments 1 or 2, indicated the presence or absence of a tank target in computer-generated landscape scenes. The target was greenish in color and identical to the target used in experiment 1. The search scenes were similar to those described in the previous experiments, except that they now depicted only tree distractors, and there were 75 trees in each scene. A different type of tree was also used in this experiment, one that permitted a more compelling color manipulation. Target–distractor color similarity was manipulated in three within-subjects conditions. In the homogeneous–similar (HOS) condition all of the trees were similar in color to the target. In the homogeneous–different (HOD) condition the trees were all similar to each other but appeared in a non-target color (reddish). In the heterogeneous (HET) condition one-third of the trees appeared in the target color, another one-third appeared in the non-target color from the HOD condition, and the remaining one-third appeared in a new non-target color (yellowish). Figure 5 shows representative samples of search scenes in each color condition. There were ninety experimental trials, which were evenly divided into the three color and present/absent conditions, and six practice trials. Conditions were randomly interleaved throughout the experiment, which lasted approximately 45 min. All other details regarding the stimuli, design, and procedure were identical to experiments 1 and 2.

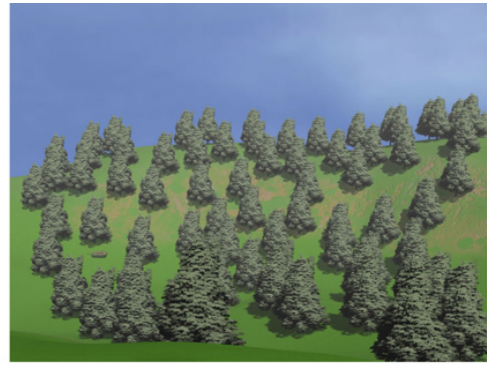
4.2 Results and discussion

As in the previous experiments, error rates were low, averaging approximately 5% in TP trials and 1% in TA trials, with no significant differences. These error trials were excluded from all subsequent analyses.

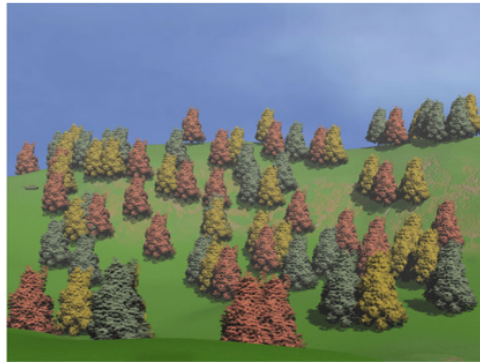
Mean RTs are shown in figure 6 as a function of target presence and the color composition of the tree distractors. Clearly, the color manipulations had no effect on search performance in either TP ($F_{2,22} = 0.31, p = 0.737$) or TA ($F_{2,22} = 1.18, p = 0.325$) trials. Observers took roughly the same amount of time to make their judgments regardless of whether the trees shared the target color (HOS), were a highly discriminable non-target color (HOD), or varied in color (HET). The fact that no differences were found between these very different color conditions is striking, and suggests that the traditional rules governing the relationship between perceptual similarity and search set-size effects do not apply to the tree distractors in our task. Contrary to theories of



(a)

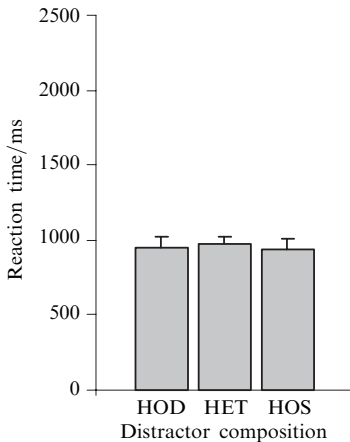


(b)

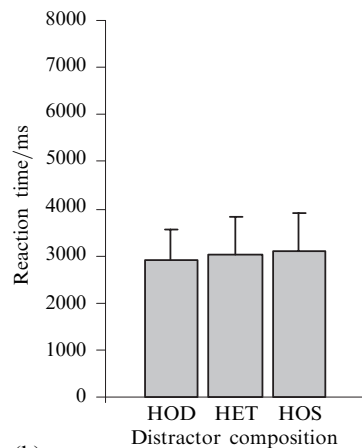


(c)

Figure 5. [In colour online.] Sample images from experiment 3 in the (a) homogeneous–different, (b) homogeneous–similar, and (c) heterogeneous conditions.



(a)



(b)

Figure 6. Mean reaction times as a function of distractor condition and target presence for experiment 3: (a) target present; (b) target absent conditions. Note: HOD = homogeneous–different; HET = heterogeneous; HOS = homogeneous–similar. Error bars indicate 1 SEM.

guided search (Wolfe 1994; Zelinsky 2008), increasing target–distractor color similarity did not make the search task more difficult, nor did increasing the color variability within the distractor group (Duncan and Humphreys 1989). This pattern also provides converging evidence for our assertion that the tree set-size effects observed in experiments 1 and 2 were caused by a second factor, one independent of the similarity relationships demonstrated to exist for shrubs and boulders. Specifically, it is not the case that color similarity relationships between the target and the trees were actually

impacting search efficiency in these tasks, but that these effects were being masked by the set-size effect arising from another source; even when all other sources of variability were held constant, target-color similarity still failed to meaningfully affect search efficiency.

In addition to ruling out a role of target–distractor color similarity, the results of experiment 3 provide indirect support for our earlier claim that the tree set-size effects reported in experiments 1 and 2 were due to changes in visual clutter. Given that set size was held constant in experiment 3, we did not expect systematic differences in clutter, as measured by edge count, to exist between the HOD, HET, and HOS color conditions; edge count should vary with the number of objects in the scene, not the color of the objects. We verified this expectation by again computing edge-count estimates for each of the scenes in all three of the color composition conditions. As expected, edge counts were roughly equivalent across all three conditions (HOD = 46 096; HET = 40 917; HOS = 47 086).⁽²⁾ Importantly, none of the observed differences even approached the very large difference found between the EQ and HT conditions from experiment 1. Taken together, the edge analyses from experiments 1 and 3 strongly suggest a relationship between visual clutter and the observation of set-size effects for tree distractors. When tree clutter increased between the EQ and HS/HB conditions from experiments 1 and 2, RTs increased as well, regardless of whether the tree color matched the target. However, when tree clutter was held roughly constant across the target–distractor similarity manipulations in experiment 3, search times were largely unaffected. We therefore conclude that there exist two largely independent sources of set-size effects in the context of scenes, with one source being the degree of visual similarity between the target and distractor (as in the case of our shrub and boulder objects) and the other being the degree of visual clutter that a distractor introduces into the scene (as in the case of our tree objects).

5 General discussion

Our objective here was to better characterize the role of perceptual factors in determining set-size effects in quasi-realistic scenes. We did this by manipulating the color and size similarity between targets and distractors, and the proportion of each type of feature in a large but quantifiable set of search objects. The relationships that emerged from this study both simplify and complicate our understanding of set-size effects in this more realistic context.

Turning first to the simplifying finding, much of what was already known about the relationship between set-size effects and perceptual factors from signal-detection theory (eg Palmer et al 2000; Vergheze 2001) also applies to set-size effects in scenes. This is particularly true for the dominant role that target–distractor color similarity is known to play in guiding search (eg Motter and Belky 1998; Rutishauser and Koch 2007; Shen et al 2000; Williams 1967). Consistent with these earlier studies we found large set-size effects for shrubs and boulders when these distractors shared the target's color. However, when these distractors were similar to the target in only size and shape, these set-size effects were dramatically reduced, or disappeared entirely. We interpret these patterns as evidence for color being a key determinant for inclusion in a relevant set size (Neider and Zelinsky 2008); when there are too many objects through which to search, we rely on color to reduce the size of this search space. As for our failure to find any evidence for a role of distractor size in guiding search, we speculate that this may be due to the scene stimuli used in our study. An object will change size depending on its depth in the scene, thereby making this feature a less reliable cue for search guidance (Aks and Enns 1996).

⁽²⁾The slightly lower estimates obtained in the HET condition resulted from our method occasionally failing to find edges between the yellow trees and the yellow-green background.

Our findings are also theoretically important in that they suggest that only a small number of features are used to guide search. Real-world objects are visually complex, meaning that the potential exists for their representation to require a very high-dimensional feature space. Indeed, recent models of change detection (Zelinsky 2003), and of bottom-up (eg Itti and Koch 2001; Parkhurst et al 2002) and top-down (eg Rao et al 2002; Zelinsky 2008) visual search have assumed such high-dimensional representations, with objects coded in terms of multiple spatial scales, orientations, and colors. Our findings inform this practice by suggesting that such high-dimensional feature representations may be unnecessary, at least for the search task used in the current study. Alternatively, target objects may indeed be coded in terms of many features, but in practice search may be guided by a small number of dominant features within this larger feature set. Object color appears to be one of these dominant features (also see Hwang et al 2009).

However, our study also complicates the relationship between search set size and scenes by identifying a second process that creates set-size effects independently of target-distractor similarity. Large set-size effects for trees were found in experiments 1 and 2 regardless of the distractor's similarity to the target. From experiment 3 we also found that varying target-tree color similarity, and the color heterogeneity of these distractors, had no meaningful effect on search efficiency, further arguing for a dissociation between tree color and the observed set-size effects. This immunity from color manipulations is inconsistent with theories of search guidance that rely on target-distractor similarity relationships to explain set-size effects (eg Duncan and Humphreys 1989; Wolfe 1994; Zelinsky 2008). Just as these theories were supported by our observations of color-specific set-size effects in the case of shrubs and boulders, these theories would predict that tree set-size effects should occur only when these distractors are perceptually similar to the target, which was clearly not the case. More broadly, the observed persistence of a tree set-size effect is inconsistent with all signal-detection conceptions of search; a gray tank target is not likely to be confused with a green or red tree, so why would the addition of these non-target-colored tree objects degrade search efficiency so dramatically?

We believe that visual clutter is responsible for the set-size effects observed for trees in this study. To support this assertion we showed that changes in edge count, a reliable predictor of clutter effects in scenes (Henderson et al 2009), was large when a tree set-size effect was observed (experiments 1 and 2) and small when tree set size was held constant (experiment 3). This finding is theoretically important in at least two respects. First, the identification of two independent sources of set-size effects means that theories of search designed to work with scenes must acknowledge the contribution of both sources. Search theories must not focus exclusively on target-distractor similarity relationships to explain set-size effects, as this will neglect the contribution of visual clutter. Likewise, search theories cannot describe set-size effects purely in terms of visual clutter or feature congestion, as this will neglect the role of target-distractor similarity relationships. Second, given that some objects exert target-distractor similarity effects and others exert only clutter effects, it follows that *objects* remain an important construct in understanding set-size effects in scenes. Approaches that seek to redefine set-size effects in terms of visual clutter do not deal with individuated objects; objects are dissolved into features, and the congestion among these features is used to estimate clutter (eg Rosenholtz et al 2005). The present data suggest that this approach may be overly simplistic. Objects in scenes sometime generate only clutter (our trees) and at other times are processed to the level of target-distractor similarity relationships (our shrubs and boulders); a search theory should be able to predict which objects will produce which set-size effects, and explain why.

What sort of theoretical framework would be needed to integrate these two sources of set-size effects? Following Neider and Zelinsky (2008), we believe that the answer to this question requires the specification of a relevant search set, one that is determined by the comparison of individual object features to the search target. In the context of our present study, we speculate that shrub, boulder, and tree patterns were segmented by low-level visual processes into perceptual objects, with each object represented as a collection of its component visual features. These object representations are then compared to the search target feature by feature. If this comparison reveals that any one object feature is sufficiently different from the target, as in the case of tree distractors differing substantially from the target in size, this object is excluded from the search set. However, if all of an object's features fail to reach this exclusion threshold, as was the case for shrubs in experiment 1 and boulders in experiment 2, these objects will be included in the set of objects to be considered for inspection during search. According to this object-based feature-threshold framework, objects included in the search set would then be ranked in terms of their visual similarity to the target and prioritized for inspection as described by theories of guided search (eg Wolfe 1994; Zelinsky 2008). Those objects that do not make it into the search set will not undergo this target-similarity prioritization because they, along with their features, would have been excluded as viable candidates for search on the basis of their single-feature target dissimilarity. Still, search efficiency might degrade with the addition of these objects to the scene as a result of the visual noise introduced by the object's features, with the cost of each object to search proportional to the degree of clutter that it introduces.

Such an object-based feature-threshold framework would explain why shrubs and boulders in our task were modulated by target similarity relationships, and why trees were largely immune to these relationships, yet still produced substantial set-size effects. Future work will attempt to use eye-movement behavior to validate the division of scene distractors into these two qualitatively different groups, with the expectation being that objects in the search set (shrubs or boulders) will attract a disproportionate number of fixations compared to objects excluded from this set (trees).⁽³⁾ Another important direction for future work will be to explore the relationship between clutter, as measured by edge counts, and low-level visual salience, as measured by feature contrast (eg Itti and Koch 2001), as they relate to set-size effects in scenes. We argued that clutter was responsible for the tree set-size effects reported in this study, but might it be the case that trees were simply large and therefore too perceptually salient to completely ignore? The data from experiment 3 suggest that clutter, and not salience, is the more likely cause of these set-size effects; the red trees against the green background in the HOD condition should have created greater feature contrast than the green trees against the green background in the HOS condition, yet this difference in salience did not affect search efficiency. Still, these possibilities should be more systematically dissociated so as to better characterize the purely bottom-up contribution to set-size effects in scenes.

Acknowledgment. This work was supported by the National Institute of Health under grant R01 MH063748 to GJZ.

⁽³⁾Note that such a demonstration would also have implications for current conceptualizations of scene search as a mixture of bottom-up and top-down processes (eg Ehinger et al 2009; Peters and Itti 2007; Zelinsky et al 2006). Rather than assuming that search is guided by a combination of bottom-up and top-down guidance signals, it may be that these signals remain largely independent and segregated by the objects in a scene. Distractors might therefore be usefully characterized as belonging to one of two groups—those that exert their distracting influence through the addition of clutter (a bottom-up factor) and those that are distracting because of their perceptual similarity to the target (a top-down factor).

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
- Bravo M J, Farid H, 2004 "Search for a category target in clutter" *Perception* **33** 643–652
- Brockmole J R, Castelhana M S, Henderson J M, 2006 "Contextual cueing in naturalistic scenes: Global and local contexts" *Journal of Experimental Psychology: Learning, Memory, and Cognition* **32** 699–706
- Canny J, 1986 "A computational approach to edge detection" *IEEE Transactions on Pattern Analysis and Machine Intelligence* **8** 679–714
- Castelhana M S, Henderson J M, 2007 "Initial scene representations facilitate eye movement guidance in visual search" *Journal of Experimental Psychology: Human Perception and Performance* **33** 753–763
- Duncan J, Humphreys G, 1989 "Visual search and stimulus similarity" *Psychological Review* **96** 433–458
- Eckstein M P, Drescher B, Shimozaki S S, 2006 "Attentional cues in real scenes, saccadic targeting and Bayesian priors" *Psychological Science* **17** 973–980
- Ehringer K A, Hidalgo-Sotelo B, Torralba A, Oliva A, 2009 "Modeling search for people in 900 scenes: A combined source model of eye guidance" *Visual Cognition* **17** 945–978
- Henderson J M, 2003 "Human gaze control in real-world scene perception" *Trends in Cognitive Sciences* **7** 498–504
- Henderson J M, Chanceaux M, Smith T J, 2009 "The influence of clutter on real-world scene search: Evidence from search efficiency and eye movements" *Journal of Vision* **9** 1–8
- 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
- Hwang A D, Higgins E C, Pomplun M, 2009 "A model of top-down attentional control during visual search in complex scenes" *Journal of Vision* **9**(5):25, 1–18, <http://journalofvision.org/9/5/25/>, doi:10.1167/9.5.25
- Itti L, Koch C, 2001 "Computational modeling of visual attention" *Nature Reviews Neuroscience* **2** 194–203
- Julesz B, 1981 "Textons, the elements of texture perception and their interactions" *Nature* **290** 91–97
- Motter B, Belky E, 1998 "The guidance of eye movements during active visual search" *Vision Research* **38** 1805–1815
- Neider M B, Zelinsky G J, 2006 "Scene context guides eye movements during visual search" *Vision Research* **46** 614–621
- Neider M B, Zelinsky G J, 2008 "Exploring set size effects in realistic scenes: Identifying the objects of search" *Visual Cognition* **16**(1) 1–10
- Palmer J, Verghese P, Pavel M, 2000 "The psychophysics of visual search" *Vision Research* **40** 1227–1268
- 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
- Peters R J, Itti L, 2007 "Beyond bottom-up: Incorporating task-dependent influences into a computational model of spatial attention" *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2007)* pp 1–8
- Rao R, Zelinsky G J, Hayhoe M, Ballard D, 2002 "Eye movements in iconic visual search" *Vision Research* **42** 1447–1463
- Rosenholtz R, Li Y, Mansfield J, Jin Z, 2005 "Feature congestion: a measure of display clutter" *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (SIGCHI 2005)* 761–770
- Rutishauser U, Koch C, 2007 "Probabilistic modeling of eye movement data during conjunction search via feature-based attention" *Journal of Vision* **7**(6) 1–20
- Shen J, Reingold E M, Pomplun M, 2000 "Distractor ratio influences pattern of eye movements during visual search" *Perception* **29** 241–250
- Torralba A, Oliva A, Castelhana M, Henderson J M, 2006 "Contextual guidance of attention in natural scenes: The role of global features on object search" *Psychological Review* **113** 766–786
- 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

-
- Verghese P, 2001 “Visual search and attention: a signal detection theory approach” *Neuron* **31** 523–535
- Williams L G, 1967 “The effects of target specification on objects fixated during visual search” *Acta Psychologica* **27** 355–360
- Wolfe J M, 1994 “Guided Search 2.0: A revised model of guided search” *Psychonomic Bulletin & Review* **1** 202–238
- Wolfe J M, 1998 “Visual search”, in *Attention* Ed. H Paschler (London: University College London Press) pp 13–71
- Wolfe J M, 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
- Zelinsky G J, 1996 “Using eye saccades to assess the selectivity of search movements” *Vision Research* **36** 2177–2187
- Zelinsky G J, 2003 “Detecting changes between real-world objects using spatio-chromatic filters” *Psychonomic Bulletin and Review* **10** 533–555
- Zelinsky G J, 2008 “A theory of eye movements during target acquisition” *Psychological Review* **115** 787–835
- Zelinsky G J, Zhang W, Yu B, Chen X, Samaras D, 2006 “The role of top–down and bottom–up processes in guiding eye movements during visual search”, in *Advances in Neural Information Processing Systems (NIPS 2005)* volume 18, Eds Y Weiss, B Scholkopf, J Platt (Cambridge, MA: MIT Press) pp 1569–1576

ISSN 0301-0066 (print)

ISSN 1468-4233 (electronic)

PERCEPTION

VOLUME 39 2010

www.perceptionweb.com

Conditions of use. This article may be downloaded from the Perception website for personal research by members of subscribing organisations. Authors are entitled to distribute their own article (in printed form or by e-mail) to up to 50 people. This PDF may not be placed on any website (or other online distribution system) without permission of the publisher.