Available online at www.sciencedirect.com

SCIENCE DIRECT®

Vision Research 46 (2006) 614–621

Vision Research

www.elsevier.com/locate/visres

Brief communication

Scene context guides eye movements during visual search

Mark B. Neider, Gregory J. Zelinsky *

Department of Psychology, State University of New York at Stony Brook, USA

Received 10 March 2005; received in revised form 8 August 2005

Abstract

How does scene context guide search behavior to likely target locations? We had observers search for scene-constrained and scene-unconstrained targets, and found that scene-constrained targets were detected faster and with fewer eye movements. Observers also directed more initial saccades to target-consistent scene regions and devoted more time to searching these regions. However, final checking fixations on target-inconsistent regions were common in target-absent trials, suggesting that scene context does not strictly confine search to likely target locations. We interpret these data as evidence for a rapid top-down biasing of search behavior by scene context to the target-consistent regions of a scene.

Keywords: Scene context; Contextual cueing; Visual search; Top-down guidance; Scene perception

1. Introduction

Real-world search tasks are performed in contextually rich environments that offer numerous high-level cues for likely target location. This is why, when searching for your car in a crowded parking lot, you are likely to confine your search to the ground level and not look to the sky or at your feet. Such scene-based influences have long been assumed in the scene perception literature (Biederman, Mezzanotte, & Rabinowitz, 1982; Yarbus, 1967), but the search literature has focused instead on visual feature guidance (Wolfe, 1994) rather than on guidance from scene context. Except for one recent demonstration of observers distributing their gaze on footpaths and doorways when looking for people (Oliva, Torralba, Castelhano, & Henderson, 2003), the effects of scene context on search remain largely unexplored.

Related work has defined context in terms of the randomly generated configurations of items in displays, and has shown that this information can be used to facilitate search. Chun and Jiang (1998) found that targets are detected more efficiently when they appear in previously viewed spatial configurations compared to new configurations, a benefit they referred to as "contextual cueing." More recent studies using the contextual cueing paradigm have shown that this benefit extends to dynamic display configurations (Chun & Jiang, 1999) and target identification (Endo & Takeda, 2004). Peterson and Kramer (2001) even documented the use of contextual cues to direct eye movements in a search task. However, although all of these studies converge on the fact that simple geometric configurations can be implicitly learned and used to improve search efficiency, one cannot generalize this finding to the use of scene-based constraints during search. Scene constraints are grounded in semantic knowledge of the world (e.g., knowing that parked cars do not float in the air), and as such are likely to be represented and accessed differently than implicitly learned configural information.

There have been studies looking at the effects of scene constraints on search (see Henderson & Hollingworth, 1999, for a review), but these studies have typically manipulated the semantic consistency of a target in a scene. For example, Henderson, Weeks, and Hollingworth (1999) showed observers line drawings of scenes and asked them

---

* Corresponding author. Tel.: +1 631 632 7827; fax: +1 631 632 7876.
E-mail address: Gregory.Zelinsky@sunysb.edu (G.J. Zelinsky).

0042-6989/$ - see front matter © 2005 Published by Elsevier Ltd.
doi:10.1016/j.visres.2005.08.025

to search for either a scene-consistent or a scene-inconsistent target. They found that scene-consistent targets were detected faster and fixated sooner relative to scene-inconsistent targets. Although their study focused on the looking behavior to scene-inconsistent targets, Henderson and colleagues speculated that this benefit might be due to observers using their semantic scene knowledge to guide their eye gaze to likely display locations when searching for scene-consistent targets.

2. Experiment 1

Building on the contextual cueing and scene-consistency work, we used only scene-consistent targets, but manipulated whether the scene context constrained the likely locations of these targets in the search display. Targets were either scene-constrained (SC), meaning that the scene provided information as to the target's general location, or scene-unconstrained (SU), meaning that the target could appear almost anywhere in the scene. For SC targets and a given scene, there were also target-consistent regions (regions that might contain the target) and target-inconsistent regions (regions where the target would never appear).

Three questions were addressed in this study. First, how much faster can SC targets be found relative to SU targets? Although our introspections tell us that we use scene context in almost all of our day-to-day search tasks, the topic has never been the focus of an experimental study, therefore basic questions regarding scene context and search remain unanswered. Second, how tightly is search constrained by scene context? One possibility is that it is strictly confined to the target-consistent region, a possibility reinforced by our introspections (e.g., we never look for a parked car in the sky). If this were the case, we would expect a preferential pattern of fixations to target-consistent regions, and few (if any) fixations to the target-inconsistent regions. Alternatively, search may only be biased toward the target-consistent regions, meaning that targets might be sought in inconsistent regions once the consistent regions have been exhausted. Third, how soon after scene onset can contextual constraints begin guiding the search process? Recent work has suggested that top-down influences require at least 100 ms following scene onset to affect oculomotor guidance (e.g., Zoest, Donk, & Theeuwes, 2004), meaning that initial saccades should not be preferentially directed to target-consistent regions in our study. Finding a disproportionate direction of initial saccades to these regions would suggest a very early guidance of search by scene context.

2.1. Method

Search displays depicted a pseudorealistic desert–mountain scene created using Discreet's 3D Studio Max. Targets and distractors consisted of blimps, jeeps, and helicopters, each of which subtended roughly 1.3° of visual angle.[1] Our goal was to create a search context approximating an observer viewing moments from a "blimp show" held in a desert. The scenes were constructed so that a mountain range acted as a divider between a low-lying desert region and a clear blue sky region. Consistent with pre-existing scene expectations, blimps appeared only in the sky and jeeps only on the ground (SC objects). The helicopter appeared equally as often in the sky as it did on the ground. There were six objects per scene with at least one object of each type present in each scene. Object color (red, green, blue, or yellow) was manipulated to avoid item duplication. In addition to target type (blimp, jeep, or helicopter), we also manipulated target presence (50% present, 50% absent) and the proportional size of the sky and ground regions in the scene (the ratios of sky to ground were 1, 2, or 0.5).[2] There were 90 experimental trials and 6 practice trials, with each object type appearing equally often as the target. Figs. 1A–C show representative search scenes at each of the three region size manipulations.

Twelve experimentally naïve SBU undergraduates served as observers. Eye movements were recorded throughout each trial using an EyeLink II eye-tracker (SR Research) sampling at 500 Hz. Search displays subtended 27° × 20° of visual angle and were presented in color on a 19 in. CRT monitor. Each trial began with a 1-s central presentation of a two-word description of the visual target, such as RED BLIMP, BLUE JEEP, or GREEN HELO.[3] The search display immediately followed the target description and remained visible until the observer made a target-present or target-absent judgment by pressing the left or right triggers of a common game pad. Observers were instructed to make their judgments as quickly as possible without sacrificing accuracy, but were not instructed as to the spatial contingencies of the blimp, jeep, and helicopter objects. They were therefore left to rely on their own scene and object knowledge to guide their search. Eye movement was completely unrestricted in this task, although gaze position at the time of search display onset was always located at or near the display's center due to the positioning of the preview target description.

2.2. Results and discussion

Errors averaged less than 5% per condition and were excluded from further analysis. An analysis of the response times (RTs) also failed to reveal significant interactions between region size and target type (SC or SU) in either the

---

[1] Because objects in scenes naturally vary in size with their positioning along the *z*-axis, the actual sizes of our object stimuli ranged from 1.2° to 1.4° of visual angle.

[2] To construct scenes having the desired proportions of sky to ground, we first found the number of pixels comprising the sky and ground regions, and then tilted the virtual 3D camera either up or down until the desired proportions were obtained.

[3] To better equate for the time needed to subvocalize the target's name during encoding (Zelinsky & Murphy, 2000), the label "HELO" was attached to the helicopter target. Subjects were explicitly informed of this designation in the experiment instructions.

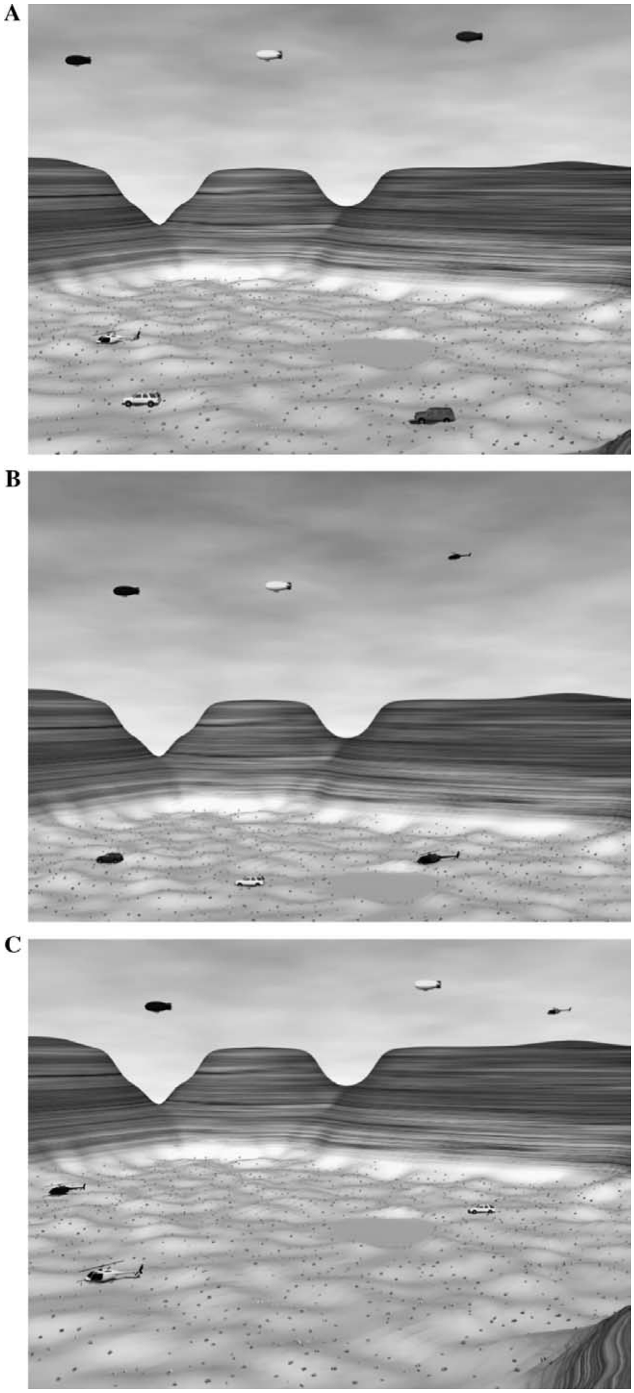

Fig. 1. Representative search displays showing scenes with (A) equal-sized sky and ground regions, (B) a sky region twice the size of the ground region, and (C) a sky region half the size of the ground region. Note that the actual scenes were in color.

target-present, $F(4, 44) = 1.79$, $p > .05$, or target-absent, $F(4, 44) = 2.01$, $p > .05$, data. All subsequent analyses therefore collapse the data over the region size manipulation. We attribute this negative result to the fact that the search relevant items could be easily segmented from their background, rendering the size of the background region relatively unimportant.

If scene context is able to guide search to likely target regions, then we would expect faster RTs for the jeep and blimp targets compared to the helicopter targets. These expectations were confirmed (Table 1). Mean target-present RTs for the SC targets were approximately 230 ms faster than for the SU target, $F(2, 22) = 8.18$, $p < .01$, a benefit amounting to roughly 17% of the average target-present search time. A similar, but non-significant trend (~80 ms) was found in the target-absent data.

To determine the degree of search restriction to a particular scene region, we analyzed for each target type the number of fixations per region and the aggregated fixation durations within a region, what we are calling the regional gaze dwell time. Fig. 2A shows this analysis for the number of fixations. To the extent that observers were able to use their knowledge of the target and scene to restrict their search, gaze should be preferentially allocated to the target-consistent scene regions. As indicated by the crossover interaction, observers indeed devoted more fixations to the sky region when searching for the blimp, and more fixations to the ground region when searching for the jeep. These interactions were reliable in both the target-present, $F(2, 22) = 61.58$, $p < .001$, and target-absent, $F(2, 22) = 6.41$, $p < .01$, data. Search for the helicopter produced an intermediate result, with fixations divided more evenly between the two regions. Our analysis of regional dwell time produced a similar result (Fig. 2B). Again, the crossover interaction in the target-present data means that observers spent more time in the sky when searching for the blimp, and spent more time on the ground when searching for the jeep, $F(2, 22) = 70.92$, $p < .001$. The fact that this interaction also appeared in the target-absent data, $F(2, 22) = 5.09$, $p < .05$, suggests that this disproportionate allocation of search to target-consistent scene regions was not solely due to holding gaze on the target object while making a manual response.

Although it is clear from Figs. 2A and B that the target-consistent regions received the most scrutiny during search, it is equally clear that search was not entirely confined to these regions. This is particularly true of the target-absent data in which observers made on average 42% of their fixations in a trial to the target-inconsistent regions. To better characterize these target-inconsistent fixations, Fig. 2C plots the proportion of final fixations as a function of scene region and target type. Consistent with the fixation number and dwell time analyses, observers were far more likely to be looking in the target-consistent regions when making their target-present search judgments. However, the opposite pattern emerged in the target-absent data, with observ-

Table 1
Mean RTs (ms) in Experiments 1 and 2 by target type

	Experiment 1		Experiment 2	
	Present	Absent	Present	Absent
BLIMP	1256 (117)	1672 (157)	1022 (89)	1142 (76)
HELO/OLEH	1500 (91)	1784 (173)	1204 (55)	1335 (74)
JEEP	1282 (111)	1737 (148)	917 (24)	1130 (57)

Note. Values in parentheses indicate one standard error of the mean.

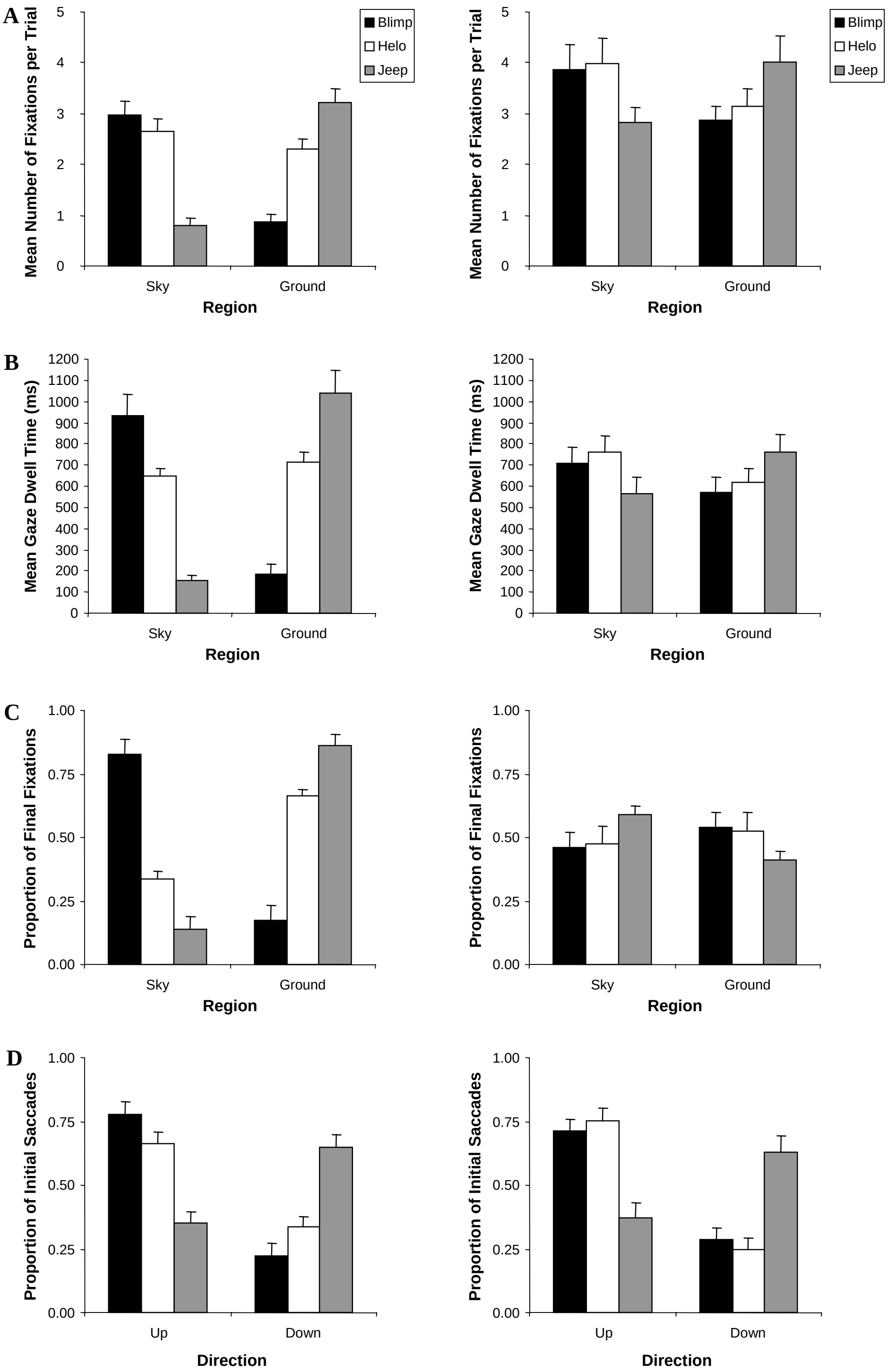

Fig. 2. Relationships between target type and scene region for four oculomotor measures in Experiment 1. Left panels show target-present data, right panels show target-absent data. (A) Mean number of fixations per trial. (B) Mean gaze dwell time. (C) Proportion of final fixations before button press. (D) Proportion of initial saccades following search display onset.

ers fixating more often on the ground when searching for the blimp, $t(11) = 4.21$, $p < .01$, and more often in the sky when searching for the jeep, $t(11) = 8.00$, $p < .001$. These patterns suggest that observers devoted their final fixations to a cursory inspection of the target-inconsistent regions prior to making a target-absent response. We believe that this conservative *checking* behavior was largely responsible for the comparatively small SC–SU benefit found in the target-absent RT data.

We now know that scene context can be used to focus search on regions likely to contain a given target, but how soon does this contextual information become available to guide eye movements during search? To answer this question we analyzed the direction of the initial saccades made by observers relative to the onset of the search display. Trials having an initial saccade less than 1° in amplitude were excluded from this analysis (approximately 5% of trials). These results are shown in Fig. 2D. When searching for a blimp, observers initially looked up towards the sky on approximately 74% of the trials; when searching for a jeep, their initial eye movements were down towards the ground on approximately 64% of the trials. In the majority of trials, the search process therefore had almost immediate access to the contextual information indicating where a given target was likely to appear in these scenes. Importantly, this Direction × Target interaction was significant in the target-absent data, $F(2, 22) = 19.76$, $p < .001$, as well as in the target-present data, $F(2, 22) = 26.53$, $p < .001$, suggesting that initial saccades were not guided towards scene regions by the target's visual features. Also interesting is the fact that observers made approximately 70% of their initial eye movements up towards the sky when looking for the helicopter target. Given that this target appeared equally often in the sky and on the ground, this bias to look up for helicopters suggests that contextual guidance during search is driven by pre-existing scene-based knowledge and not simply by probability matching to target-location contingencies specific to an experiment.

3. Experiment 2

Our findings from Experiment 1 showed that the search for SC targets can be guided to target-consistent scene regions. Observers looked preferentially to the ground when searching for a jeep, and they looked preferentially to the sky when searching for a blimp. However, our characterization of this context effect in terms of SC and SU targets is complicated by the fact that observers also initially looked to the sky when searching for the SU helicopter (Fig. 2D). One explanation for this behavior might be that our observers had a strong pre-existing expectation that helicopters should appear in the sky, and that this association could not be easily overridden by the actual locations of the target in the search displays. If this were the case, helicopters may have been treated as SC targets, even though they were probabilistically SU targets.

To better characterize search in terms of scene-constrained and scene-unconstrained performance, in Experiment 2 we used a nonsense object as the SU target. Nonsense objects cannot have pre-existing semantic associations with specific scene regions, thereby allowing a cleaner comparison between SC and SU conditions. If the search for SU targets is unaffected by scene context, we should find a pattern of initial saccades more closely approximating the 50/50 distribution of targets in the sky and ground regions. However, if we continue to find a strong sky or ground bias in the direction of the initial saccades, this would suggest that the context of the scene imposes a semantic structure on the scene objects such that search targets become assigned to scene regions even in the absence of pre-existing target-region associations.

3.1. Method

Eight experimentally naïve SBU undergraduates served as observers. None of these observers participated in Experiment 1, and all had normal or corrected-to-normal vision, by self-report. The stimuli, apparatus, design, and procedure for this experiment were identical to the descriptions provided for Experiment 1, with the following exceptions. The helicopter objects in each search scene were replaced with a fictional object type, referred to in the experiment as an "Oleh." Olehs were created by separating the helicopter object from Experiment 1 into four parts (fuselage, tail, rotor, and landing strut), then recombining these parts into a new spatial arrangement. This method preserved many of the visual features of the helicopter object (including the object's color and the area of the colored parts), while minimizing any pre-existing association between the object and a particular scene region. The name OLEH, which is HELO spelled backwards, was chosen to partially control for linguistic differences between the target designations used in Experiments 1 and 2. Observers were instructed that the olehs were fictional objects, but they were not instructed as to the spatial contingencies between these objects and the scene regions.

3.2. Results and discussion

Errors averaged less than 6% per condition and were excluded from further analysis. As in Experiment 1, an analysis of RTs did not reveal significant interactions between region size and target type in either the target-present, $F(4, 28) = 2.58$, $p > .05$, or target-absent data, $F(4, 28) = 1.37$, $p > .05$. All subsequent analyses collapse the data across the region size manipulation.

Consistent with our findings from Experiment 1, observers were approximately 235 ms faster when searching for SC targets compared to the SU target (Table 1) in target-present trials, $F(2, 14) = 8.48$, $p < .005$, and approximately 200 ms faster in target-absent trials, $F(2, 14) = 14.32$, $p < .001$. This SC benefit was obtained despite RTs being overall faster in Experiment 2, possibly due to the slightly higher error rate tolerated by this population of observers. The Experiment 1 patterns of gaze behavior to the SC targets also replicated in Experiment 2. Figs. 3A and B show that observers devoted more fixations to the target-consistent scene regions (target-present, $F(2, 14) = 142.56$, $p < .001$; target-absent, $F(2, 14) = 12.43$, $p < .005$) and looked longer in these regions (target-present, $F(2, 14) = 92.44$, $p < .001$; target-absent, $F(2, 14) = 4.37$, $p < .05$), and Fig. 3D indicates a similar pattern for the direction of the initial saccades (target-present, $F(2, 14) = 12.22$, $p < .005$; target-absent, $F(2, 14) = 4.68$, $p < .05$). However, contrary to Experiment 1, gaze patterns to the SU target did not reliably differ by scene region. Except for a small but significant tendency to fixate longer on the ground in the target-present trials, $t(7) = 2.46$, $p < .05$, corresponding analyses conducted for the SU target failed to reveal

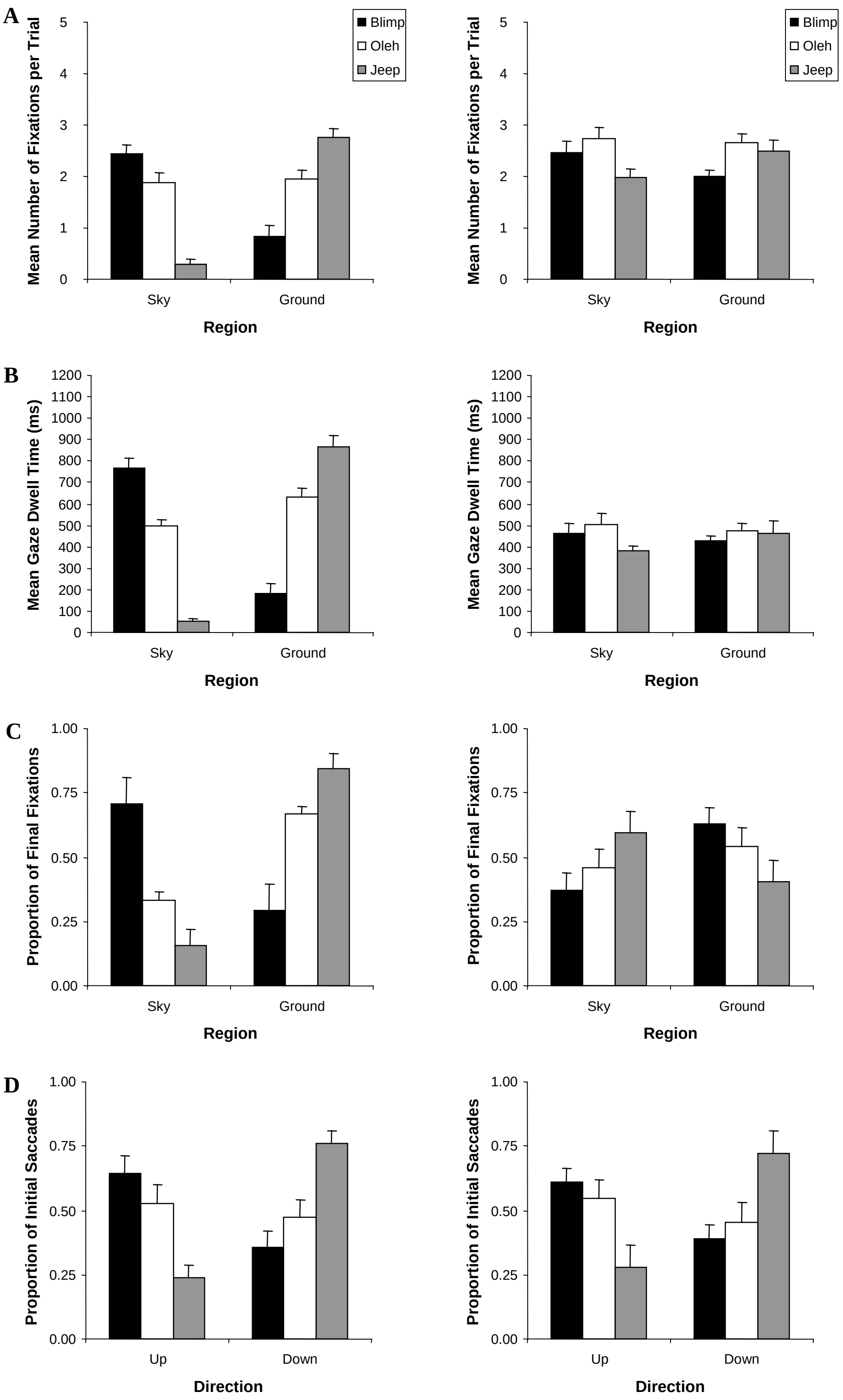

Fig. 3. Relationships between target type and scene region for four oculomotor measures in Experiment 2. Left panels show target-present data, right panels show target-absent data. (A) Mean number of fixations per trial. (B) Mean gaze dwell time. (C) Proportion of final fixations before button press. (D) Proportion of initial saccades following search display onset.

regional fixation preferences (all *p*s > .05). This absence of a regional difference is particularly notable in the initial saccade analysis. Whereas observers in Experiment 1 were as likely to look towards the sky for the helicopter target as they were for the blimp target, initial saccades in Experiment 2 showed no such directional bias for the oleh object (target-present, $t(7) = .39$, $p > .05$; target-absent, $t(7) = .59$, $p > .05$). Observers searching for the oleh target were about as likely to initially look up towards the sky (54%) as they were to look down towards the ground (46%).

Taken together, the data from Experiment 2 suggest that, in the absence of a clear expectation for where an object will likely appear in a scene, the search for SU targets will be distributed fairly evenly between the scene's regions.

4. General discussion

Previous work on contextual cueing has shown that spatial configuration information can be used implicitly to guide search to a target (Chun & Jiang, 1998, 1999). Other work has thoroughly documented the effects of target-scene consistency on search (Henderson et al., 1999) and many of the semantic factors constraining the placement of objects in scenes (Biederman et al., 1982). Our study bridges these two literatures by showing that contextual information based on one's expectation of a target's location in a scene can also be used to guide search behavior. As in the case of Yarbus's (1967) demonstration that picture viewing depends on the task assigned to the observer, we have shown that observers will search a scene differently depending on the target of their search and their expectation of where this target might appear in the scene.

The current study broadens our understanding of scene-based contextual guidance in three respects. First, by comparing SC and SU conditions in Experiments 1 and 2, we were able to quantify the size of the contextual guidance benefit, which in the case of the target-present data was approximately 17% of the average search time in Experiment 1 and 21% in Experiment 2. Although the actual size of this benefit was, at most, only 235 ms in our relatively easy search task, extrapolating this benefit to a moderately difficult real-world search task can easily result in several seconds of savings. Second, we determined that this guidance process, although clearly able to focus search on the target-consistent scene regions (Figs. 2A and B and 3A and B), did not confine search to these regions. As indicated by the final fixation analyses (Figs. 2C and 3C), observers often ended their target-absent search by looking in the inconsistent region after determining that the target was not present in the consistent region. This behavior suggests that scene context works by biasing search towards target-consistent scene regions (Navalpakkam & Itti, 2005); not by excluding target-inconsistent scene regions from the search process (see Desimone & Duncan, 1995, for additional discussion of attentional biasing). Third, the time-course of the contextual guidance process is informed by our finding that initial saccades were preferentially directed to target-consistent scene regions (Figs. 2D and 3D). Although this guidance was not perfect, our data suggest that scene-based guidance can be available almost immediately to the search process. This evidence for early guidance also casts doubt on those theories of gaze control claiming that initial fixations are driven largely by bottom-up processes (Mannan, Ruddock, & Wooding, 1995; Parkhurst, Law, & Niebur, 2002). Clearly, the scene-context effects reported in the current study indicate a top-down process.

We broadly interpret our findings as evidence for a fast-acting and pronounced top-down bias of the search process by scene context, a bias that has been relatively neglected in the search literature. Just as search may be guided by low-level visual features (Wolfe, 1994), search is also guided by scene constraints imposed on the locations of objects. However, we do not see this top-down guidance process as being incompatible with bottom-up theories of search and gaze control. Rather, we envision a common underlying search representation that integrates control signals from a wide variety of informational sources (see Itti & Koch, 2001; and Corbetta & Schulman, 2002, for possible brain areas responsible for this integration function). One source of control arises from feature-contrast in a scene, as described by saliency-based theories of search (Itti & Koch, 2000), and another source of control stems from feature matches between the target and the search scene (Rao, Zelinsky, Hayhoe, & Ballard, 2002; Zelinsky, 2005). Other forms of control undoubtedly arise from the demands specific to a given task (Land & Hayhoe, 2001; Navalpakkam & Itti, 2005; Turano, Geruschat, & Baker, 2003). The current study describes yet another form of top-down control, one based on scene context. In the absence of scene-based guidance, oculomotor search is driven by signals from bottom-up saliency and top-down feature matches spanning the entirety of the search display. However, with the addition of scene-based constraints, it may be possible to bias the search space to regions of the display most likely to contain the search target.

One limitation of the current study is that scene context was defined rather narrowly (for only a single type of scene) and for targets in which there exist clear and over-learned scene constraints (jeeps and blimps). Future work will explore the boundary conditions of scene-based guidance by using less-constraining scenes and targets with weaker scene associations. We will also seek to better understand the use of scene constraints in the real world, which we suspect may be far more pronounced than those reported here. Whereas observers viewing a computer-generated scene may harbor some uncertainty as to whether scene constraints will be violated from trial to trial, observers searching for actual objects in the real world may be more inclined to trust and use this information to guide their search.

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.

References

Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: detecting and judging objects undergoing relational violations. *Cognitive Psychology, 14*, 143–177.

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology, 36*, 28–71.

Chun, M. M., & Jiang, Y. (1999). Top-down attentional guidance based on implicit learning of visual covariation. *Psychological Science, 10*, 360–365.

Corbetta, M., & Schulman, G. L. (2002). Control of goal-oriented and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3*(3), 201–215.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*, 193–222.

Endo, N., & Takeda, Y. (2004). Selective learning of spatial configuration and object identity in visual search. *Perception & Psychophysics, 66*(2), 293–302.

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.

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.

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

Mannan, S., Ruddock, K. H., & Wooding, D. S. (1995). Automatic control of saccadic eye movements made in visual inspection of briefly presented 2-D images. *Spatial Vision, 9*, 363–386.

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

Oliva, A., Torralba, A., Castelhano, M.S., & Henderson, J. M. (2003). Top-down control of visual attention in object detection. In *Proceedings of the IEEE international conference on image processing* (Vol. I, pp. 253–256). Barcelona, Spain: IEEE.

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.

Peterson, M. S., & Kramer, A. S. (2001). Attentional guidance of the eyes by contextual information and abrupt onsets. *Perception & Psychophysics, 63*(7), 1239–1249.

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

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.

Wolfe, J. M. (1994). Guided Search2.0: A revised model of guided search. *Psychonomic Bulletin and Review, 1*, 202–238.

Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum.

Zelinsky, G., & Murphy, G. (2000). Synchronizing visual and language processing: An effect of object name length on eye movements. *Psychological Science, 11*, 125–131.

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). Amsterdam: Elsevier.

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