Attentional capture is contingent on scene region: Using surface guidance framework to explore attentional mechanisms during search

Effie J. Pereira & Monica S. Castelhano

Springer

Springer

BRIEF REPORT

Attentional capture is contingent on scene region: Using surface guidance framework to explore attentional mechanisms during search

Effie J. Pereira[1] · Monica S. Castelhano[2]

Abstract

Studies have established that scene context guides attention during visual search, but it is not yet clear how. In this study, we examined how attention is deployed across scenes using an attentional capture paradigm. Using the Surface Guidance Framework (Castelhano & Pereira, 2019), we defined target-relevant and target-irrelevant scene regions for each target object and compared how attentional capture of a suddenly onsetting distractor differs for object and letter searches. We found an enhancement of capture effects when distractors appeared within target-relevant regions, with greater proportions of distractors fixated and greater proportions of saccades made toward the distractor for object searches, but not for letter searches. Thus, attention in the real world can be flexibly and spatially distributed on the basis of contextual information, with the Surface Guidance Framework presenting a powerful tool for exploring attentional guidance in real-world scenes.

Keywords Real-world attention · Attentional capture · Scene context · Visual search

Whether navigating through our environment or searching for a friend in a crowd, attentional mechanisms allow us to selectively process a subset of the available information with relative ease. For example when driving, we may pay attention to road signs on the sidewalk or we may be captured by an indicator light on the dashboard. Researchers have posited that attentional capture is dependent on two types of constraining factors: top-down influences that voluntarily direct attention (e.g., road signs) and bottom-up properties that can involuntary capture attention (e.g., indicator light). In this study, we examined how these factors interact during visual search in real-world scenes.

Early research has examined the dichotomy between voluntary and involuntary deployment of attention. From a top-down perspective, attentional deployment is deliberative and occurs when knowledge about the task or target guides selection (Posner, 1980). From a bottom-up perspective, attentional deployment is directed by external factors (i.e., saliency, motion, abrupt onsets) regardless of task relevance (Theeuwes, 1994; Yantis & Jonides, 1984). Although there is disagreement about how top-down filtering occurs (Gaspar & MacDonald, 2014), most agree that attentional capture can reveal how information is prioritized for processing.

Real-world scenes differ from visual arrays in a multitude of ways, but primarily through the influence of general knowledge and prior experience on processing (Brockmole & Henderson, 2005). Researchers have long theorized that real-world scenes contain a richly diverse set of visual characteristics that can differentially affect attentional guidance. For example, visual search in scenes can be restricted to likely target locations (Castelhano & Heaven, 2011; Castelhano & Henderson, 2007; Neider & Zelinsky, 2006; Pereira & Castelhano, 2014; Torralba, Oliva, Castelhano, & Henderson, 2006), such that processing is prioritized based on broader scene regions rather than specific locations.

In the current study, we investigated the degree to which attentional deployment is informed by spatial associations between objects and scenes during visual search. Thus, for a specific

✉ Effie J. Pereira
effie.pereira@mail.mcgill.ca

1 Department of Psychology, McGill University, 1205 Avenue Docteur Penfield, Montreal, QC H3A 1B1, Canada

2 Department of Psychology, Queen's University, 62 Arch Street, Kingston K7L 3L3, ON, Canada

_Springer

target, we sought to operationalize target-relevant regions of the scene (i.e., where a target object is likely to appear) and target-irrelevant regions (i.e., where a target object is unlikely to appear). To do so, we developed the Surface Guidance Framework. Based on the output of the Contextual Guidance Model, Torralba and colleagues (2006) found that an object's position is captured by their likely vertical position in the scene, with horizontal positioning being less informative as an object is equally likely found to the left or right of a spatial region (e.g., paintings are found in upper regions and mugs are found on midlevel regions such as countertops; Torralba et al., 2006, Fig. 11).

Drawing from these results, we posit that the surface regions at different vertical positions capture the spatial relationships between objects and scenes. For instance, if we divide scenes into three horizontal surfaces—(1) upper (e.g., ceiling, upper walls), (2) middle (e.g., countertops, tabletops, desktops, stovetops), and (3) lower (e.g., floor, lower walls)—then we can specify the target objects associated with each region—(1) upper (e.g., painting, wall clock), (2) middle (e.g., blender, book), and (3) lower (e.g., garbage bin, shoes). Using this method of defining targets in relation to their spatial region, we are able to designate "target-relevant regions" based on the likelihood of finding the target in this region. Thus, the Surface Guidance Framework enables the exploration of how attentional deployment differs across target-relevant or target-irrelevant regions (see Fig. 1).

In the present study, we used eye tracking to examine whether attentional deployment in real-world scenes is spatially modulated by scene context, such that attentional capture is dependent on the task-relevance of the spatial location of a sudden-onset distractor. We explored this question using two types of searches—object and letter—in a between-participants design. For "object" searches, participants searched for an object in a scene in order to take advantage of strong spatial associations between target objects and their expected scene location. Conversely, "letter" searches, which are not associated with any specific scene regions, were used as a control to compare attentional capture effects when scene context does not affect search. For each type of search, an abruptly onsetting distractor object occurred on 50% of trials. Of theoretical importance, we manipulated the location of the distractor object to be either within or outside the target-relevant region (target-relevant and target-irrelevant conditions, respectively). If attention were deployed to target-relevant regions within object searches, we would expect the greatest degree of attentional capture for distractors appearing in target-relevant versus target-irrelevant conditions, but no differences between the two conditions for letter searches.

Further, because there is no theoretical reason to expect a difference between target-relevant and target-irrelevant regions for letter searches, we additionally collapsed across these conditions to use as a single baseline to explore the type of attentional mechanism present. If a facilitation of processing in the target-relevant region is responsible for the effect, then the target-relevant condition should show a systematically higher level of capture than the letter-control baseline. If an inhibitory process in the target-irrelevant region is responsible for the effect, then the target-irrelevant condition should show a systematically lower level of capture than the letter-control baseline.

Method

Participants

Sample size was determined in a two-fold manner. First, an a priori power analysis (G*power; Faul, Erdfelder, Lang, &

Fig. 1 (a) Examples of search scenes with (b) highlighted surface regions (red = upper, yellow = middle, green = lower) as per the Surface Guidance Framework. (Color figure online)

Buchner, 2007) estimated effect sizes. Based on the Brockmole and Henderson (2005) study, we estimated a medium-to-large effect (Cohen's d_z ranging from .92 to 3.42), and with power set at a minimum of .95, we determined that four to18 participants per group were required. Taking into account counterbalancing, 16 participants per object and letter search group would be appropriate within this range.

Second, because there is always a degree of uncertainty regarding the sample size needed to observe a statistically significant effect (Lakens & Evers, 2014), we also applied a sequential analysis to determine the final stopping point for our study (Lakens, 2014). Sequential analyses allow for conducting interim analyses without undermining null hypotheses significance testing and while controlling for Type I error rate.

With sequential analysis, we declared the number of stopping points (i.e., intervals) in order to determine the alpha boundaries for each of these declared points, which were then used to determine when data collection was complete.[1] Using a linear spending function (i.e., a power family function with a phi of 1) from the GroupSeq package on R (Pahl, 2018), we computed the corrected alpha boundaries for two-sided interim analyses at three intervals with an initial alpha level = .05. Thus, we aimed to collect 32 participants for our first interval (16 per object and letter search group), with a critical alpha boundary = .017. If we failed to reject the null hypothesis at this stage, we then planned to collect an additional 32 participants (64 in total, 32 per group) for our second interval at a critical alpha boundary = .022, and if necessary, a further 32 participants would be collected (96 in total, 48 per group) for our final interval at a critical alpha boundary = .028.

In the current study, we collected data for two intervals (see Supplementary Material for analysis of the first interval). Sixty-four (32 per group) Queen's University undergraduate students participated and were compensated either with course credit or $10/hr for their participation.

Apparatus and stimuli

Eye movements were tracked using an EyeLink 1000 (SR Research; Mississauga, ON) at 1000 Hz. Stimuli were presented on a 21-in. CRT monitor (800 × 600 pixels), with a refresh rate of 100 Hz, controlled by Experiment Builder (SR Research). Participants were seated 60 cm away, stabilized by a head and chin rest, and although viewing was binocular, only the right eye was tracked.

The stimuli consisted of 36 real-world indoor scene photographs obtained from various sources. The scenes subtended a visual angle of 38.1° × 28.6°; targets and distractors had an average size of 2.5° × 2.6°. The relevant contextual regions were counterbalanced across the three scene regions: upper,

middle, and lower regions (Castelhano & Pereira, 2019; Torralba et al., 2006, Fig. 11).

The search target was manipulated as a between-subjects factor across two groups: (1) *object search*: a target object search; and (2) *letter search*: a target letter search. For object searches, objects were defined as smaller scale discrete entities that were easily moveable within the scene (e.g., books, trash cans; Henderson & Hollingworth, 1999).[2] For letter searches, letter targets (gray, Times New Roman, 11-pt font) were placed in the central x–y coordinates of the target object, which were digitally removed (using Adobe Photoshop CS5).

For each type of search, distractor presence was manipulated as a within-subjects factor across two conditions: (1) *absent*: no distractor onset; and (2) *present*: an abruptly onsetting distractor occurred on 50% of trials. Of theoretical importance, for distractor-present trials, we manipulated the location of the distractor object as a within-subjects factor across two conditions: (1) *target-relevant condition*: the distractor object appeared in the same contextual region as the target; and (2) *target-irrelevant condition*: the distractor object appeared outside the contextual region of the target. Distractor objects were placed in the scene at an equidistant location from the target (see Fig. 2).

Procedure

Participants were instructed to search the scene for the prespecified target and to press a response button once it was found. They were then calibrated on the eye tracker using a nine-point calibration screen to ensure high accuracy (average spatial error ≤.4°, maximum spatial error ≤.7°). Calibration was checked prior to every trial using a five-point calibration screen.

For each trial, participants were presented with a target word or target letter in the center of the screen for 2 s, followed by a fixation cross for 500 ms. The search scene was then displayed until a response was made or until 20 s had elapsed. On 50% of the trials, no distractor object was present. On the other 50%, a distractor object appeared 50 ms after the beginning of the first voluntary fixation on the scene (split evenly between target-

---

[1] Preregistration of the sequential analysis plan can be found at the Open Science Framework: osf.io/zufe3.

[2] The position of the target objects (in the upper, middle, and lower regions) was independently verified by a separate group of participants ($N = 10$). They were shown the scene in which the target had been erased (using Adobe Photoshop CS5) and asked to draw polygons to denote where they thought the named target would be found. Participants were not given any limitations on the number or the shape of the polygons to draw and used a mouse to select the edges of the polygon until a closed shape was formed. A majority of the participants drew polygons with four points, and on average they drew four unique regions overall. The regions on average subtended 36% of the area of the entire image. Polygons drawn overlapped with the exact target position on 78% of trials and corresponded to the vertical position of the target on 89% of trials. There were no systematic differences between images, and any differences between analyses were due to, for example, participants highlighting the walls across the room but not highlighting all the walls or all sections of the walls as potential target locations.

Fig. 2 Example search images across contextual and distractor search conditions. For illustrative purposes, target letters are depicted in higher contrast and are not drawn to scale. The targets are highlighted in blue, and the distractor objects are highlighted in red. (Color figure online)

relevant and target-irrelevant region conditions; see Fig. 3). Participants completed four practice trials prior to 36 experimental trials. Conditions were counterbalanced across participants, and each participant saw each search scene only once. The experiment took ~30 mins.

Results

Data analysis

To examine the effect of distractors on processing, we analyzed the data in two ways: (1) visual search analyses, reflecting overall task performance, and (2) attentional capture analyses, reflecting the immediate effects of the abruptly onsetting distractor on processing. Because we were interested in processes involved in active and ongoing search, we imposed a minimum performance criterion of 80% search accuracy (zero participants excluded). For all eye-movement measures in both analyses, fixation durations ≤ 90 ms and $\geq 2,000$ ms were excluded as outliers (3,015 of 21,878 fixations; 13.8%). Target and distractor objects were defined by a rectangular region 1° from their outermost edge.

For visual search analyses, we examined overall task performance across reaction time (RT), latency to target, and number of fixations to target. For each measure, we compared the distractor present and absent conditions for object and letter searches using a 2 × 2 mixed ANOVA. This analysis allowed for a comparison of the current study with previous research on attentional capture.

For the attentional capture analyses, we focused on the immediate effects of the distractor by examining the time period immediately before and after its onset for distractor-present trials only. Measures included the proportion of fixations on the distractor, proportion of saccades toward the distractor (regardless of where the fixation landed), and fixation durations immediately before and just after the distractor onset. For each measure, we compared the target-relevant and target-irrelevant distractor conditions for each target search using two-tailed paired-samples t tests. Additionally, after checking and finding no differences between target-relevant and target-irrelevant conditions for letter searches, we collapsed across these conditions to form an overall letter-control condition to use as a baseline. This allowed us to examine whether attentional capture effects were facilitative, inhibitory, or neutral when compared with baseline
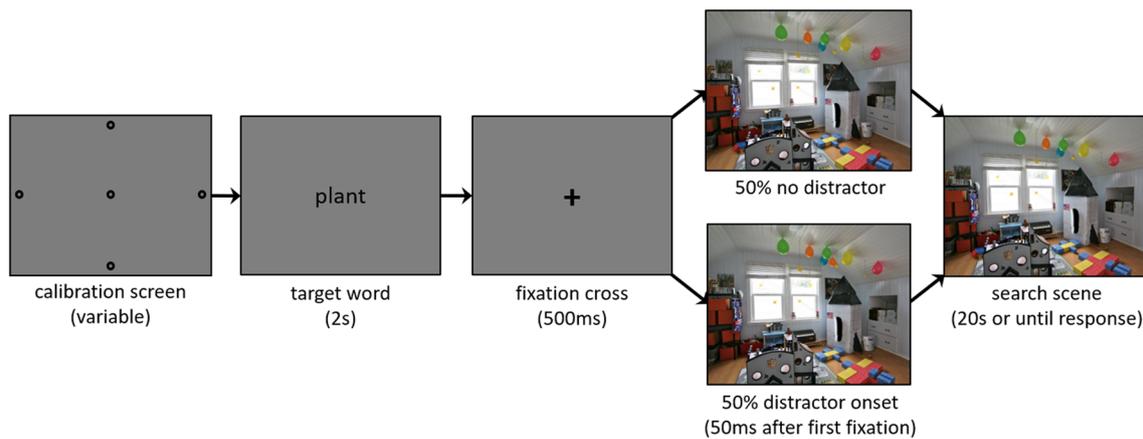
Fig. 3 The trial sequence for object search with the distractor in a target-irrelevant region. Participants began by fixating on the central point of the calibration screen. A word describing the target was presented for 2 s, followed by a fixation cross for 500 ms. The search scene was then shown for a maximum of 20 s or until the participant made a response. On 50% of trials, no distractor appeared; for the other 50%, a distractor object would onset 50 ms after the first voluntary fixation on the scene

performance. For this, we compared the target-relevant and target-irrelevant object search conditions with the letter-control baseline using two-tailed independent-sample t test.[3] The critical alpha was set at $\alpha = .022$ for all analyses. Data analysis plans were preregistered and can be found at the Open Science Framework (osf.io/zufe3).

Visual search analysis

Overall, participants' accuracy (defined as a fixation on the target and a button press within three fixations of target fixation) was high at 92%. Measures (means and standard deviations) are presented in Table 1.

Reaction time Response time was defined as elapsed time from the onset of the search scene until a response was made. A significant main effect of target search was found, $F(1, 62) = 89.53$, $p < .001$, $\eta_p^2 = .59$, with participants taking longer to find letters than objects. We also found a main effect of distractor presence, $F(1, 62) = 7.00$, $p = .01$, $\eta_p^2 = .10$, with longer searches when the distractor was present versus absent. No interaction was found between target search and distractor presence, $F(1, 62) = .53$, $p = .47$, $\eta_p^2 = .01$.

Latency to target Target latency was defined as the elapsed time between the onset of the search scene until first fixation on target (excluding the first fixation). Similar to RT, we found a significant main effect of target search, $F(1, 62) = 229.17$, $p < .001$, $\eta_p^2 = .79$, with longer latencies for letter versus object searches, and a significant main effect of distractor presence, $F(1, 62) = 5.53$, $p = .022$, $\eta_p^2 = .08$, with longer

searches when the distractor was present, but no interaction between target search and distractor presence, $F(1, 62) = .03$, $p = .86$, $\eta_p^2 = .001$.

Number of fixations to target Number of fixations to target was defined as the number of individual fixations made until the first fixation on the target. As with previous measures, a significant main effect of target search was found, $F(1, 62) = 215.23$, $p < .001$, $\eta_p^2 = .78$, along with a significant main effect of distractor presence, $F(1, 62) = 6.36$, $p = .014$, $\eta_p^2 = .09$, with no significant interaction between the two, $F(1, 62) = .03$, $p = .86$, $\eta_p^2 = .001$.

Thus, across global visual search measures, search performance was affected by distractor presence, with less efficient searches occurring for letter than for object searches, when scene context could be used to predict target location.

Attentional capture analysis

Proportion of distractors fixated Proportion of distractors fixated was calculated as the proportion of trials in which participants fixated on the distractor after its onset (within three fixations to account for corrective or reprogrammed fixations). Greater proportion of distractors were fixated in target-relevant versus target-irrelevant distractor conditions for object searches, $t(31) = 3.80$, $p = .001$, $d_z = .67$, 95% CI [.07, .24], but not for letter searches, $t(31) = .31$, $p = .76$, $d_z = .06$, 95% CI [−.08, .11]. When collapsing across letter searches ($M = .39$, $SD = .17$), target-relevant regions were facilitated compared with letter-control, $t(62) = 2.35$, $p = .022$, $d = .59$, 95% CI [.02, .22], however no differences were found between target-irrelevant and letter-control, $t(62) = .90$, $p = .37$, $d = .23$, 95% CI [−.13, .05].

---

[3] See Supplementary Materials for additional exploratory analyses.

Table 1 Visual search measures as a function of target search and distractor presence

| | Object search | | | | Letter search | | | |
| | Distractor present | | Distractor absent | | Distractor present | | Distractor absent | |
	M	SD	M	SD	M	SD	M	SD
Reaction time (ms)	1,562	452	1,425	365	2,580	589	2,398	540
Latency to target (ms)	1,014	204	867	196	2,086	566	1,929	471
# of fixations to target	3.63	0.79	3.06	0.67	7.23	1.84	6.67	1.57

We also examined ordinal fixations directly after distractor onset (Fixation$_d$). A greater proportion of distractors were fixated immediately following its onset for target-relevant versus target-irrelevant distractor conditions for object searches, $t(31)$ = 2.96, p = .006, d_z = .52, 95% CI [.04, .20], but not for letter searches, $t(31)$ = .23, p = .82, d_z = .04, 95% CI [−.09, .07]. When collapsing across letter searches (M = .24, SD = .16), no difference was found between target-relevant and letter-control, $t(62)$ = 1.88, p = .07, d = .47, 95% CI [−.01, .17], nor between target-irrelevant and letter-control, $t(62)$ = .79, p = .44, d = .20, 95% CI [−.12, .05]. Means are presented in Fig. 4.

Proportion of saccades toward the distractor Proportion of saccades was calculated as the number of trials in which participants saccaded toward the distractor immediately after its onset (within ~10°). Greater proportion of saccades were directed toward target-relevant versus target-irrelevant distractor conditions for object searches, $t(31)$ = 2.89, p = .007, d_z = .51, 95% CI [.03, .19], but not for letter searches, $t(31)$ = .39, p = .70, d_z = .07, 95% CI [−.06, .09]. When collapsing across letter searches (M = .18, SD = .13), no differences were found between target-relevant and letter-control, $t(62)$ = 2.20, p = .03, d = .55, 95% CI [.01, .16], nor between target-irrelevant and letter-control, $t(62)$ = .74, p = .46, d = .19, 95% CI [−.10, .04]. Means are presented in Fig. 5.

Fixation duration before versus after distractor onset Fixation duration before versus after onset was calculated as the difference in fixation duration directly before and immediately after onset of the distractor. Here, we found no differences for target-relevant vs. target-irrelevant distractor conditions for object searches, $t(31)$ = .55, p = .59, d_z = .10, 95% CI [−10.98, 19.06], nor for letter searches, $t(31)$ = .97, p = .94, d_z = .17, 95% CI [−11.45, 32.24]. However, after collapsing into letter-control (M = −8.31, SD = 29.10), we found a facilitative effect for target-relevant distractors versus letter-control searches, $t(62)$ = 2.53, p = .014, d = .63, 95% CI [4.32, 36.63]. No differences were found between target-irrelevant and letter-control, $t(62)$ = 1.93, p = .06, d = .48, 95% CI [−.57, 33.45]. Means are presented in Fig. 6. Thus, the overall pattern of results suggests that when searching for an object within a scene, distractors in target-relevant regions affected fixation planning but not distractor processing.[4]

Discussion

The current study investigated the degree to which scene context can modulate attentional deployment and affect attentional capture during visual search in natural scenes. Using a standard capture paradigm in real-world scenes, we found that participants fixated distractors more frequently and had greater proportion of saccades toward distractors when they could use scene context to guide their visual search. Our findings also revealed no differences across fixation duration measures directly after distractor onset, suggesting that attentional capture may affect global fixation planning rather than specific individual processing measures. These results are consistent with previous work showing that abruptly onsetting objects do capture attention within scenes (Brockmole & Henderson, 2005) and that scene context modulates attentional deployment during visual search (Castelhano & Henderson, 2007; Neider & Zelinsky, 2006; Pereira & Castelhano, 2014). However, our findings further indicate that scene context constrains top-down deployment of attention based on broad spatial boundaries informed by the task-relevance of scene regions.

The pattern of results demonstrates that attentional capture in scenes can be flexibly biased based on knowledge of the likely target location (i.e., spatial associations between the

---

[4] To ensure that attentional capture measures were not driven by proximity to the distractor, we examined the distance in visual angle from the fixation location to the distractor at the moment of distractor onset. We found no differences between target-relevant versus target-irrelevant distractor conditions for object searches, $t(31)$ = 1.26, p = .22, d_z = .22, 95% CI [−1.47, .35]; however, letter searches showed that current fixations were closer to the distractor for target-irrelevant versus target-relevant searches, $t(31)$ = 4.28, p < .001, d_z = .76, 95% CI [1.00, 2.83]. Although we believe this to be an artifact of search, in that some fixations were likely to be closer to the distractor by way of chance, if the distractor was closer to the fixation upon onset, it theoretically should have resulted in greater capture effects given fixation proximity to the event. This, however, is not the pattern of results seen across all attentional capture measures and lends more credence to the hypothesis that context modulates attentional capture.
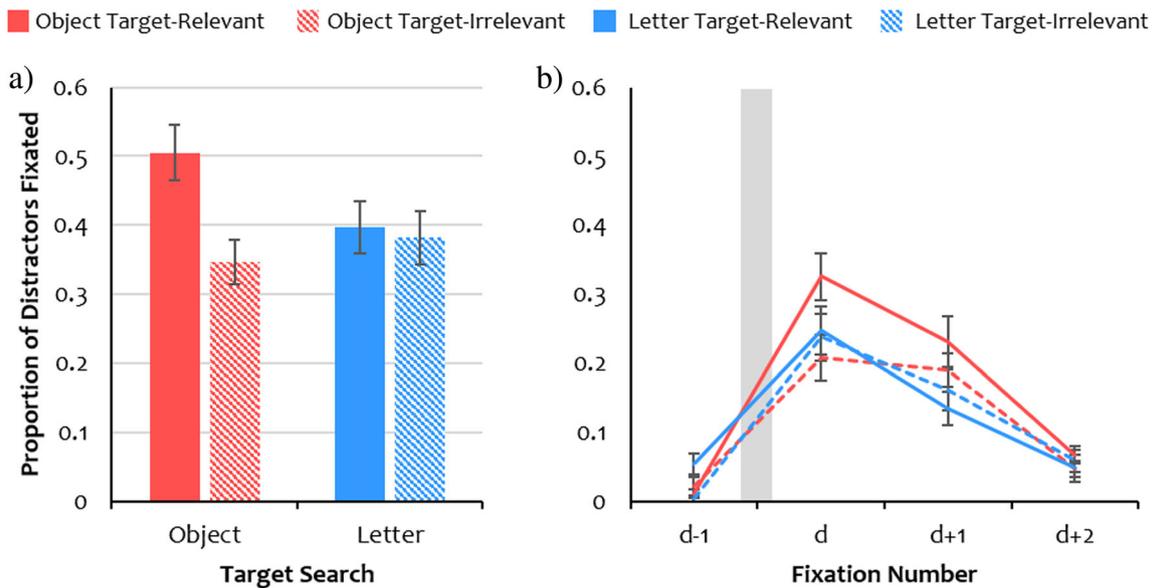
Fig. 4 Mean ± 1 *SE* for proportion of distractors fixated (**a**) overall and (**b**) by fixation number after distractor onset (Fixation$_d$), as a function of distractor search condition. The shaded area represents the occurrence of the sudden-onset distractor. (Color figure online)

target object and scene). These findings cannot be solely accounted for by theories that suggest that attention is tuned to detect uniquely onsetting distractors (e.g., Yantis & Jonides, 1984) or by theories that selectively modulate attention based on target features or specific spatial locations (e.g., Folk, Remington, & Johnston, 1992). Instead, we posit that attention in natural scenes is spatially distributed on the basis of contextually-defined relevant regions, thus providing important information about how attention is distributed in the real world. In contrast to traditional visual search arrays, real-world scenes do not have specific locations demarcated or highlighted beforehand, but they have been shown to affect

how attention and eye movements are directed (e.g., Neider & Zelinsky, 2006). The current findings demonstrate that attention is differentially deployed across scenes, with an enhanced focus on task-relevant regions.

The results from the present study also have interesting implications for how eye movement planning and attentional deployment occur in the real world. This is particularly significant in light of ongoing discussions aimed at reconceptualising attentional control from a dichotomy of top-down and bottom-up mechanisms to an integrative system that is influenced by perceptual and task information, as well as knowledge and past experience (Awh, Belopolsky, &
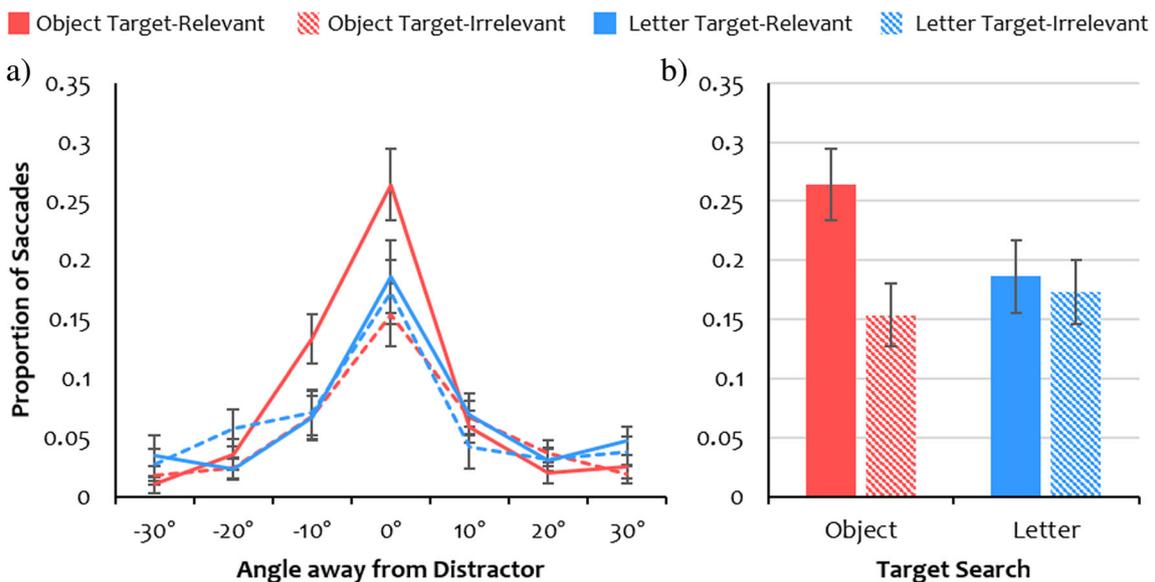
Fig. 5 Mean ± 1 *SE* for (**a**) proportion of saccades directed toward the distractor (normalized to 0°) within an angular deviation of 10° for each bin, and (**b**) proportion of saccades at 0°, as a function of distractor search condition. (Color figure online)
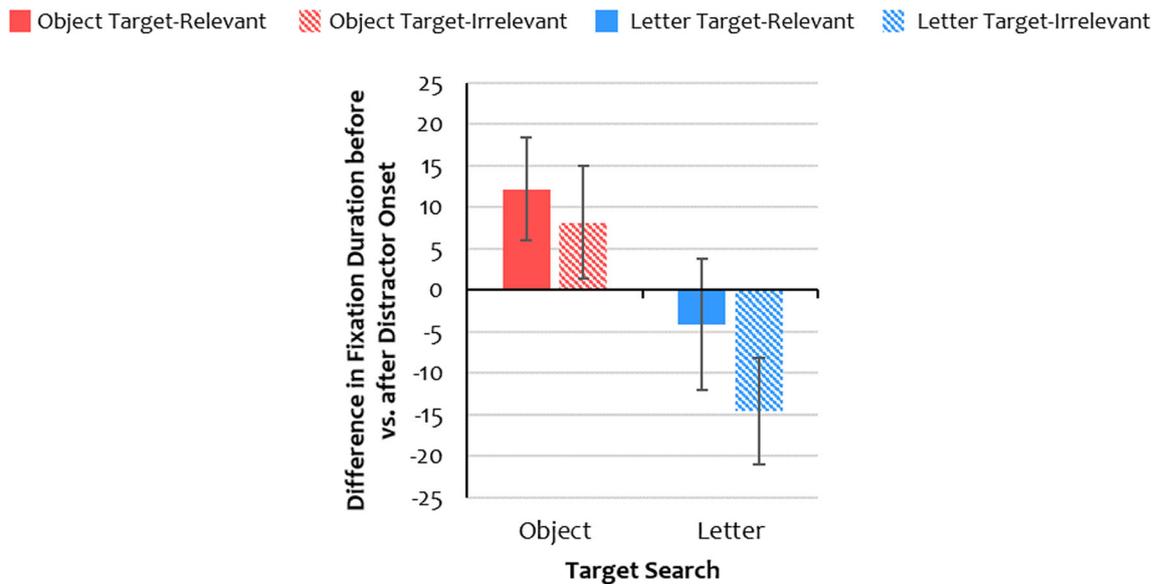
Fig. 6 Mean ± 1 *SE* for difference in fixation durations before versus after distractor onset as a function of distractor search condition. (Color figure online)

Theeuwes, 2012). Because the onset of the distractor occurred during the first voluntary fixation on the scene, it suggests that relevant regions can be prioritized for processing immediately after search onset. This is consistent with the notion that contextual information is acquired quickly and can help define task-relevant regions early in processing (Castelhano & Henderson, 2007; Võ & Henderson, 2009). This early focus adds to the literature on the effects of prior knowledge and experience on attentional control. Other studies have found that both trained and long-term associations automatically affect attentional deployment (Leber & Egeth, 2006). In a similar vein, our study looked at constraints imposed by prior knowledge of how objects are typically arranged in scenes (Castelhano & Witherspoon, 2016), thus suggesting that attentional capture effects can be modulated based on long-established associations. These findings suggest that attentional control settings can be based on a combination of prior knowledge and stimulus structure, further supporting the view of attentional control as a dynamic integrative system.

Our findings also speak to the investigative power of specifying target-relevant and target-irrelevant regions using the Surface Guidance Framework. With this framework, we were able to balance regions that were relevant across scenes (equally distributed across upper, middle, and lower), as well as contrast how processing differs across regions based on the expectancy of target location. Rather than specific *x–y* coordinates, regions were based on the position of different surfaces relative to the overall scene image. As such, the Surface Guidance Framework supports previous attentional capture work in real-world settings that has found that central objects capture fixations during driving (Underwood, Chapman, Berger, & Crundall, 2003) and that having expertise in

specific visual environments can heighten attentional focusing within task relevant regions (Crundall, Underwood, & Chapman, 1999). Thus, by operationalizing a target's expected location, the Surface Guidance Framework represents an effective tool for exploring the effects of context on attentional guidance in more natural settings.

In conclusion, we demonstrate that the influence of context on the dynamic and integrative nature of attention is much broader than previously thought, furthering our understanding of these mechanisms in the real world and highlighting the necessity of studying attention through an integrative lens that ties together sensory information, current goals, prior knowledge, and experience.

References

Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: Q failed theoretical dichotomy. *Trends in Cognitive Sciences,16*(8),437–443.

Brockmole, J. R., & Henderson, J. M. (2005). Prioritization of new objects in real-world scenes: Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 31,* 857–868.

Castelhano, M. S., & Heaven, C. (2011). Scene context influences without scene gist: Eye movements guided by spatial associations in visual search. *Psychonomic Bulletin & Review, 18*(5), 890–896.

Castelhano, 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*(4), 753–763.

Castelhano, M. S., & Pereira, E. J. (2019). *Searching through the clutter: Using surface guidance framework to explore set size effects in scenes.* Manuscript under review.

Castelhano, M. S., & Witherspoon, R. L. (2016). How you use it matters: Object function guides attention during visual search in scenes. *Psychological Science, 27*(5), 606–621.

Crundall, D., Underwood, G., & Chapman, P. (1999). Driving experience and functional field of view. *Perception, 28*(9), 1075–1087.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191.

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 1030–1044.

Gaspar, J. M., & McDonald, J. J. (2014). Suppression of salient objects prevents distraction in visual search. *The Journal of Neuroscience, 34*(16), 5658–5666.

Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology, 50*(1), 243–271.

Lakens, D. (2014). Performing high-powered studies efficiently with sequential analyses. *European Journal of Social Psychology, 44*(7), 701–710.

Lakens, D., & Evers, E. R. K. (2014). Sailing from the seas of chaos into the corridor of stability: Practical recommendations to increase the informational value of studies. *Perspectives on Psychological Science, 9*(3), 278–292.

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

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

Pahl, R. (2018). GroupSeq: A GUI-based program to compute probabilities regarding group sequential designs (R Package Version 1.3.5) [Computer software]. Retrieved from https://CRAN.R-project.org/package=GroupSeq

Pereira, E. J., & Castelhano, M. S. (2014). Peripheral guidance in scenes: The interaction of scene context and object content. *Journal of Experimental Psychology: Human Perception and Performance, 40*(5), 2056–2072.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*(1), 3–25.

Theeuwes, J. (1994). Stimulus-driven capture and attentional set: Selective search for color and visual abrupt onsets. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 799–806.

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review, 113,* 766–786.

Underwood, G., Chapman, P., Berger, Z., & Crundall, D. (2003). Driving experience, attentional focusing, and the recall of recently inspected events. *Transportation Research Part F: Traffic Psychology and Behaviour, 6*(4), 289–304.

Võ, M. L.-H., & Henderson, J. M. (2009). Does gravity matter? Effects of semantic and syntactic inconsistencies on the allocation of attention during scene perception. *Journal of Vision, 9*(3), 1–15.

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