Neural Correlates of Divided Attention in Natural Scenes

Sabrina Fagioli[1] and Emiliano Macaluso[1,2]

Abstract

■ Individuals are able to split attention between separate locations, but divided spatial attention incurs the additional requirement of monitoring multiple streams of information. Here, we investigated divided attention using photos of natural scenes, where the rapid categorization of familiar objects and prior knowledge about the likely positions of objects in the real world might affect the interplay between these spatial and nonspatial factors. Sixteen participants underwent fMRI during an object detection task. They were presented with scenes containing either a person or a car, located on the left or right side of the photo. Participants monitored either one or both object categories, in one or both visual hemifields. First, we investigated the interplay between spatial and nonspatial attention by comparing conditions of divided attention between categories and/or locations. We then assessed the contribution of top–down processes versus stimulus-driven signals by separately testing the effects of divided attention in target and nontarget trials. The results revealed activation of a bilateral frontoparietal network when dividing attention between the two object categories versus attending to a single category but no main effect of dividing attention between spatial locations. Within this network, the left dorsal premotor cortex and the left intraparietal sulcus were found to combine task- and stimulus-related signals. These regions showed maximal activation when participants monitored two categories at spatially separate locations and the scene included a nontarget object. We conclude that the dorsal frontoparietal cortex integrates top–down and bottom–up signals in the presence of distractors during divided attention in real-world scenes. ■

INTRODUCTION

In many everyday life situations, such as driving a car, it is necessary to concurrently pay attention to multiple events/objects in different locations, such as a car coming toward us from the left while a pedestrian is about to cross the road to our right. Several studies have demonstrated that attention can flexibly handle multiple stimuli at non-adjacent spatial locations (McMains & Somers, 2004; Tong, 2004; Awh & Pashler, 2000; Castiello & Umiltà, 1992). However, most research on divided spatial attention have only made use of artificial experimental stimuli, such as simple geometrical shapes. How we divide attentional resources in more complex, natural environments remains largely unexplored.

Although dividing attention between spatial locations might be advantageous for several everyday life activities, extensive evidence indicates that this comes at a cost (Bichot, Cave, & Pashler, 1999; Castiello & Umiltà, 1990). Most of the previous studies on divided spatial attention contrasted conditions when participants attended to a single stream of stimuli in one visual hemifield with conditions when the participants attended to two streams in separate hemifields (McMains & Somers, 2005; Castiello & Umiltà, 1990; Eriksen & St. James, 1986). This approach leaves open the question of whether any resulting costs reflect processes specifically associated with divided spatial attention or more general interference effects because of the simultaneous monitoring of multiple streams of information (dual-task performance). In a previous fMRI study, we had addressed this issue by manipulating the number of attended positions and the number of task-relevant stimulus categories, operationally defining dual-task performance as attending to multiple objects at the same time, irrespective of spatial position (Fagioli & Macaluso, 2009). The results highlighted the role of the dorsal fronto-parietal attention network, where monitoring two locations interacted with the monitoring of two objects. These findings led us to suggest that nonspatial processes (e.g., dual-task interference) might contribute to mechanisms underlying divided spatial attention in the dorsal fronto-parietal cortex. However, this previous study employed simple stimuli (geometrical shapes) leaving the question open as to whether analogous interactions between spatial and nonspatial processes also take place with more naturalistic visual scenes, where prior knowledge and visual experience are likely to influence mechanisms of attention control (e.g., see Wolfe, Alvarez, Rosenholtz, Kuzmova, & Sherman, 2011).

Although classic research on visual attention suggests that the efficacy of visual search depends on simple attributes such as the number of distractor items or their similarity with the target (Wolfe & Horowitz, 2004; Wolfe,

[1]IRCCS Santa Lucia Foundation, Rome, Italy, [2]Lyon Neuroscience Research Center

1998; Treisman & Gelade, 1980), evidence from studies using real-world scenes indicate that visual search in cluttered naturalistic environments can be remarkably fast despite the overwhelming amount of information (Peelen & Kastner, 2011; Wolfe et al., 2011). This suggests that factors characterizing real-world situations, such as prior knowledge and expectations of objects in real-world contexts—in other words, visual experience—affect attention control in such naturalistic conditions (Peelen & Kastner, 2014; Wu, Wick, & Pomplun, 2014). In particular, there are at least two ways in which visual experience might influence attention control. First, real-world objects are recognized more efficiently when they are familiar, a phenomenon termed "ultra-rapid categorization" (Thorpe, Fize, & Marlot, 1996). For example, participants can detect the presence of highly familiar object categories, such as animals or vehicles, with a single glance (Li, VanRullen, Koch, & Perona, 2002; VanRullen & Thorpe, 2001a; Thorpe et al., 1996). A possible explanation of this is that object categorization occurs in a pre-attentive manner, with little requirements of top–down control (Li et al., 2002; Rousselet, Fabre-Thorpe, & Thorpe, 2002; VanRullen & Thorpe, 2001b). Second, prior knowledge about the spatial arrangement of objects within natural scenes—for example, we expect cars to be located on the street, rather than floating in the sky—contributes to guiding visual search (Wu et al., 2014). Furthermore, prior information about the visual context can facilitate the recognition of targets embedded in complex real-world displays, also known as the "contextual cueing effect" (Chun, 2000; Hollingworth & Henderson, 1998).

Here, we investigated the interplay between spatial and nonspatial divided attention while viewing images of natural scenes. As in our previous study (Fagioli & Macaluso, 2009), in different blocks, we instructed participants to monitor either one or two relevant object categories (now, car and/or person), which appeared either in one or both visual hemifields (left and/or right side of images). This allowed us to test the effects of divided spatial attention (one location vs. two locations) and of divided nonspatial attention (one category vs. two categories) and, critically, the interaction between these two types of top–down control. Here, we hypothesized that information about the spatial organization of real-world scenes, as well as the fast categorization of familiar, real-world objects, would affect the interplay of spatial and nonspatial processes when attention is divided while viewing natural scenes. Specifically, we expected that the efficient selection of real-world object categories would reduce demands on top–down control. This in turn would reduce the influence of nonspatial processes on mechanisms of divided spatial attention and the interaction between spatial and nonspatial factors in the dorsal fronto-parietal cortex (Fagioli & Macaluso, 2009).

On the other hand, fast object categorization might have a detrimental effect when, rather than detecting a target, the participants are required to filter out irrelevant objects. Previous behavioral studies have shown that object categorization is impaired when real-world scenes contain several foreground distractor objects (Walker, Stafford, & Davis, 2008; but see Lavie, Beck, & Konstantinou, 2014, and Cohen, Alvarez, & Nakayama, 2011, about the effect of task load on real object categorization). The role of distractor processing on attentional selection has been extensively investigated (e.g., see Duncan & Humphreys, 1989). Stimuli that are irrelevant to the task at hand, but are nevertheless within the same "attentional task set," have been associated with mechanisms of attention capture (contingent capture of attention; see Folk, Remington, & Wright, 1994; Folk, Remington, & Johnston, 1992). In the framework of visual search tasks with naturalistic stimuli, Reeder and Peelen (2013) asked participants to detect real-world object categories (i.e., car or people) in natural scenes. In a subset of the trials, participants responded to a lateralized dot probe that was cued by task-irrelevant stimuli (silhouettes of cars or people). Participants responded more quickly when the probe was presented in the same location as the template-matching cue compared with when it was presented on the opposite side (Reeder & Peelen, 2013). Contingent attention capture was only observed for category exemplars (silhouettes of the object) or characteristic parts of the object (such as the wheels of a car or limbs of a person) but not for objects semantically related to the cued category (e.g., air freshener, car radio, bracelets, hats). These findings suggest that attentional capture in naturalistic scenes is contingent on a search template composed of visual features characterizing the target (Reeder, van Zoest, & Peelen, 2015; Evans & Treisman, 2005). Several studies have suggested that different neural structures are responsible for target and distractor processing, namely, in frontal and parietal areas, respectively. For instance, it has been shown that the attentional selection of relevant stimuli relies on the activity of frontal areas (e.g., FEFs), whereas the suppression of competitive distractors is mediated by parietal areas, including the intraparietal sulcus (IPS) and the TPJ (Painter, Dux, & Mattingley, 2015; Akyurek, Vallines, Lin, & Schubo, 2010; Geng & Mangun, 2009).

Therefore, it is likely that the process of correctly identifying a target category in a real-world scene relies both on top–down control (e.g., implementation of an attentional template) and the bottom–up or stimulus-driven capture of attention by irrelevant objects in the scene. In the current study, we differentiated the overall effect of (top–down) task from the impact of the task set on the processing of distractors by considering separately the activity associated with divided spatial/nonspatial attention in target and distractor trials. Specifically, each trial/photo included only a single object (a car or a person) on the left or right side of the scene. A given object in a given position (e.g., a car on the left side) was then either a "target" or a "distractor," depending on the task instructions. This allowed us to ask whether any activation of frontoparietal regions reflects endogenous processes

associated with the maintenance of the current task set (which should lead to common activation both in target and distractor trials) or instead reflects more specific operations triggered during one or another type of trial (e.g., detection of the targets vs. filtering of the distractors). We predicted that top–down and bottom–up processes would jointly influence stimulus processing (Macaluso & Doricchi, 2013; Beck & Kastner, 2009; Indovina & Macaluso, 2007) and that bottom–up mechanisms would capture attention toward objects that share task-relevant characteristics with the target. The latter implies filtering and selection operations specific to trials with distractors.

In summary, we presented participants with natural scenes including either one car or one person, on the left or right side of visual scenes. In different blocks, we manipulated top–down voluntary attention, asking participants to monitor one of or both visual hemifields (focused vs. divided spatial attention, respectively) and to detect one or two object categories at the attended location/s (one relevant category vs. two relevant categories). The factorial crossing of spatial and nonspatial tasks enabled us to disentangle any effect associated with divided spatial attention (attending to two nonadjacent positions vs. a single position) from those associated with dual-task performance (attending to multiple object categories vs. one object category) and to highlight any interaction between these two types of top–down selective attention processes (see also Fagioli & Macaluso, 2009). Moreover, our design allows us to assess the interplay between these top–down task-related effects and stimulus-driven mechanisms of attention control by assessing the effects of the divided attention tasks both in target and distractor trials.

METHODS

Participants

Sixteen right-handed healthy volunteers (seven men, mean age = 28.75 years, range = 21–37 years) took part in the experiment. Participants gave written informed consent. The study was approved by the independent ethics committee of the Santa Lucia Foundation (Scientific Institute for Research Hospitalization and Health Care).

Paradigm

In all conditions, participants were presented with photos of real-world scenes (cityscapes and landscapes; see also next section below). These scenes contained one of two possible study-relevant object categories (i.e., people or cars), located either on the left or right side of the scene. A subset of the scenes (20%) did not contain any study-relevant object category (catch trials).

On a block-by-block basis, participants were instructed to monitor one or two object categories, either in one hemifield or in both hemifields, thus performing one of four distinctive attentional tasks (Figure 1A):

(A) Attend to one object category in one hemifield (Foc1); for example, "respond only to cars on the left side of the picture"

(B) Attend to both object categories in the same hemifield (Foc2); for example, "respond to both cars and people on the left side"

(C) Attend to one single object category, but monitor both hemifields at the same time (Div1); for example, "respond to cars, both on the left and on the right side of the picture"

(D) Attend to one object category in one hemifield and the other object category in the opposite hemifield (Div2); for example, "respond to cars presented on the left and to people presented on the right"

In addition, as each scene contained no more than a single study-relevant item (car or person), each of the four attention tasks could be further categorized into different trial types. For example, depending on the content of the picture, trials in the Foc1 condition can be categorized as follows: (1) target (T), an object of the relevant category was presented at a task-relevant/attended location; (2) object (O) distractor, an object of the relevant category was presented at the unattended location; (3) space (S) distractor, an object of the irrelevant category was presented at the attended location; or (4) catch (C) trial, the picture did not include any object of the study-relevant categories (see Figure 1B, top).

Thus, O-distractor trials contained an object that matched the category dimension of the current target template, as defined by the task instructions, whereas S-distractor trials contained an object that matched the position dimension of the target template. O distractors can only exist when attention was focused on one single location (Foc1 and Foc2), and S distractors existed only when participants were asked to monitor one single category (Foc1 and Div1). In the Div2 task, the distractor conditions entailed the presentation of a task-relevant object but at a task-incorrect, albeit attended, location (e.g., a car on the left while monitoring for left-people and right-cars). These trials are labeled as O&S distractors, because the distractor object matched both a relevant category and a relevant position of the target template. In total, the experiment included 13 possible trial types (see Figure 1B, bottom). The task of the participants was to press a response button as quickly as possible whenever they were presented with a target (T).

These conditions allowed us to address two main research questions. First, we investigated the effects associated with top–down divided attention by comparing the four task conditions (Div2, Div1, Foc2, and Foc1), irrespective of trial type/picture presented during the task. We tested for the main effects of divided spatial attention (Div > Foc, i.e., effect of monitoring two locations vs. one location) and of monitoring two categories versus

Figure 1. (A) Example displays used to instruct the participants on the four attention tasks. The Div2 task required monitoring of two different object categories in both visual hemifields (in the example, "attend cars on the left and people on the right side of the scene"). The Div1 task required monitoring of the same object category in both hemifields (in the example, "attend cars on both the left and the right side of the scene"). The Foc2 task required monitoring of two different object categories in the same hemifield (in the example, "monitor cars and people on the left side of the scene"). Finally, the Foc1 task required monitoring of one object category in one

hemifield (in the example, "attend cars on the left side of the scene"). (B, top) An example showing the four possible trial types associated with one of the Foc1 tasks: (1) target (T), relevant object category at the attended location; (2) distractor (O), relevant object category at the unattended location; (3) distractor (S), nonrelevant object category at the attended location; (4) catch (C), no study-relevant object in the scene. (B, bottom) The 13 trial types as defined by the combination of the four attention tasks and the object present in the scene. The O&S distractors in the Div2 task entailed an object of a relevant category presented at an attended but incorrect side (e.g., a car on the left, while monitoring the left-people and right-cars; see also Methods). (C) Mean RTs (in milliseconds) for the four attention tasks. Error bars are *SEM*.
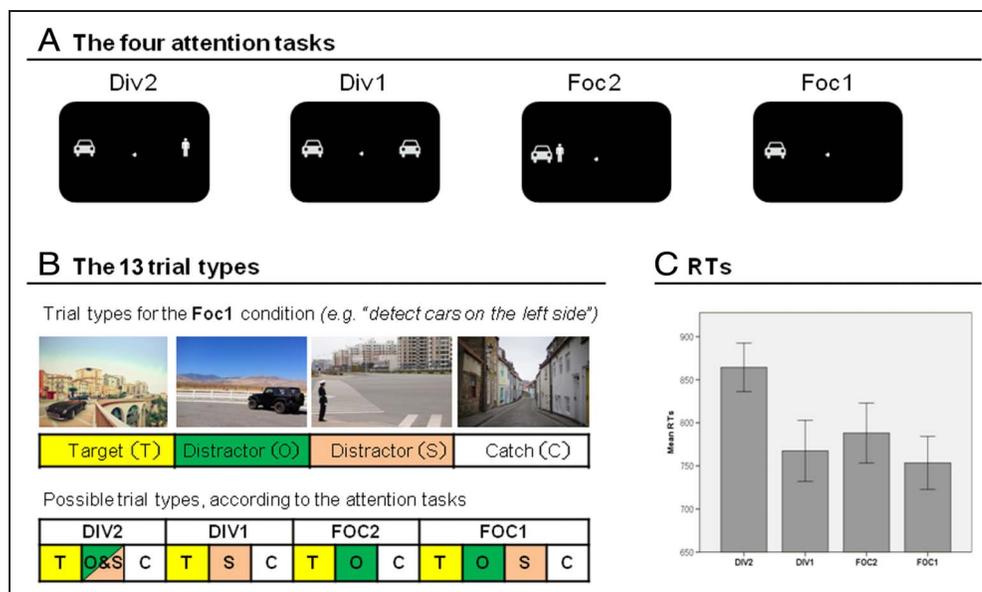
one object category (2 > 1). We also tested the interaction between these two types of top–down control processes to assess the contribution of nonspatial processes (here, an increase of the number of relevant categories) on brain activity associated with divided spatial attention (see also Fagioli & Macaluso, 2009).

Second, we asked to what extent any effect of divided attention was dependent on the specific trial type/picture, that is, the interaction between top–down and stimulus-driven signals for attention control in natural scenes. For this, we retested the relevant main effects and interactions but now separated target trials from distractor trials. These tests allowed us to assess whether any task-specific effect was related to the selection of the targets (T-trials) or the filtering out of the distractors (O&S-, O-, and S-distractor trials) or represented a stimulus-independent effect associated with the current task set (common to T-distractor and all distractor trials). It should be noted that, because the attention task determined the distractor type (O&S, O, or S; see Figure 1B), the comparisons involving the distractor trials entailed differences in terms of both the attention task and the distractor type (space and/or location matching; see also section above). However, because the attention tasks were performed both with targets and with distractors, we could identify any distractor-specific effect by analyzing the interactions between the attention task type (space/category divided attention) and the stimulus type (target vs. distractors). With this, we could confirm that any distractor-specific

effects were indeed driven by distractor processing rather than differences in the attention tasks.

Stimuli and Task

Participants lay in the scanner in a dimly lit environment. All visual stimuli were projected onto a translucent screen at the back of the MR bore and were visible through a mirror mounted on the head coil. At the beginning of each block, an instruction display informed participants about the current attention task (i.e., relevant location/s and object category/s). The instruction display (duration = 3 sec) consisted of a white central fixation point displayed on a black background, plus one or two white stick figures (depicting a car or a man) that indicated the relevant side and the relevant object category for the upcoming block of trials (figure width = approx 1.5°, centered at 6° visual angle from central fixation). For example, the Foc1 task was prompted by a single figure (car or person) displayed on the left or right side of central fixation, whereas the Div2 instruction was composed of two figures (the car and the person), one displayed on the left side and the other on the right side of fixation (see also Figure 1A).

Photos were selected from the SUN database (Xiao, Hays, Ehinger, Oliva, & Torralba, 2010), the Graz-02 Database (Marszałek & Schmid, 2007; Opelt, Pinz, Fussenegger, & Auer, 2006), and the Internet. Three hundred eighty-four photographs were selected for the experiment. Specifically, we chose pictures that only contained a single object of a

study-relevant category (i.e., a car or a person), located either on the left or right side of the image and not crossing the midline. The position, viewpoint, and size of the people and cars in the pictures were highly variable, mimicking real-world viewing conditions. To rule out the possibility that the task performance was biased by the specific arrangement of the objects in the scene, each picture was used twice over the entire experiment, once as a target and once as a distractor. This allowed us to minimize any possible stimulus-related confound on the task-related effects.

The final set of photographs contained pictures with a person on the left ($n = 64$) or right ($n = 64$) side and a car on the left ($n = 64$) or right ($n = 64$) side, plus 128 additional photographs without any foreground object and any car or person. The latter were used for the catch trials. The assignment of the pictures to the different task conditions and trial types was counterbalanced across participants to minimize the possibility that some uncontrolled characteristics of the pictures systematically affected our comparisons between conditions.

Each task block was composed of 10 trials. In each trial, a photo was displayed centrally for 50 msec (horizontal visual angle = 20°), followed by an ISI ranging from 2000 to 4000 msec. We used a brief stimulus presentation time to discourage any overt eye movements (Jans, Peters, & De Weerd, 2010). Each 10-trial block included four targets (T), four distractors (O&S, O, and/or S), and two catch (C) trials, randomly intermixed within each block.

All participants underwent four fMRI scanning runs (lasting approximately 12 min each). Every fMRI run was composed of 16 blocks. Over the entire experiment, each participant was presented with 640 trials: 512 trials with pictures including a car or a person (i.e., each of 256 photos with objects presented twice, once as a target and once as a distractor), plus 128 catch trials with photos that did not include any car or person.

Image Acquisition and Preprocessing

Functional images were acquired with an Allegra scanner operating at 3 T (Siemens, Erlangen, Germany). BOLD contrast was obtained using T2*-weighted EPI. The acquisition of 32 transverse slices (2.5 mm thick, 50% distance factor), with a repetition time of 2.08 sec, provided coverage of the whole cerebral cortex. The in-plane resolution was 3 × 3 mm.

The fMRI data were processed with SPM8 (www.fil.ion. ucl.ac.uk). The first four image volumes of each run were discarded to allow for stabilization of longitudinal magnetization. For each participant, the remaining 1480 volumes were realigned with the first volume, and the acquisition timing was corrected using the middle slice as reference. To allow interparticipant analysis, all images were normalized to the Montreal Neurological Institute standard space (Collins, Neelin, Peters, & Evans, 1994), using the mean of

all 1480 images. All images were smoothed using an isotropic Gaussian kernel (FWHM = 8 mm).

Data Analysis

Statistical inference was based on a random effects approach (Penny & Holmes, 2003). This was composed of two steps. First, for each participant, data were best-fitted (least square fit) at every voxel using a linear combination of the effects of interest. These were the onsets of the stimulus presentation for each of the 13 event types: Div2 T, O&S, and C; Div1 T, S, and C; Foc2 T, O, and C; and Foc1 T, O, S, and C (see Figure 1B, bottom). In addition, the model included one regressor modeling the instruction display, one for error trials irrespective of condition, and the six head movement parameters estimated during realignment. These additional effects were not considered further in the group analyses. All event types were convolved with the SPM8 standard hemodynamic response function. Linear compounds (contrasts) were used to determine the effect of the 13 relevant trial types, averaged across the four fMRI runs. The corresponding 13 contrast images (per participant) were entered in a within-participant ANOVA for statistical inference at the group level. Correction for nonsphericity (Friston et al., 2002) was used to account for possible differences in error variance across conditions and any nonindependent error terms from repeated measures. Within the group-level ANOVA, we addressed our two main research questions.

First, we assessed the main effects of dividing attention between two locations and between two categories and the interaction between these two factors. Because of the different number of distractor types available for each task (see Figure 1B), the contrast weights averaged the effects of O and S distractors in the Foc1 condition. For example, for the main effect Div > Foc, we used the contrast weights [1 1 1 1 1 1 −1 −1 −1 −1 −0.5 −0.5 −1] (see Figure 1B, bottom, for the order of the 13 trial types). The SPM threshold was set to $p < .05$, FWE-corrected at the cluster level (cluster size estimated a p uncorrected = .001), considering the whole brain as the volume of interest.

Second, we asked whether any task-related effect was modulated by trial type (i.e., targets vs. distractors). Accordingly, we retested the effects of dividing attention between two locations and between two categories and the interaction between these two factors, but now separately for the three trial types. To statistically evaluate any differences because of trial type, we formally tested for the two-way interaction between the effects of the divided attention tasks (Div > Foc and 2 > 1) and trial type (target vs. distractor) and the three-way interaction among the three factors. Again, all contrasts averaged the weights of O and S distractors in the Foc1 condition; for example, for the three-way interaction, the weights were [−1 1 0 1 −1 0 1 −1 0 −1 0.5 0.5 0] (see Figure 1B).

The SPM threshold was set to $p < .05$, FWE-corrected at the cluster level (cluster size estimated a p uncorrected = .001), considering the whole brain as the volume of interest.

Additional fMRI Analyses with Central Fixation Controlled

The eye position was recorded during fMRI using an ASL eye-tracking system, adapted for use in the scanner (Applied Science Laboratories, Bedford, MA; Model 504, sampling rate = 60 Hz). Eye position traces were examined in a 500-msec window, starting at the onset of the picture. Losses of fixation were identified as changes in horizontal eye position greater than 2° of the visual angle. The eye position traces were adjusted only with respect to the overall mean position across the entire scanning session, that is, without removing any eventual condition-specific biases (e.g., participants sustaining gaze toward the relevant hemifield throughout blocks of focused left/right attention). Because of technical difficulties, we were able to obtain reliable eye position data in only eight participants. For these participants, we could confirm central fixation in at least two thirds of the trials, at least in three of the four fMRI runs.

For these participants, we performed an additional control fMRI analysis that considered three fMRI runs and modeled any loss of fixation as a separate event type. For the second-level ANOVA, we used the parameter estimates of the 13 conditions of interest that now included only trials with fixation controlled. For each participant, linear contrasts were used to average the condition-specific parameter estimates across the three runs. With these additional analyses, we sought to replicate our main results, while including only trials with central fixation confirmed. Note that these tests were nonindependent of our main analyses and included only a small pool of participants. Thus, we report only peaks of activation (p uncorrected < .001) located within the clusters that we identified in our main analysis (see Table 1, rightmost columns) and emphasize that the corresponding results are confirmatory only. Future studies should attempt to obtain more reliable eye-tracking data, aiming to perform the main statistical analyses (full sample size) with central fixation controlled at the single-trial level in all participants.

RESULTS

Behavioral Data

The mean (*SEM*) RTs to the target stimuli for the four attention tasks were 864 (14) msec in Div2, 767 (18) msec in Div1, 788 (17) msec in Foc2, and 753 (15) msec in Foc1. A two-way ANOVA, considering the number of monitored positions (i.e., divided vs. focused spatial attention, Div vs. Foc) and number of monitored categories (2 vs. 1), re-vealed a main effect of Position ($F(1, 15) = 57.3, p < .001$) and a main effect of Category ($F(1, 15) = 91.1, p < .001$). Participants were slower when attending to two spatial positions compared with one position as well as when monitoring two categories compared with one category (see Figure 1C). Moreover, the analysis showed a significant interaction term ($F(1, 15) = 33.7, p < .001$): The participants were significantly slower in monitoring two categories at two nonadjacent locations (Div2) than in the other attention tasks (Div2 vs. Div1, Foc2, and Foc1; all ps < .001). The post hoc analysis also revealed that, during focused spatial attention, monitoring one category was easier in terms of RTs than monitoring two categories (Foc2 vs. Foc1, $p < .001$). No other effect reached statistical significance.

Overall, the accuracy of the target detection task was high (95%). The error rates (false alarms + missing responses) for the four attention tasks were as follows: Div2, 11% (2%); Div1, 7% (1%); Foc2, 6% (1%); and Foc1, 7% (1%). Analysis of the error rates revealed a main effect of Position ($F(1, 15) = 16.3, p = .001$) and a significant interaction term ($F(1, 15) = 15.9, p = .001$). Consistent with the RT data, post hoc analyses showed that participants were less accurate in the Div2 task than in the other three attention tasks (Div2 vs. Div1, Foc2, or Foc1; all ps < .05).

fMRI Analyses

Dividing Attention in Space and across Object Categories

First, we compared the effects of paying attention to both hemifields versus one hemifield (divided spatial attention) and monitoring two categories versus one category (divided category-based attention), irrespective of the trial type (target, distractor, or catch). Unlike our previous study with nonnaturalistic stimuli (Fagioli & Macaluso, 2009), we now found that the frontoparietal network activated only for divided category-based attention but not for divided space-based attention (Figure 2A). The effect of divided category-based attention (2 > 1) showed a cluster of activation in the left premotor cortex, extending from the lateral surface of the hemisphere to the depth of the precentral sulcus and likely including the FEFs. The activation also extended ventrally to the ventral premotor areas and rostrally to the pFC. Another cluster was located in the pars triangularis of the left inferior frontal gyrus (IFG; Table 1). In the parietal lobe, a significant cluster of activation was found in the IPS, extending to the posterior and anterior parietal lobes, bilaterally in both hemispheres. By contrast, the main effect of attending to two separate spatial positions versus spatially focused attention (Div > Foc) did not show any significant effects.
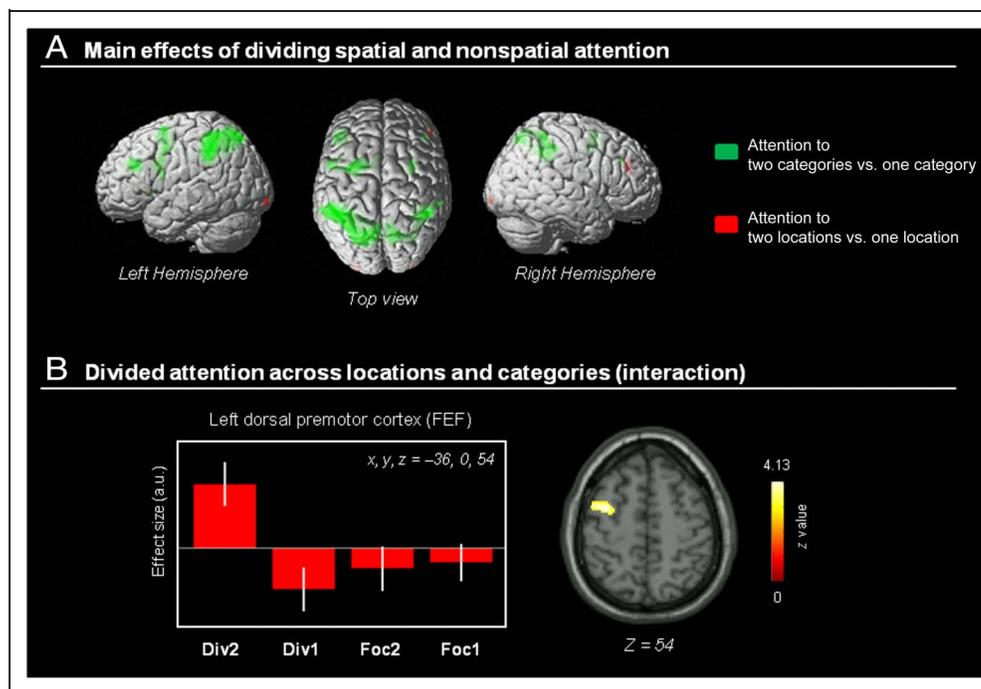
For completeness, we computed the effects of dividing category-based attention separately under spatially

Table 1. Anatomical Location and Statistical Scores for the Regions That Activated during Selection of Multiple Categories and Multiple Locations

| | *All Trial Types* | | | | | | *Target Trials Only* | | | | *Distractor Trials Only* | | | |
| | *Main Analysis* | | | | *Fixation Controlled* | | | | | | | | | |
	p *Corrected*	*Size*	*Coord*	*z* *Value*	*Coord*	*z* *Value*	*p* *Corrected*	*Size*	*Coord*	*z* *Value*	*p* *Corrected*	*Size*	*Coord*	*z* *Value*
Main effect of dividing attention between categories														
Left aIPS	<.001	5146	−46, −42, 46	5.83	−40, −48, 56	3.33	<.001	5115	−48, −42, 50	5.32	<.001	1530	−46, −42, 52	5.42
Left PPC			−20, −72, 56	5.66					−16, −72, 56	4.56			−10, −68, 64	3.88
Right PPC			14, −66, 60	4.33					10, −70, 48	4.38				
Right aIPS			46, −42, 38	4.20			<.001	849	32, −38, 44	4.07	.002	531	34, −44, 42	4.28
Left FEF	<.001	1160	−28, −2, 46	4.86	−26, 4, 50	3.23	<.001	2950	−26, −2, 50	4.61	.008	412	−28, 4, 62	4.32
Left PreC			−48, 4, 20	3.92					−42, 2, 44	4.89			−28, −2, 46	4.22
Left IFG	.022	322	−44, 32, 28	4.57	−48, 32, 32	3.43			−40, 36, 28	4.89				
Interaction between spatial and nonspatial divided attention														
Left FEF	.010	395	−36, 0, 54	4.13	−40, 4, 56	3.35					<.001	2211	−32, 4, 52	5.81
Left PreC			−50, 2, 46	4.00	−38, 2, 40	3.49							−44, 2, 36	5.31
Left pIPS											<.001	1298	−28, −56, 44	4.72

The *p* values are corrected for multiple comparisons at the whole-brain level (cluster-level correction: cluster size estimated at *p* uncorrected = .001). Cluster sizes are in number of voxels. Blank spaces indicate that the corresponding area did not show any significant effect at the relevant threshold (*p* corrected = .05, for the main analyses; *p* uncorrected = .001, for the additional and nonindependent analyses with fixation controlled). PPC = posterior parietal cortex; pIPS = posterior IPS; PreC = precentral gyrus.

Figure 2. (A) Main effects of divided spatial and nonspatial attention. Three-dimensional rendered projections of the activations associated with monitoring two object categories versus one object category (green blobs) and monitoring two locations versus one spatial location (red blobs). Activations are rendered at p uncorrected = .001, showing also a few nonsignificant voxels for the main effect of monitoring multiple locations (cluster size threshold for the display = 200 voxels). Please see Table 1 for the statistics associated with the significant main effect of monitoring multiple object categories versus monitoring a single category. (B) Horizontal section and signal plot for the region in the dorsal premotor cortex (FEF) that showed a significant interaction between divided spatial and nonspatial attention. The signal plot shows that the interaction effect was driven by high activity when participants were asked to monitor two different object categories at two nonadjacent locations (Div2 task; Bar 1). The level of activation for the four attention tasks is mean adjusted (i.e., the four values sum to zero) and is expressed in arbitrary units (a.u.; ±90% confidence interval). SPM display threshold: p uncorrected = .001, minimum cluster size = 300 voxels.

focused attention (Foc2 > Foc1) and under spatially divided attention (Div2 > Div1). The Foc2 > Foc1 comparison did not reveal any significant activations. At an uncorrected threshold ($p < .001$), we found bilateral activation of the posterior parietal cortex bilaterally ($x, y, z = -20, -72, 54$; z value = 4.16; $x, y, z = 28, -72, 54$; z value = 3.75), consistent with the main effect of monitoring two categories versus one category in these regions (see Figure 2A and Table 1). The Div2 > Div1 comparison revealed a bilateral cluster of activation in the anterior IPS (left aIPS: $x, y, z = -32, -56, 52$; z value = 5.76, p corrected < .001; right aIPS: $x, y, z = 34, -46, 46$; z value = 3.70, p corrected = .004) and FEFs (left FEF: $x, y, z = -28, 0, 48$; z value = 5.66, p corrected < .001; right FEF: $x, y, z = 30, 0, 48$; z value = 4.60, p corrected = .046).

Thus, dividing spatial attention in naturalistic scenes appears to have little or no impact on the level of activity of the dorsal frontoparietal cortex or other regions in the brain. However, the behavioral data showed a clear cost of dividing spatial attention, particularly when monitoring two categories (see Figure 1C). Indeed, the corresponding comparison in the imaging analyses (i.e., the interaction between space- and category-divided attention) revealed a significant effect in the left premotor cortex. In this region, the BOLD signal increased specifically when participants monitored two categories at two non-contiguous locations (Div2 task). In the premotor cortex of the right hemisphere, the same effect was observed

only at a lower statistical threshold ($x, y, z = 30, -2, 48$; z value = 2.96, p uncorrected < .005). The pattern of activity for the left premotor cortex is shown in Figure 2B. The signal plot shows maximal activation in Div2 condition (Bar 1 in Figure 2B), driving the interaction between location-based and category-based attention in this region.

These results were confirmed by the additional, but nonindependent, fMRI analysis with central fixation controlled (see Methods). For the main effect of dividing attention across object categories, the additional analysis highlighted the activation of three clusters located within the left aIPS, the left FEF, and the left IFG (see Table 1, "All Trial Types/Fixation Controlled" column). The analysis with fixation controlled confirmed the activation of the left premotor cortex associated with the interaction between location-based and category-based divided attention (see Table 1, bottom).

Distractor Processing: Modulation According to Trial Type

After having examined the effect of dividing attention across multiple spatial locations and categories with target, distractor, and catch trials pooled together, we retested the main effect of dividing attention between two categories (see Figure 2A) and the interaction between location-based and category-based divided attention (see Figure 2B), separately for target and distractor trials. This allowed us

to assess whether the effects associated with divided attention highlighted in the section above are related to the current task set (common to target and distractor trials), the selection of the task-relevant stimuli (specific for target trials), or the filtering out of irrelevant stimuli (specific for distractor trials).

The contrasts testing for the main effect of monitoring multiple categories separately in target and distractor trials revealed that regions within the frontoparietal network activated both when the trial included a target or a distractor (see Table 1, "Target Trials Only" and "Distractor Trials Only," top). This is consistent with the hypothesis that the effect of object-based divided attention in these areas was primarily driven by the current attentional task set, rather than any specific operations related to the detection of the targets or the filtering of the distractors.

By contrast, the activation associated with the interaction between category-based and space-based divided attention was significant only for the distractor trials (see Table 1, bottom). We confirmed the specificity of this effect for the distractor trials to be beyond any effect associated with the attention tasks by formally testing the three-way interaction among the two effects of dividing attention and the trial type (target vs. distractor; see also Methods). This demonstrated that, while monitoring two stimulus categories at two nonadjacent locations (Div2), activity in both the left frontal and left parietal regions was greater during distractor than target trials. Specifically, clusters of activation within the frontal lobe were located in the left FEFs ($x, y, z = -32, 4, 52$; z value = 4.82, p corrected = .039) and in the left precentral gyrus ($x, y, z = -46, 6, 34$; z value = 4.35, p corrected = .001). In the

parietal cortex, the cluster of activation was located in the left posterior IPS and extended to the left posterior parietal cortex ($x, y, z = -30, -60, 38$; z value = 4.18, p corrected = .002). Whereas significant effects were restricted to the left hemisphere, analogous patterns were detected also in the right hemisphere, but at a lower statistical threshold (FEF: $x, y, z = 36, 12, 58$; z value = 3.61, p uncorrected < .001; IPS: $x, y, z = 32, -60, 40$; z value = 2.87, p uncorrected < .005).

The signal plots in Figure 3 show the effects of the attention tasks in the left FEF and in the left posterior IPS, separately for target and distractor trials. In distractor trials, the task of attending to multiple categories at noncontiguous spatial locations was associated with a selective increase of the BOLD signal, over and above the other three attention tasks (in Figure 3, compare Bar 1 with Bars 2, 3, and 4 in each left panel of the "Distractor" plot). Instead, in the target trials, the increase of activity within the frontoparietal network was not selectively modulated by the Div2 task but was associated with the monitoring of two categories versus one category (compare Bar 1 with Bar 2 and Bar 3 with Bar 4 in each right panel of the "Target" plot). These results indicate that the involvement of these regions in the Div2 condition reflected some operation required selectively during distractor processing, such as the filtering out of nontarget stimuli presented at a relevant location under conditions of divided attention (see Discussion section).

The results of the three-way interaction among the two effects of dividing attention and trial type were confirmed by the additional, but nonindependent, fMRI analysis with central fixation controlled. This showed the activation of three clusters located within the left posterior

Figure 3. Horizontal section and signal plots for the two regions in the left dorsal frontoparietal cortex where the interaction between spatial and nonspatial divided attention was specific to the distractor trials. The signal plots show the pattern of activation in the left FEFs and the left posterior IPS (pIPS) separately for target and distractor trials. In distractor trials, there was a specific increase of activity when the participants attended to two different categories at noncontiguous spatial locations (see Div2 condition in the "Distractors" plots, in red). By contrast, in target trials, there was only a main effect of monitoring two object categories (compare Bar 1 with Bar 2 and Bar 3 with Bar 4 in the "Targets" plots, in green). The signal plots show the parameter estimates at the peak voxels of the three-way interaction among the two divided attention tasks and the trial types (see also main text). The parameter estimates are mean adjusted across the four conditions in each plot and are expressed in arbitrary units (a.u.; ±90% confidence interval). SPM display threshold: p uncorrected = .001, minimum cluster size = 250 voxels.
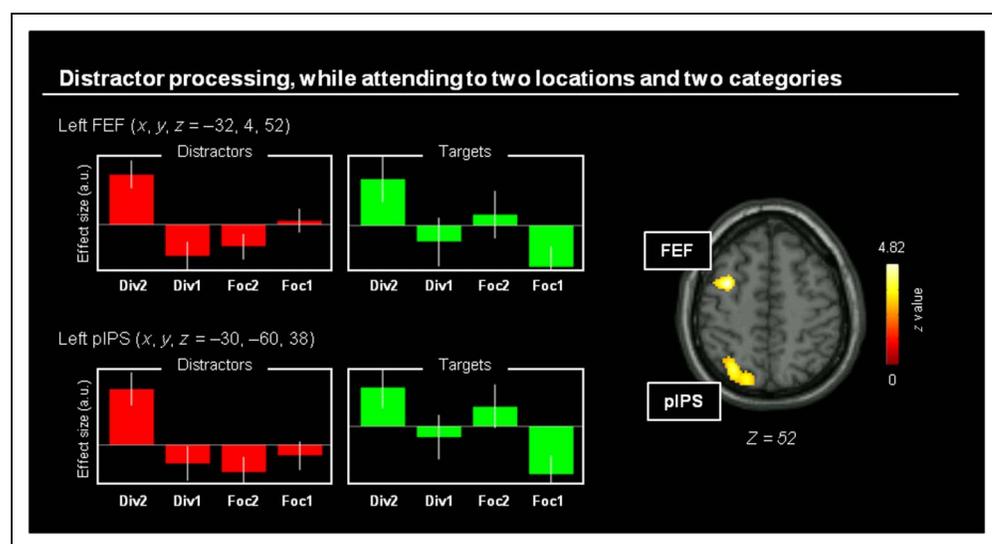
IPS, extending to the left posterior parietal cortex (x, y, z = −32, −66, 48; z value = 3.44), the left FEFs (x, y, z = −38, 4, 58; z value = 2.97), and the left precentral gyrus (x, y, z = −38, 8, 34; z value = 3.30).

DISCUSSION

We investigated divided spatial and nonspatial attention in naturalistic scenes and the interplay between these top–down task-related effects and stimulus-related signals associated with the content of the scenes (target vs. distractor trials). We found that areas located within the frontoparietal network activated when dividing attention between multiple object categories and that the interaction of location- and category-based attention involved the dorsal premotor cortex, which was maximally activated when participants were asked to monitor two different categories at spatially separate locations. When we considered the effects of task as a function of trial type, we found that the dorsal frontoparietal cortex showed an interaction between category-based and space-based divided attention specifically for distractor trials, that is, when the participants had to ignore a task-relevant object presented at the wrong location (e.g., a car presented on the left while monitoring for cars on the right and people on the left). These results indicate that divided attention under complex and naturalistic conditions involves greater requirements of selective processing to successfully filter out competing but potentially relevant distractor objects, which we associate with increased activation of the dorsal frontoparietal network.

Unlike our previous study with simple stimuli (Fagioli & Macaluso, 2009), here, regions of the dorsal fronto-parietal network activated primarily when participants divided attention between categories, with no overall effect of divided spatial attention (see Figures 2A and 3, "target trials" in green). This might appear to conflict with our initial expectations that prior knowledge and fast processing of real-world objects would reduce the impact of monitoring multiple streams (i.e., divided nonspatial attention) on the activity of regions engaged during divided spatial attention. A possible explanation for this might be that, when participants were asked to divide spatial attention while monitoring a single category (Div1 condition), fast object categorization permitted detecting the presence of a target object without engaging top–down voluntary resources to split attention between the two hemifields (Li et al., 2002; Rousselet et al., 2002). However, this advantage for the processing of real-world objects can turn into a cost when participants had to ignore distractor objects while monitoring two different categories at separate locations (O&S distractors, in Div2). Therefore, fast object categorization would lead to in-depth processing of stimuli that share features of the target such that a stimulus sharing both category and position dimensions of the target (i.e., object of a relevant category, presented at a relevant [but incorrect]

location) would require additional processing resources to be filtered out, as will be discussed further below.

The significant interaction found between spatial and nonspatial divided attention specifically in distractor trials rules out the simple explanation that the Div2 task entailed high demands both on attention control and on working memory (Nebel et al., 2005). The close relationship between working memory and visual attention, as well as their overlapping neural substrates in frontal and parietal regions (Silk, Bellgrove, Wrafter, Mattingley, & Cunnington, 2010; Corbetta, Kincade, & Shulman, 2002), has been extensively discussed (Chun, 2011; Theeuwes, Belopolsky, & Olivers, 2009; Awh, Vogel, & Oh, 2006; Awh & Jonides, 2001). However, any effect associated with the number of task rules or the maintenance of a complex attentional template (Desimone & Duncan, 1995; Duncan & Humphreys, 1989) should lead to a sustained and trial-type independent response throughout the entire block requiring attention to two positions and two categories (Div2 task). Instead, the current findings indicate that frontoparietal activity relates specifically to selection mechanisms associated with distractor stimuli, which we suggest are linked to the suppression of highly interfering nontarget objects, that is, distractors that belong to a task-relevant category presented at a task-relevant (but incorrect) location.

The specificity of the distractor effect for the Div2 condition points to a major role of task set. The Div2 task involved a complex combination of spatial and nonspatial dimensions. The rapid categorization of real objects can facilitate attentional selection under simple task constraints, such as conditions requiring the simple detection of foreground objects, which engages bottom–up/preattentive mechanisms. However, Evans and Treisman (2005) found that, when participants were asked not only to detect the target object but also to report its position in a scene, they frequently failed to localize the targets that they had correctly detected. These results suggest that bottom–up processing might not be sufficient to simultaneously encode both the category and spatial dimensions of objects in natural scenes. The same study also found that target detection was impaired by the presence of distractors that contained similar features as the target (e.g., detecting animals among humans, both animate objects, as opposed to detecting animals among vehicles), consistent with the view that distractors that overlap with target features require additional processing resources (Evans & Treisman, 2005).

Therefore, we argue that the distractor effect in Div2 condition results from the complex task set stemming from the high-level of similarity between the distractors and the targets: The Div2 distractors shared both (spatial) location- and (nonspatial) category-related dimensions with the targets. The impact of complex task sets on the processing of distractor stimuli has been recently studied by Painter, Dux, Travis, and Mattingley (2014). Their study made use of simple geometrical shapes and showed

that distractors presented at irrelevant/unattended locations but matching the color of the target evoked larger electrophysiological responses during conjunction search compared with searches for a single feature. These results highlight enhanced processing of distractors that match complex combinations of target-defining features and suggest increased attention demands needed to filter out such irrelevant stimuli. In our study, the Div2 distractors entailed the presentation of stimuli that matched both a relevant location and a relevant category, as defined by the current task set. We suggest that this combination required active feature-based suppression of irrelevant stimuli at attended locations. It should be noted that our design did not allow us to tease apart the full hierarchy of spatial and nonspatial signals that in principle might contribute to attentional selection of task-relevant stimuli. To investigate the full hierarchy of the relevant signals would involve asking participants to also detect targets of one or both categories irrespective of location and respond to all stimuli at one or both locations irrespective of category. Such design would then enable the assessment of how the inclusion of specific combinations of features in the task set modulates the response of the frontoparietal regions, for example, by comparing "detect car and people irrespective of side" versus "detect left-people and right-car" (Div2 condition here). However, the investigation of these hierarchical combinations of selection signals was beyond the aim of the current study, which instead focused on the interplay between spatial and nonspatial factors when monitoring two spatially separate locations (see also Fagioli & Macaluso, 2009).

In line with the suggestion that specific mechanisms might engage during distractor suppression, previous studies identified distinct patterns of activity associated with target enhancement compared with distractor suppression (Painter et al., 2015; Seidl, Peelen, & Kastner, 2012; Akyurek et al., 2010; Andersen & Muller, 2010). Painter and colleagues (2015) found that different parietal regions were associated with the capture of attention by distractor and target stimuli. Using TMS, they showed that stimulation of parietal regions, including IPS, selectively interfered with attentional capture by distractor stimuli, suggesting that distinct neural circuits underlie distractor and target capture. Similarly, an fMRI experiment found that distractor suppression activated parietal regions, whereas target enhancement activated frontal regions (Akyurek et al., 2010). Recently, using naturalistic stimuli, Seidl and colleagues (2012) found that detecting object categories that are relevant for the task at hand relies not only on the neural enhancement of the current "task set" but also on the active suppression of the previous attentional set that is no longer relevant for the current behavioral goal. Using fMRI, they asked participants to detect object categories (i.e., car, people, tree) embedded in natural scenes and measured the brain activity associated with the processing of (1) the current study-relevant object category (target category), (2) the

object category that was previously but no longer relevant (distractor category), and (3) an object category that was never study-relevant (neutral category). They found increased processing of the target category and decreased processing of the distractor category in higher-order visual areas (both relative to the neutral category), suggesting that, when the distractor objects strongly interfere with the identification of the target, then the visual system can act by suppressing their processing to prevent the erroneous selection of, or interference from, irrelevant objects (Seidl et al., 2012).

Our current data show that the process of suppressing irrelevant distractors relies on a specific neural network that does not apply to target stimuli, consistent with the existence of an active mechanism of distractor suppression. Thus, although the detection of familiar target objects in natural scenes might be relatively efficient (Li et al., 2002; Thorpe et al., 1996), the filtering out of familiar nontarget objects may require additional processes, especially when these share task-relevant characteristics with the target (the distractors in the Div2 condition; see also discussions above). The latter might entail active inhibition of eye movements toward distractor stimuli (Belopolsky & Theeuwes, 2012; Theeuwes, Kramer, Hahn, Irwin, & Zelinsky, 1999; Folk et al., 1992). Previous studies showed that task-irrelevant abrupt onsets elicit reflexive eye movements toward the location of the irrelevant visual event ("oculomotor capture"), followed by a corrective eye movement toward the target (Theeuwes et al., 1999). Notably, distractors sharing some task-relevant characteristics with the target lead to longer gaze dwell times (Mulckhuyse, Van der Stigchel, & Theeuwes, 2009; Ludwig & Gilchrist, 2003). These results have been interpreted as evidence of a top–down inhibitory process that is engaged to resolve the competition between the representation of the target and of the distractor (Born, Kerzel, & Theeuwes, 2011; Mulckhuyse et al., 2009).

In the current study, the Div2 condition entailed distractor objects belonging to a relevant category and presented at a relevant location, for example, picture including a car on the left, while participants monitored cars on the right and persons on the left. We propose that these highly interfering distractors generated an oculomotor program and that dorsal premotor cortex engaged in a time-consuming process needed to suppress spatial orienting toward these nontarget stimuli. In agreement with this, studies of overt visual search in nonhuman primates showed that neurons in FEF respond to distractor stimuli and that these responses are modulated according to the similarity of the distractors with the target stimuli (Thompson, Bichot, & Sato, 2005; Bichot & Schall, 1999). Neurons in the FEF show stronger responses to distractors similar to the target compared with distractors dissimilar to the target (Bichot et al., 1999). Moreover, enhanced activity in FEF neurons has been associated with curved saccades toward distractor locations, indicating that

gaze capture from irrelevant stimuli is triggered before competition between the target and distractor neural representations is resolved by the FEF (McPeek, 2006). Several studies have suggested that successful inhibition of unwanted eye movements relies on the FEF also in humans (Van der Stigchel, van Koningsbruggen, Nijboer, List, & Rafal, 2012; Curtis & D'Esposito, 2003; Kimmig et al., 2001).

An unexpected aspect of our current results was that the interactions between spatial and nonspatial divided attention were fully significant only in the left hemisphere (see Figures 2B and 3). The corresponding regions in the right hemisphere showed analogous effects but only at a lower, uncorrected threshold (see Results section). Top–down attention control in dorsal frontoparietal cortex is typically associated with bilateral patterns of activation (Corbetta & Shulman, 2002) or, if anything, lateralized to the right hemisphere (Nobre et al., 1997; Corbetta, Miezin, Shulman, & Petersen, 1993; Mesulam, 1981). Nonetheless, several previous studies demonstrated that the type of stimulus material can influence hemispheric dominance in attention control (Fink et al., 1997). For example, Fink and colleagues (1997) showed that the classical pattern of hemispheric lateralization during attention to hierarchical stimuli (i.e., right hemisphere for attended to global elements vs. left hemisphere for attended to local elements) fully reversed when using object- rather than letter-based material. In the current study, the stimulus material (i.e., natural scenes) entailed rich prelearned contextual information (e.g., cars are located on the street) that can play a role in attentional control (Wu et al., 2014; see also Introduction). The left hemisphere has been shown to play a dominant role in tasks that involve acquired conceptual knowledge (Seger et al., 2000) and when participants are explicitly asked to form thematic categories (e.g., respond "garage" when presented with the word "car"; Sachs, Weis, Krings, Huber, & Kircher, 2008). We propose that the predominant involvement of the left hemisphere in the current study might relate to the stimulus material, with participants making use of their knowledge about the location of cars and people in natural scenes when allocating attentional resources. The lack of any such pattern of lateralization in our previous study (Fagioli & Macaluso, 2009), which involved analogous attention tasks but different stimulus materials (streams of geometric shapes, without any contextual information or prior associations), would also fit this interpretation.

In summary, the current study demonstrated that participants can divide attentional resources between object categories and spatial positions in real-world scenes. We found that the task of simultaneously dividing attention between multiple categories and separate spatial positions incurred a behavioral processing cost. This might reflect the joint effect of attention and memory demands for the maintenance of a particularly complex task set, along with an increase of top–down control needed for the sup-

pression of irrelevant stimuli in cluttered natural scenes. The imaging results showed that dividing attention between multiple categories and positions involved the dorsal frontoparietal regions (FEF and IPS), indicating common control processes for object- and space-based selection in these areas. Most importantly, this effect was found to be specific to distractor trials, highlighting an interplay between top–down task-based control and stimulus-driven signaling related to the current sensory input. We suggest that, in complex and naturalistic scenes, category- and location-based selection depends both on the sustained enhancement of the attentional target template and on the ability to resist the bottom–up stimulus-driven capture of attention by irrelevant objects in the scene. In conclusion, we provide the first evidence that attentional resources can be effectively divided in real-world scenes and that the dorsal frontoparietal network integrates top–down and stimulus-driven signals to handle the competitive interactions arising during divided attention in these complex conditions.

REFERENCES

Akyurek, E. G., Vallines, I., Lin, E. J., & Schubo, A. (2010). Distraction and target selection in the brain: An fMRI study. *Neuropsychologia, 48,* 3335–3342.

Andersen, S. K., & Muller, M. M. (2010). Behavioral performance follows the time course of neural facilitation and suppression during cued shifts of feature-selective attention. *Proceedings of the National Academy of Sciences, U.S.A., 107,* 13878–13882.

Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends in Cognitive Sciences, 5,* 119–126.

Awh, E., & Pashler, H. (2000). Evidence for split attentional foci. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 834–846.

Awh, E., Vogel, E. K., & Oh, S. H. (2006). Interactions between attention and working memory. *Neuroscience, 139,* 201–208.

Beck, D. M., & Kastner, S. (2009). Top–down and bottom–up mechanisms in biasing competition in the human brain. *Vision Research, 49,* 1154–1165.

Belopolsky, A. V., & Theeuwes, J. (2012). Updating the premotor theory: The allocation of attention is not always accompanied by saccade preparation. *Journal of Experimental Psychology: Human Perception and Performance, 38,* 902–914.

Bichot, N. P., Cave, K. R., & Pashler, H. (1999). Visual selection mediated by location: Feature-based selection of

noncontiguous locations. *Perception & Psychophysics, 61,* 403–423.

Bichot, N. P., & Schall, J. D. (1999). Effects of similarity and history on neural mechanisms of visual selection. *Nature Neuroscience, 2,* 549–554.

Born, S., Kerzel, D., & Theeuwes, J. (2011). Evidence for a dissociation between the control of oculomotor capture and disengagement. *Experimental Brain Research, 208,* 621–631.

Castiello, U., & Umiltà, C. (1990). Size of the attentional focus and efficiency of processing. *Acta Psychologica, 73,* 195–209.

Castiello, U., & Umiltà, C. (1992). Splitting focal attention. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 837–848.

Chun, M. M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Sciences, 4,* 170–178.

Chun, M. M. (2011). Visual working memory as visual attention sustained internally over time. *Neuropsychologia, 49,* 1407–1409.

Cohen, M. A., Alvarez, G. A., & Nakayama, K. (2011). Natural-scene perception requires attention. *Psychological Science, 22,* 1165–1172.

Collins, D. L., Neelin, P., Peters, T. M., & Evans, A. C. (1994). Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *Journal of Computer Assisted Tomography, 18,* 192–205.

Corbetta, M., Kincade, J. M., & Shulman, G. L. (2002). Neural systems for visual orienting and their relationships to spatial working memory. *Journal of Cognitive Neuroscience, 14,* 508–523.

Corbetta, M., Miezin, F. M., Shulman, G. L., & Petersen, S. E. (1993). A PET study of visuospatial attention. *Journal of Neuroscience, 13,* 1202–1226.

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

Curtis, C. E., & D'Esposito, M. (2003). Success and failure suppressing reflexive behavior. *Journal of Cognitive Neuroscience, 15,* 409–418.

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

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

Eriksen, C. W., & St. James, J. D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception & Psychophysics, 40,* 225–240.

Evans, K. K., & Treisman, A. (2005). Perception of objects in natural scenes: Is it really attention free? *Journal of Experimental Psychology: Human Perception and Performance, 31,* 1476–1492.

Fagioli, S., & Macaluso, E. (2009). Attending to multiple visual streams: Interactions between location-based and category-based attentional selection. *Journal of Cognitive Neuroscience, 21,* 1628–1641.

Fink, G. R., Marshall, J. C., Halligan, P. W., Frith, C. D., Frackowiak, R. S., & Dolan, R. J. (1997). Hemispheric specialization for global and local processing: The effect of stimulus category. *Proceedings of the Royal Society of London, Series B: Biological Sciences, 264,* 487–494.

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.

Folk, C. L., Remington, R. W., & Wright, J. H. (1994). The structure of attentional control: Contingent attentional capture by apparent motion, abrupt onset, and color.

Journal of Experimental Psychology: Human Perception and Performance, 20, 317–329.

Friston, K. J., Penny, W., Phillips, C., Kiebel, S., Hinton, G., & Ashburner, J. (2002). Classical and Bayesian inference in neuroimaging: Theory. *Neuroimage, 16,* 465–483.

Geng, J. J., & Mangun, G. R. (2009). Anterior intraparietal sulcus is sensitive to bottom–up attention driven by stimulus salience. *Journal of Cognitive Neuroscience, 21,* 1584–1601.

Hollingworth, A., & Henderson, J. M. (1998). Does consistent scene context facilitate object perception? *Journal of Experimental Psychology: General, 127,* 398–415.

Indovina, I., & Macaluso, E. (2007). Dissociation of stimulus relevance and saliency factors during shifts of visuospatial attention. *Cerebral Cortex, 17,* 1701–1711.

Jans, B., Peters, J. C., & De Weerd, P. (2010). Visual spatial attention to multiple locations at once: The jury is still out. *Psychological Review, 117,* 637–684.

Kimmig, H., Greenlee, M. W., Gondan, M., Schira, M., Kassubek, J., & Mergner, T. (2001). Relationship between saccadic eye movements and cortical activity as measured by fMRI: Quantitative and qualitative aspects. *Experimental Brain Research, 141,* 184–194.

Lavie, N., Beck, D. M., & Konstantinou, N. (2014). Blinded by the load: Attention, awareness and the role of perceptual load. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences, 369,* 20130205.

Li, F. F., VanRullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences, U.S.A., 99,* 9596–9601.

Ludwig, C. J., & Gilchrist, I. D. (2003). Goal-driven modulation of oculomotor capture. *Perception & Psychophysics, 65,* 1243–1251.

Macaluso, E., & Doricchi, F. (2013). Attention and predictions: Control of spatial attention beyond the endogenous-exogenous dichotomy. *Frontiers in Human Neuroscience, 7,* 685.

Marszałek, M., & Schmid, C. (2007). *Accurate object localization with shape masks.* Paper presented at the IEEE Conference on Computer Vision & Pattern Recognition.

McMains, S. A., & Somers, D. C. (2004). Multiple spotlights of attentional selection in human visual cortex. *Neuron, 42,* 677–686.

McMains, S. A., & Somers, D. C. (2005). Processing efficiency of divided spatial attention mechanisms in human visual cortex. *Journal of Neuroscience, 25,* 9444–9448.

McPeek, R. M. (2006). Incomplete suppression of distractor-related activity in the frontal eye field results in curved saccades. *Journal of Neurophysiology, 96,* 2699–2711.

Mesulam, M. M. (1981). A cortical network for directed attention and unilateral neglect. *Annals of Neurology, 10,* 309–325.

Mulckhuyse, M., Van der Stigchel, S., & Theeuwes, J. (2009). Early and late modulation of saccade deviations by target distractor similarity. *Journal of Neurophysiology, 102,* 1451–1458.

Nebel, K., Wiese, H., Stude, P., de Greiff, A., Diener, H. C., & Keidel, M. (2005). On the neural basis of focused and divided attention. *Brain Research, Cognitive Brain Research, 25,* 760–776.

Nobre, A. C., Sebestyen, G. N., Gitelman, D. R., Mesulam, M. M., Frackowiak, R. S., & Frith, C. D. (1997). Functional localization of the system for visuospatial attention using positron emission tomography. *Brain, 120,* 515–533.

Opelt, A., Pinz, A., Fussenegger, M., & Auer, P. (2006). *Generic object recognition with boosting.* Paper presented at the IEEE Transactions on Pattern Recognition and Machine Intelligence (PAMI).

Painter, D. R., Dux, P. E., & Mattingley, J. B. (2015). Distinct roles of the intraparietal sulcus and temporoparietal junction in attentional capture from distractor features: An individual differences approach. *Neuropsychologia, 74,* 50–62.

Painter, D. R., Dux, P. E., Travis, S. L., & Mattingley, J. B. (2014). Neural responses to target features outside a search array are enhanced during conjunction but not unique-feature search. *Journal of Neuroscience, 34,* 3390–3401.

Peelen, M. V., & Kastner, S. (2011). A neural basis for real-world visual search in human occipitotemporal cortex. *Proceedings of the National Academy of Sciences, U.S.A., 108,* 12125–12130.

Peelen, M. V., & Kastner, S. (2014). Attention in the real world: Toward understanding its neural basis. *Trends in Cognitive Sciences, 18,* 242–250.

Penny, W., & Holmes, A. (2003). Random effects analysis. In R. S. J. Frackowiak (Ed.), *Human brain function II*. New York: Elsevier.

Reeder, R. R., & Peelen, M. V. (2013). The contents of the search template for category-level search in natural scenes. *Journal of Vision, 13,* 13.

Reeder, R. R., van Zoest, W., & Peelen, M. V. (2015). Involuntary attentional capture by task-irrelevant objects that match the search template for category detection in natural scenes. *Attention, Perception, & Psychophysics, 77,* 1070–1080.

Rousselet, G. A., Fabre-Thorpe, M., & Thorpe, S. J. (2002). Parallel processing in high-level categorization of natural images. *Nature Neuroscience, 5,* 629–630.

Sachs, O., Weis, S., Krings, T., Huber, W., & Kircher, T. (2008). Categorical and thematic knowledge representation in the brain: Neural correlates of taxonomic and thematic conceptual relations. *Neuropsychologia, 46,* 409–418.

Seger, C. A., Poldrack, R. A., Prabhakaran, V., Zhao, M., Glover, G. H., & Gabrieli, J. D. (2000). Hemispheric asymmetries and individual differences in visual concept learning as measured by functional MRI. *Neuropsychologia, 38,* 1316–1324.

Seidl, K. N., Peelen, M. V., & Kastner, S. (2012). Neural evidence for distracter suppression during visual search in real-world scenes. *Journal of Neuroscience, 32,* 11812–11819.

Silk, T. J., Bellgrove, M. A., Wrafter, P., Mattingley, J. B., & Cunnington, R. (2010). Spatial working memory and spatial attention rely on common neural processes in the intraparietal sulcus. *Neuroimage, 53,* 718–724.

Theeuwes, J., Belopolsky, A., & Olivers, C. N. (2009). Interactions between working memory, attention and eye movements. *Acta Psychologica, 132,* 106–114.

Theeuwes, J., Kramer, A. F., Hahn, S., Irwin, D. E., & Zelinsky, G. J. (1999). Influence of attentional capture on oculomotor control. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 1595–1608.

Thompson, K. G., Bichot, N. P., & Sato, T. R. (2005). Frontal eye field activity before visual search errors reveals the integration of bottom–up and top–down salience. *Journal of Neurophysiology, 93,* 337–351.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature, 381,* 520–522.

Tong, F. (2004). Splitting the spotlight of visual attention. *Neuron, 42,* 524–526.

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

Van der Stigchel, S., van Koningsbruggen, M., Nijboer, T. C., List, A., & Rafal, R. D. (2012). The role of the frontal eye fields in the oculomotor inhibition of reflexive saccades: Evidence from lesion patients. *Neuropsychologia, 50,* 198–203.

VanRullen, R., & Thorpe, S. J. (2001a). Is it a bird? Is it a plane? Ultra-rapid visual categorisation of natural and artifactual objects. *Perception, 30,* 655–668.

VanRullen, R., & Thorpe, S. J. (2001b). The time course of visual processing: From early perception to decision-making. *Journal of Cognitive Neuroscience, 13,* 454–461.

Walker, S., Stafford, P., & Davis, G. (2008). Ultra-rapid categorization requires visual attention: Scenes with multiple foreground objects. *Journal of Vision, 8,* 21.1–21.12.

Wolfe, J. M. (1998). Visual search. In H. Pashler (Ed.), *Attention* (pp. 13–74). Hove: Psychology Press.

Wolfe, J. M., Alvarez, G. A., Rosenholtz, R., Kuzmova, Y. I., & Sherman, A. M. (2011). Visual search for arbitrary objects in real scenes. *Attention, Perception, & Psychophysics, 73,* 1650–1671.

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience, 5,* 495–501.

Wu, C. C., Wick, F. A., & Pomplun, M. (2014). Guidance of visual attention by semantic information in real-world scenes. *Frontiers in Psychology, 5,* 54.

Xiao, J., Hays, J., Ehinger, K., Oliva, A., & Torralba, A. (2010). *SUN database: Large-scale scene recognition from abbey to zoo*. Paper presented at the IEEE Conference on Computer Vision and Pattern Recognition.