Orienting of Visuo-Spatial Attention in Complex 3D Space: Search and Detection

Akitoshi Ogawa* and Emiliano Macaluso

Neuroimaging Laboratory, Santa Lucia Foundation, Via Ardeatina 306, Rome, Italy

◆————————————————————————◆

Abstract: The ability to detect changes in the environment is necessary for appropriate interactions with the external world. Changes in the background go more unnoticed than foreground changes, possibly because attention prioritizes processing of foreground/near stimuli. Here, we investigated the detectability of foreground and background changes within natural scenes and the influence of stereoscopic depth cues on this. Using a flicker paradigm, we alternated a pair of images that were exactly same or differed for one single element (i.e., a color change of one object in the scene). The participants were asked to find the change that occurred either in a foreground or background object, while viewing the stimuli either with binocular and monocular cues (bmC) or monocular cues only (mC). The behavioral results showed faster and more accurate detections for foreground changes and overall better performance in bmC than mC conditions. The imaging results highlighted the involvement of fronto-parietal attention controlling networks during active search and target detection. These attention networks did not show any differential effect as function of the presence/absence of the binocular cues, or the detection of foreground/background changes. By contrast, the lateral occipital cortex showed greater activation for detections in foreground compared to background, while area V3A showed a main effect of bmC vs. mC, specifically during search. These findings indicate that visual search with binocular cues does not impose any specific requirement on attention-controlling fronto-parietal networks, while the enhanced detection of front/near objects in the bmC condition reflects bottom-up sensory processes in visual cortex. *Hum Brain Mapp 36:2231–2247, 2015.* © 2015 The Authors Human Brain Mapping Published by Wiley Periodicals, Inc.

Key words: stereoscopic viewing; fMRI; change detection; lateral occipital complex; V3A

◆————————————————————————◆

INTRODUCTION

In everyday life, we can detect and recognize objects embedded within complex scenes that contain many elements competing for processing resources. The ability to detect changes within such complex environments is important to guide behavior in natural dynamic environments. Many factors contribute to the selection and detection processes, including the location of the change in depth. Changes occurring near to the organism should be registered and processed with a higher level of priority, compared with events occurring at far distances [Ozkan and Braunstein, 2010, see also Pomplun et al., 2013]. Experimentally, the ability to detect changes has been

investigated using the flicker paradigm [Rensink, 2002; Rensink et al., 1997]. In this paradigm, two pictures are presented in alternation with a blank image briefly presented between them. The pictures are identical, but for one change that is the target of the detection task. The visual disruption associated with the presentation of the intervening blank image reduces the ability to detect the difference between the two pictures (i.e., change blindness). However, when attention is focused on the target—e.g., by informing the participant about the target location—the change can be easily detected [Cavanaugh et al., 2004; Rensink, 2002 see also Lathrop et al., 2011; Seegmiller et al., 2011].

Here we used the flickering paradigm to investigate behavioral and neuro-physiological correlates of the allocation of visuo-spatial attention in complex three-dimensional (3D) environments. Previous studies using simple visual stimuli showed that changes in the foreground are more detectable than those in the background, even if the changes in the background are larger than those in the foreground [Mazza et al., 2005; Turatto et al., 2002]. This is consistent with the notion that foreground (nearer) locations receive higher processing priority, possibly via enhanced allocation of spatial attention there compared to background objects [Ozkan and Braunstein, 2010, see also Pomplun et al. 2013].

In these previous studies, the foreground/background position of the target was rendered using monocular cues only, i.e., texture and perspective. This leads to the question of whether binocular depth signals would further modulate change detection at front/back locations. A few studies that investigated the role of binocular cues in visual search revealed that these cue can make visual search in 3D space more efficient [Nakayama and Silverman, 1986; see also Finlayson et al., 2013]. Binocular signals can be used to segregate objects in 3D visual space [Chau and Yeh, 1995], making the allocation of attention more efficient and facilitating the detection of the search targets. However, the question remains of whether such enhanced foreground–background segregation also results in facilitated change detection. Thus, here we crossed the factors of target-position (foreground/background) and viewing-condition (bmC/mC: binocular and monocular depth cues vs. monocular cues only) asking whether these would interact with each other or would jointly contribute to task performance, but via separate mechanisms.

From the neuroimaging perspective, visual search has been studied extensively, but typically using visual displays with simple items (e.g., bars, letters, and faces) and without any depth cues. Overall these studies associated serial/inefficient visual search with the activation of the dorsal fronto-parietal attention network that includes the superior parietal lobule, the intraparietal sulcus (IPS) and the frontal eye fields (FEF) [Corbetta and Shulman, 1998; Fairhall and Macaluso, 2009; Shulman et al., 2003, 2007]. By contrast, the detection of behaviorally relevant targets has been associated with the activation of a ventral fronto-parietal system that includes the inferior parietal lobule (angular gyrus [AG], supramarginal gyrus [SMG]) and the

inferior frontal gyrus [Linden et al., 1999; Shulman et al., 2007, 2009; see also Corbetta et al., 2008]. These areas are also involved in attention reorienting in depth, when binocular cues are available (close-to-far or far-to-close, in 3D space [Chen et al., 2012]). Accordingly, it can be hypothesized that change detection in complex 3D space may further tax on these fronto-parietal networks. This would demonstrate that these networks can make use of disparity signals for the control of visuo-spatial attention.

Alternatively, disparity signals may contribute to the processing of complex scenes primarily via modulation of activity in occipital visual cortex. Change detection tasks, without any disparity cue, have consistently highlighted the role of the visual areas that represent the target-defining feature. Beck et al. [2001] reported activation of the fusiform gyrus for face-changes versus a more medial and anterior region for place-changes. The lateral occipital complex (LOC) has been associated with the detection of object repetitions [Schwarzkopf et al., 2010], while change detection of landmark buildings was associated with the activity of single neurons in human parahippocampal gyrus [Reddy et al., 2006]. If depth cues affect processing of foreground/background objects via modulation of sensory representations, it can be hypothesized that performing a change detection task in 3D space will involve occipital regions processing the binocular cues (e.g., V3A [Brouwer et al. 2005; Neri et al., 2004; Ogawa and Macaluso, 2013; Tsao et al., 2003]) and/or regions involved in object representation (e.g., LOC [Grill-Spector, 2003; Grill-Spector et al., 1999; Malach et al., 1995]). Moreover, binocular cues may influence brain activity during search, i.e., when these cues can help selecting objects for attention focusing (e.g., via object-background segmentation [Cottereau et al., 2011]), and/or brain activity associated with change-detection, when the participant becomes aware of the target position in 3D space.

Accordingly, in this functional magnetic resonance imaging (fMRI) study we asked participants to detect color changes of objects placed within natural scenes that were presented either with binocular and monocular cues (bmC) or with monocular cues only (mC). The target object was selected either in the foreground or in the background. Behaviorally, we expected faster and more accurate performance for "front" than "back" targets [cf. Mazza et al., 2005; Ozkan and Braunstein, 2010; Turatto et al., 2002], with a further advantage when the display also included the binocular cues [cf. Finlayson et al., 2013; Nakayama and Silverman, 1986]. The fMRI analyses separated search-related activity and detection-related activity. We expected that search, and any contribution of binocular cues to this, would primarily affect activity in the dorsal fronto-parietal attention network; while detection-related effects should engage the ventral fronto-parietal network [Shulman et al., 2007]. Detection-related activation in the ventral system may be further modulated according to the location in depth of the target object [Chen et al., 2012]. In addition, we hypothesized that target-position and viewing-

condition would influence the "figure on ground" segregation of objects presented at different depths, with the presence of binocular cues further facilitating the processing of foreground objects. The latter may entail additive effects of "front" target-position (e.g., segmentation processes in LOC [Finlayson et al., 2013]) and "bmC" viewing-condition (e.g., processing of binocular signals in V3A [Cottereau et al., 2011; Tsao et al., 2003]), or the two factors interacting in the same brain region (e.g., LOC and/or V3A).

MATERIALS AND METHODS

Subjects

Twenty subjects (aged 20–31, mean = 23.6 years, 11 females and 9 males) with no history of neurological or psychiatric illness participated in the experiment. All had normal or corrected-to-normal visual acuity with contact lenses. They reported no difficulty to perceive stereoscopic depth when viewing 3D pictures. All participants gave written informed consents prior to the experiment. The study was approved by ethical committee of Santa Lucia Foundation.

Experimental Paradigm

We used a change detection task [Beck et al., 2001; Huettel et al., 2001] with pictures of natural scenes. Each trial consisted in the repeated presentation of two pictures, with an interleaved blank image (Fig. 1A). The two pictures could be either identical (i.e., target-absent) or they could differ for one single item (i.e., target-present). The target was created by changing the color of one object, either in the foreground or in the background of the scene ("front vs. back" target-condition: F vs. B). In different blocks the pictures were presented with or without stereoscopic depth cues. Given that even in the absence of any binocular cue, the monocular cues can contribute to the perception of object position in depth (e.g., occlusions, perspective, and relative size; see also discussion section), we labeled the two viewing-conditions: "bmC" vs. "mC". The task of the participants was to report the presence of the target, by pressing a button as soon as they detected the change. The analyses of the imaging data assessed the effect of target-position (F/B), viewing-condition (bmC/mC), and any interaction between these, separately on brain activity associated with the "search" of the target and with the "detection" of the target. The experiment included a localizer fMRI scan, aiming to individually identify the disparity-selective V3A and the object-selective LOC in the occipital visual cortex.

Trial Structure and Conditions

Each trial consisted in the repeated presentation (six times) of two images, with an interleaved blank screen. On target-present trials (F/B conditions), the two images were different, while in the "same" condition (S) only the original picture was used. Each image presentation lasted for 400 ms, while the blank screen was shown for 100 ms (see also Fig. 1A). Accordingly, the overall duration of a trial was 6 s and the sequence of alternating images and blank screens continued even if the participant detected the target and pressed the response button (i.e., the stimulus duration was fixed, irrespective of individual performance). The intertrial interval was variable between 2 and 4 s (mean = 3 s).

The experiment included 192 trials, with the six main conditions repeated 32 times each (bmC-F, bmC-B, bmC-S; and mC-F, mC-B, and mC-S, with "S" standing for "same", i.e., target-absent trial). The stimulus set included 96 pictures (see also below), with each picture presented to the subject twice: once with binocular cues (bmC: either with F, B or S target-condition) and once without the binocular cues (mC: with a different target-condition than during the bmC presentation). The mC/bmC viewing-conditions alternated every six trials in a blocked-design manner. Within each of these six-trials mini-blocks, F/B/S conditions were randomized. The use of each picture for the F/B/S conditions was counterbalanced across subjects. The 192 trials were presented in four separate fMRI runs, including 48 trials each.

Stimulus Material

We collected 96 stereoscopic images by taking pictures with a 3D camera (MHS-FS3, Sony corp.) and by downloading images from specialized websites (www.3d-foto-shop.de and www.panoramio.com). The picture set comprised 6 indoor and 90 outdoor naturalistic scenes, including people, animals, objects (e.g., statues, ships, fruits, vegetables, flowers, and vehicles), buildings, etc. All pictures were resampled to a resolution of 960 pixels width × 540 pixels height. The target for the change detection task was created by manipulating the hue of the selected object, between 20° and 180° using Photoshop CS (Adobe Systems). In each picture we identified two objects to be used as targets: one in the foreground and one in the background. There was at least one object behind/further away from the "foreground" target objects, and at least one object nearer/in front of the selected "background" objects. Thus, note that here the definition of "foreground" and "background" targets was not related to the zero disparity plane but instead reflected to the relative position of an object with respect to the other objects in the scene (please see also below).

For each picture we computed a depth-map and further verified the relative position in depth of the two selected objects. The depth-maps were calculated using the HL-SIFT flow between left and right images [Liu and Yuen, 2011; Lowe, 1999; see also, Ogawa et al., 2013]. The raw disparity values of the foreground targets ranged between −22.1 and 30.2 pixels (mean = 2.97, SEM = 2.15)
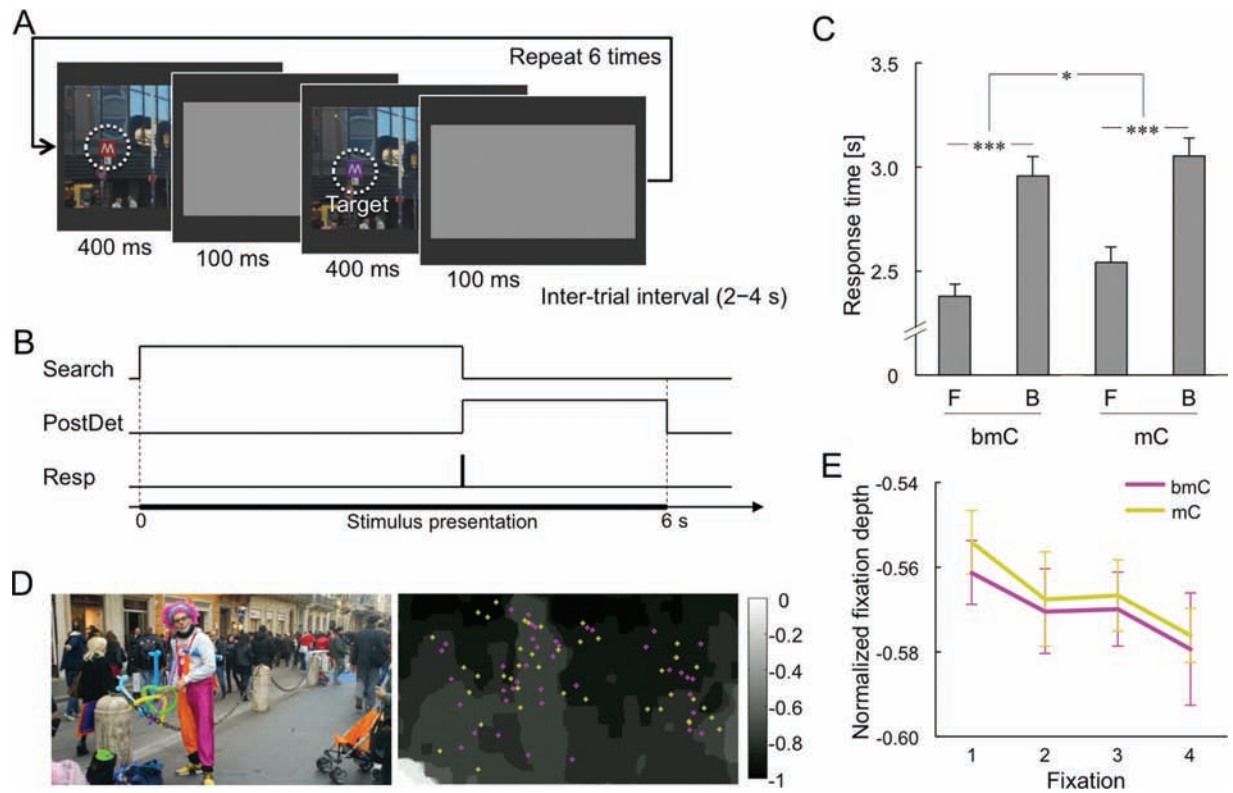
Figure 1.

Change-detection task and behavioral results. **A**: Time course of a trail. Two pictures were presented sequentially (400 ms), with an interleaved blank image (100 ms). The cycle was repeated 6 times, for total trial duration of 6 s. The two pictures could differ because the color of one object was changed (see "target" in the example). In different conditions, the target object was in the foreground (F), in the background (B), or there was no change ("same" condition). The participants were asked to press a button as soon as possible when they found the change. **B**: Modeling of the blood-oxygen-level dependent response. Each trial was modeled using a combination of predictors (see also Table I). The Hit trials (target-present and correct detection by the participant) were dived in a "search" phase (Search), a "response" event (Resp, i.e., target detection in analysis) and a "post-detection" phase (postDet). The Miss and CR trials included only the "search" phase. **C**:The response times on Hit trials. Reponses in the bmC conditions were significantly faster than in the mC conditions, and changes in the foreground were detected significantly faster than changes in the background. Error bars are standard error. $*P < 0.05$, $***P < 0.001$. **D**: Example of the computation of the fixation depth-values for one image. The depth-map (lower panel) was computed using the algorithm of HL-SIFT flow (Lowe, 1999) and normalized to values between 0 (nearest) and -1 (farthest). The violet dots (bmC-viewing) and the yellow dots (mC-viewing) show the fixation-positions across all subjects. **E**: The time course of fixation depth-values. Fixation depth gradually moved from close to far both in bmC and mC conditions, indicating that the participants searched foreground/near locations first. Error bars are standard error.

corresponding to -12 to 45.4 arcmin (mean = 3.56, SEM = 2.58), while the range of disparities for the background targets was between -10.0 and 37.9 pixels (mean = 10.5, SEM = 2.27), that is, from -26.5 to 36.2 arcmin (mean = 12.6, SEM = 2.72). However, it should be noted that the distance between the lenses of 3D camera was unknown for most of the pictures. Therefore, we normalized the disparity values before comparing across pictures/conditions. The normalized depth-maps included values ranging from 0 (nearest) to -1 (farthest; see also Fig. 1D). These normalized depth-values represent the relative depth of an object within the range of depths in each picture, but not the absolute depth-position of the object or its distance from the zero-disparity plane. We formally confirmed that the "foreground" targets had a smaller relative disparity that the "background" targets using as Wilcoxon signed-rank test ($z = 3.22$, $P < 0.01$). Indeed, except for two images where a partial occlusion of the background target-object made the results of this computational approach unreliable, the normalized depth-value of

the foreground target was always larger than that of the background target in the same image.

Moreover, we carried out additional tests to check that there was no systematic difference in "target size" and "target eccentricity" between the objects selected as front and back targets. For each picture, we created two target-images by computing the RGB difference between the original picture and the two modified pictures that included the targets. The target images were transformed to grayscale with an R:G:B ratio = 0.2989:0.5879:0.1140, and the value of each pixel was normalized in the 0 to 1 range, i.e., from black to white. The images were then binarized (threshold = 0.05) resulting in the final target-image, with pixel value = 1 at the position of the color change (i.e., target-position) and pixel values = 0 everywhere else. With these target-images, we estimated the target-size as the number of pixels with value = 1; and the target-eccentricity considering the point of gravity of the target area (x_0, y_0). The point of gravity was calculated as follows.

$$x_0 = \frac{\sum_{i=1}^{540}\sum_{j=1}^{960} jB(i,j)}{A}, y_0 = \frac{\sum_{i=1}^{540}\sum_{j=1}^{960} iB(i,j)}{A} \qquad (1)$$

where $A = \sum_{i=1}^{540}\sum_{j=1}^{960} B(i,j)$ and $B(i, j) = 1$ if (i, j) was in the target area, otherwise $B(i, j) = 0$.

We compared the target sizes and the target eccentricities between front (F) and back (B) conditions using paired t tests. No significant difference was found neither for target-size ($t_{95} = 0.107$, $p = 0.91$) nor for target-eccentricity ($t_{95} = 1.18$, $p = 0.24$). Accordingly, any effect of target-position (e.g., faster detection and/or greater brain activity in F vs. B or vice versa) should not reflect any trivial confound related to the choice of the pictures or of the target objects. Please note that despite these attempts to equate the targets' characteristics across conditions, there are a number of other untested aspects that may have differed. For example, previous behavioral work showed that objects in the background are perceived as being larger than foreground objects [Arnold, et al., 2008]. This may predict that changes in background would be more detectable, which is the opposite of our current results (see Fig. 1C; and Results section, below). Nonetheless, we acknowledge that a limitation of using complex and naturalistic stimuli is that these do not allow controlling for all possible low-level aspects of the sensory input.

Stimulus Presentation and Eye-Tracking

The presentation of the stimuli and the recording of the button-presses were controlled using "psychophysics toolbox" [Brainard, 1997] running on MATLAB 7.1 (Math-Works). The visual stimuli were presented using a LCD projector (NP216G, NEC Corp.) operating at 120 Hz, which was synchronized with a linear polarizer (DepthQ®, Light-

speed Design). The stimuli were projected on a semi-opaque screen positioned inside the magnet and that was viewed via a mirror system. The participants wore a MR-compatible passive 3D eyewear allowing them to view the polarized images just with the left or the right eye. This generated the stereoscopic stimuli, when different images were projected to the two eyes (i.e., in the bmC condition). Presenting the same image to both eyes generated in the mC condition. The pictures were viewed at a visual angle of 10.8° height and 19.2° width.

During the fMRI experiment the eye-position was recorded using an MR-compatible eye-tracker operating at 60 Hz (ASL 5000, Applied Science Laboratories). Before scanning, calibration was performed using a grid of nine fixation points. Data were processed using in-house software (www.slneuroimaginglab.com/mt-tools). We used a saccade velocity threshold ($> 60°/s$) and a fixation duration threshold (>100 ms) to classify saccades and fixations in each trial. The eye-tracking data of two subjects were excluded because of low quality due to reflection artifacts (contact lens), leaving 18 subjects for the eye-position analyses.

Detection Performance and Eye-Movement Analyses

There were four types of possible behavioral responses: correct detection in target-present trials (Hit), incorrect responses in target-absent trials (False alarm), no response in target-present trials (Miss), and no response in target-absent trials (correct rejection [CR]). On target-absent trials the performance was at ceiling (False alarms = 3.4%; CR = 96.6%), thus the analyses of the behavioral responses (accuracy and reaction times) focused on the target-present trials. With these, we performed two-way repeated-measure analysis of variances (ANOVAs) assessing the effect of "mC/bmC viewing-condition" and "F/B target-position" on hit rates (performance = Hits / (Hits + Misses)) and detection times (Hit trials only).

The analyses of the eye-movement data aimed to assess whether there was any systematic relationship between the patterns of eye movements and the presence of binocular depth signals. We considered only Miss and CR trails, i.e., when the participants searched the display for the entire duration of the trial. Previous studies highlighted the tendency of participants to make saccades to the nearer objects [Ozkan and Braunstein, 2010; Pomplun et al., 2013]. For example, Pomplun et al. [2013] used a visual search task with slanted and colored geometric objects under the stereoscopic viewing. They investigated the search strategy when the targets and distractors were presented across different depths, and found that participants trended to perform the initial saccades towards objects in the closest plane. Accordingly, here we hypothesized that participants would prioritize the exploration of near/front positions compared with far/back positions, and we asked

TABLE I. Regressors used in Model A, B, and C (first-level analyses)

Model A					Target-present				Target-absent	
	Search		PostDet			Detection			CR	
	bmC[a]	mC[a]	bmC[a]	mC[a]	bmC-F	bmC-B	mC-F	mC-B	bmC	mC

Model B					Target-present					Target-absent	
	Search				PostDet				Detection		CR
	bmC-F	bmC-B	mC-F	mC-B	bmC-F	bmC-B	mC-F	mC-B	bmC*	mC*	bmC	mC

Model C					Target-present				Target-absent	
	bmC				mC				CR	
	F-Hit	F-Miss	B-Hit	B-Miss	F-Hit	F-Miss	B-Hit	B-Miss	bmC	mC

[a]F and B conditions were collapsed.

whether the mC/bmC viewing-conditions influenced this pattern of exploration. For this, we considered the depth-map associated with each picture (cf. above) and determined the depth-values at subsequent fixations (cf. Parkhurst et al., 2002 for a related approach, but using saliency-maps rather than depth-maps; see also Fig. 1E). For each trial and each subject, we examined the first four fixations and averaged the depth-values within a 21-pixel square, separately for each fixation. For each fixation in the sequence, the depth-values were averaged across images separately in the mC and bmC conditions. The choice of four fixations' sequences ensured that these analyses included at least the data of 10 trials for each subject at each fixation in the sequence. The normalized depth-value data were assessed using a two-way repeated-measures ANOVA, with the factors of: "fixation" (1–4) and viewing condition (mC/bmC). Please note that these tests made used of normalized depth-values (cf. above) and thus assessed relative changes "within-picture," rather than evaluating shifts of gaze between specific/absolute depth-positions.

Image Acquisition

A 3T scanner Allegra (Siemens Medical Systems, Erlangen, Germany) equipped for echo-planar imaging (EPI) was used to acquire functional magnetic resonance images. A head-sized quadrature volume coil was used for radio frequency transmission and reception. Mild cushioning minimized head movement. Thirty-two slices of functional images were acquired using blood oxygenation level dependent imaging (192 mm × 192 mm × 120 mm, in-plane resolution = 64 × 64, pixel size = 3 mm × 3 mm, thickness = 2.5 mm, 50% distance factor, TR = 2.08 s, TE = 30 ms), covering the entire cortex. We acquired 296 scans in the functional localizer (see below) and 216 scans in each of four runs of the main experiment. The first four scans in each fMRI run were discarded to ensure magnetization equilibrium.

Image Preprocessing

We used SPM8 (Wellcome Department of Cognitive Neurology, University College London) on MATLAB to process the acquired images. In preprocessing, we performed slice-timing correction using the middle slice as a reference, scan-to-scan realignment, normalization to the EPI template of SPM8, resampling the images with the voxel size of 3 mm × 3 mm × 3 mm, and spatial smoothing (FWHM = 8 mm). A high-pass filter of 128 s was used to remove low-frequency noise.

Modeling of the fMRI Responses

The aim of the current study was to investigate the effect of target-position (F/B) and viewing-condition (mC/bmC), plus any interaction between the two, specifically on activity associated with search and with change detection. For this, our main analyses included multiple trial phases with variable durations (see Fig. 1B; and Models A and B in Table I); but we also carried out an additional control analysis using fixed 6-s durations and sorting each trial into a specific experimental condition (Model C in Table I, and see below).

For the analyses including multiple trial phases, the individual blood-oxygen-level dependent (BOLD) time series were modeled with two separate general linear models (GLM), because including of all the effects of interest (F/B × bmC/mC) on all phases of the trial lead to high correlations between predictors. Accordingly, one analysis focused on the detection-related activity and included four separate regressors for the detection phase (F/B × mC/bmC), but only two regressors for the search-related activity (mC and bmC conditions), see Model A in Table I. The second main analysis focused on search-related activity, thus now including four separate regressors for the search phase (F/B × mC/bmC) and two regressors for the detection-phase, see Model B.

Model A included a total of 10 predictors (cf. Table I). Four predictors modeled the correct detections in target-present trials (i.e., Hits), using single events time-locked to

participant's responses. The four predictors accounted separately for detection-related activation of front and back targets, in mC and bmC viewing-conditions. The remaining six predictors modeled the other phases of the trial (Fig. 1B). For Hit trials, we considered a "search" phase that started at the stimulus onset and ended at the time of target detection (i.e., the button-press), and a "post-detection" phase that started at the button-press and lasted until the end of the stimulus presentation. In both the cases, the predictor consisted of mini-blocks with a variable, trial-specific duration that depended on the detection-time of the participant. The "search" and the "post-detection" ("postDet") phases were modeled with two regressors each, accounting for the effect of mC/bmC viewing but not of F/B target-position. The model included two additional predictors modeling the target-absent trials (duration = 6 s) separately for mC and bmC viewing-conditions.

The Model B included 12 predictors and was conceptually similar to Model A but now focusing on search-related effects. The detection phase included only two event-related predictors corresponding to target-detections in mC and bmC conditions but pooling F and B targets. The search phase in target-present trials was modeled using four separate predictors corresponding to trials including F and B targets, presented in mC and bmC conditions (miniblocks with a duration corresponding to the actual, trial-specific duration of the search). The "postDet" phase of these trials also included four predictors given by the crossing of the F/B × mC/bmC conditions (duration = 6 s minus the trial-specific duration of the search phase, see Fig. 1B). Again, the target-absent trials were modeled separately for bmC and mC viewing-conditions (miniblocks with duration = 6 s).

Because these multiple phases models included different trial-durations across conditions (e.g., shorter search for front- vs. back-targets, see Fig. 1C), and did not model the target-position by viewing-condition interaction for the phase of no interest (e.g., the search phase in Model A, which was used to test detection-related effects), we carried out an additional control analysis seeking to replicate our main results but now using a single-phase fixed duration model (Model C, Table I). The single-subject GLMs included 10 conditions corresponding to F-Hit, F-Miss, B-Hit, B-Miss and CR; separately for the two viewing-conditions (bmC/mC). Each experimental trial was assigned to one of these conditions, and the corresponding events were modeled with a fixed duration of 6 s (i.e., the duration of the visual presentation), time-locked to the stimulus onset. We investigated search- and detection-related effects, and any influence of target-position and viewing-condition on these, by directly comparing different trial-types (see *group-level analyses* below).

V3A and LOC Functional Localizer

Together with the main change detection experiment, we also performed a localizer fMRI run with the aim to identify area V3A (visual disparity) and area LOC (object recognition) on a subject-by-subject basis. The localizer consisted of blocks of bmC intact-objects (bmC-i), mC intact-objects (mC-i), and mC scrambled-objects (mC-s) that alternated every 20 s. In each block, 12 pictures were presented. Each picture was presented for 1.2 s, followed by a blank screen lasting for 0.5 s. The pictures were shot using a 3D camera (MHS-FS3, Sony co.). Each picture included a single familiar object (e.g., banana, mug, stapler, etc.). The object was put on a gray paper sheet in front of a black background. The distance between the camera and the object was approximately 50 cm, with the camera down-facing at an angle of approximately 30°. Overall, we used 72 different pictures (= 24 objects × 3 conditions).

During the functional localizer scan, the task of the participant was to press a button as soon as possible, when the same picture was presented twice in a row (i.e., one-back task). Each block included 12 pictures, with three repeated pictures (25% of targets). There were 10 blocks for each of the three conditions (bmC-i, mC-i, and mC-s). The total duration of the functional localizer was approximately 10 min.

Group-Level Analyses of the Functional Data

Statistical inference at the group level was performed by computing contrast-images in each subject and assessing these with a set of one sample t tests and repeated-measures ANOVAs [Friston et al., 2002]. First, we used Model A to identify brain activity associated with search and with change detection, irrespective of F/B and mC/bmC conditions. For each subject, we computed two contrast-images: one corresponded to the difference between the "search" minus the "post-detection" ("postDet") conditions, and one averaged the four detection conditions (see target-present column in Table I, Model A). Using two separate one sample t tests, we assessed the overall activations and deactivations associated with search ("search vs. postDet", cf. Fig. 2A,B) and the overall effect of change-detection (mean of the four "detection" conditions, Fig. 2C).

Then we turned to the effects of F/B and mC/bmC, using Model B for the search phase and Model A for the detection phase (see above, and Table I). Two separate repeated-measures ANOVAs included either the four search-related contrast images from Model B (F/B × mC/bmC), or the four corresponding contrasts for the detection-related effects from Model A. Within these ANOVAs we tested for the main effects of target-position ("F vs. B" and vice versa), the main effects of viewing-condition ("bmC vs. mC" and vice versa) and the interactions between the two factors. All analyses above were performed at the whole brain level, using a statistical threshold of p-FWE = 0.05, corrected for multiple comparisons at the voxel-level (minimum cluster size = 4 voxels).
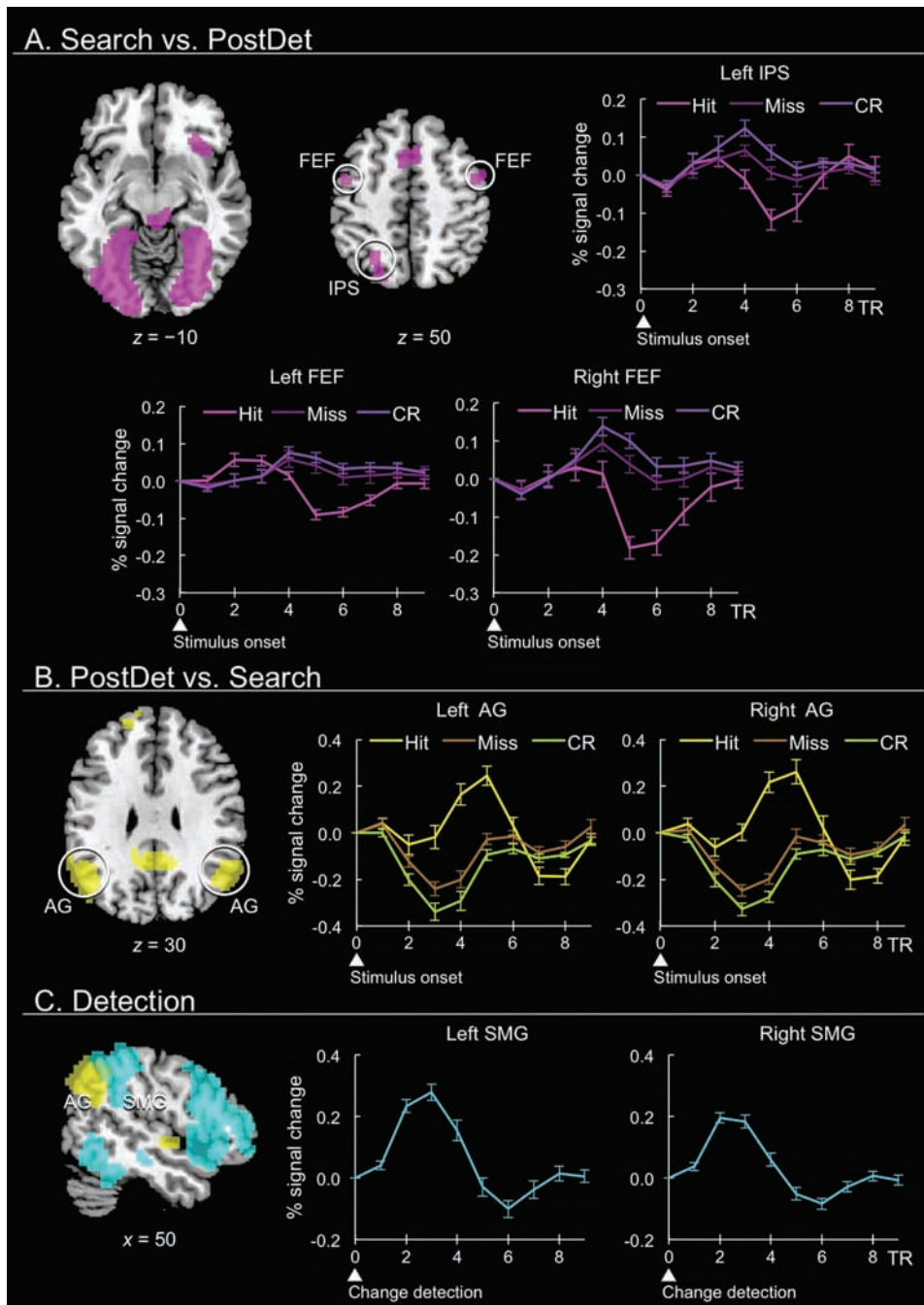
Figure 2.

Results of the whole brain analyses: search and detection. **A**: During the search phase of the trial, we found the activation in dorsal fronto-parietal regions including IPS and FEF. The signal plots show the BLOD response over the entire duration of the trial, separately for Hit, Miss, and CR trials. While in Miss and CR trials the activity remained elevated for the duration of the trial, on Hit trials the BOLD signal decreased earlier reflecting to the interruption of the search process. Note that the signal plots are presented here for illustrative purpose only. Data analyses and statistics were based on GLM models using specific predictors convolved with the hemodynamic response function (see Methods, Fig. 1B and Table I). **B**: In the AG we found the opposite pattern of activity. This area deactivated during search, with an earlier return to baseline on Hit compared with Miss and CR trials. **C**: The detection-related brain activation (rendered in cyan) that comprised the SMG plus several other regions (see Table II). The signal plots illustrate the time course of activity in the SMG on Hit trials, time-locked to the participant response (i.e., the change detection). The parietal region that deactivated during search (cf. panel B) is rendered together with the effect of change detection. This highlighted an anterior (SMG, in cyan) / posterior (AG, in yellow) segregation within the inferior parietal lobule. None of the regions reported in this figure showed a significant effect of target position (F/B) or viewing condition (mC/bmC).

Next, we performed more targeted analyses of activity in visual areas V3A and LOC, considering subject-specific regions of interest (ROIs). In each subject, we identified area V3A by comparing the "bmC-i > mC-i" conditions in the localizer; and area LOC by pooling activity in the two intact-object conditions and comparing these with the scramble condition (i.e., "bmC-i + mC-i > mC-s"). We identified peaks of individual activation within 10 mm from the corresponding group-level effect. Individual ROIs were defined as 8-mm radius spheres, and the effects of interest were extracted and averaged using MarsBaR [Brett et al., 2002]. Group-level analyses comprised t tests and ANOVAs identical to those described above for the main whole brain analyses. Significance levels were Bonferroni corrected for the number of ROIs ($n = 4$; P-corr. $= 0.0125$).

Finally, we sought to replicate the results obtained with the multiple-phases models, but now using a single-phase fixed-duration model. The contrast-images corresponding to the 10 conditions of Model C were entered in a repeated-measure ANOVA. To identify search-related activations, we compared "(Misses + CRs) > Hits": i.e., trials when the subjects searched for the entire trial duration (6 s) vs. trials when search was interrupted earlier on (i.e., when the target was detected, on average 2.7 s after stimulus-onset; cf. Fig. 1C). To identify detection-related activation, we compared "Hits > (Misses + CRs)": i.e., trials when the subjects detected the target vs. trials that did not include any detection. It should be noticed that the latter will also identify any search-related "deactivation" (please, see also discussion section). As for the main analyses with multiple trial phases, we investigated the main effects of viewing-condition (bmC/mC) and target-position (F/B), and any interaction between the two factors, examining these effects both at the whole-brain level and within the a priori defined LOC and V3A ROIs (cf. sections above).

RESULTS

Behavioral responses

The detection accuracy and the response times were assessed using two-way repeated-measures ANOVAs, with the factors of viewing-condition (mC/bmC) and target-position (F/B). Hit rates on target-present trials (i.e., Hit vs. Miss trials) were: 67.2% (s.e.m. = 2.1) in bmC-F; 51.2% (2.9) in bmC-B; 64.5% (1.8) in mC-F; and 48.7% (2.8) in mC-B condition. The corresponding ANOVA showed that the detection accuracy was higher in the "Front" than the "Back" condition ($F_{1,19} = 55.76$, $P < 0.001$) and that participants detected more targets with bmC than with mC viewing ($F_{1,19} = 4.48$, $P < 0.05$). The interaction effect was not significant. The results of the analysis of the detection times (Hit trials only) were consistent with the accuracy data: subjects were faster to detect changes in the foreground compared to changes in the background ($F_{1,19} = 35.0$, $P < 0.001$; Fig. 1C) and were faster with bmC viewing than

with two-dimensional (2D) viewing ($F_{1,19} = 5.23$, $P < 0.05$). Again the two factors did not interact. Thus, overall front-targets presented in bmC viewing were detected faster and more accurately than targets in any of the other three conditions, but the factors of target-position and viewing-condition contributed to this enhanced detection performance in an independent manner (additive effects).

Next, we asked whether participants fixated near/front positions first, and whether the presence of binocular cues modulated the fixation patterns in eye movement. Figure 1E shows the average depth-values (see Methods section) plotted for subsequent fixations and separately for the bmC and mC conditions. The eye-tracking results showed that the participants did indeed look at near-objects first, and then moved progressively towards objects located further away in space ($F_{3,51} = 3.38$, $P < 0.05$). However, this search behavior was not modulated by bmC/mC viewing-condition ($F_{1,17} = 0.23$, $P = 0.64$).

Whole Brain Analyses

First, we investigated brain activity associated with search by comparing the "search" phase versus the "postDet" phase of the target-present trials, averaging F/B target-position and mC/bmC viewing-conditions (see Fig. 1B and Methods section). This showed activation of the IPS and the superior frontal gyrus (SFG), plus the visual occipital cortex, the superior colliculus, the anterior insular cortex, and the medial frontal cortex (Fig. 2A; see also Table II). The activation of SFG putatively included the FEF. For display purposes, we extracted the time course of the BOLD signal from of the areas belonging to the dorsal attention network (i.e., IPS and FEF). These showed that on Hit trials there was first an increase of activity that, after the detection of the target, was followed by a deactivation. The time courses on Miss trials did not include any deactivation in the late phase of the trial, consistent with sustained search for the entire duration of the trial (see signal plots in Fig. 2A).

The reverse comparison (i.e., lower signal during "search" compared to the "postDet" part of the trial) revealed deactivation in the AG bilaterally. Analogous effects were found also in the precuneus, the superior medial gyrus, and left SFG (Fig. 2B, see also Table II). The time course of the BOLD signal in the AG showed that, at the beginning of the trial, the activity decreased irrespectively of trial type (i.e., both in Hit and Miss trials). However, in Hit trials this deactivation turned into activation after the target detection, while activity remained low throughout the search duration in Miss trials. Accordingly, during active search there was activation of the dorsal attention-controlling network (IPS and FEF) and the deactivation in the AG. This pattern was maintained for the entire duration of the trial when the subjects failed to detect the target (i.e., Miss trials), while in Hit trials the pattern of activation/deactivation reversed after the detection of the target.

TABLE II. Summary of the results of the whole brain analysis

Contrast/regions	MNI coordinates of the peak (mm)			z-score (peak)	p-FEW (peak)	Number of voxels
	x	y	z			
Search vs. postdetection						
L SFG/FEF	−45	−3	50	5.38	0.002	6
R SFG/FEF	51	0	56	5.48	0.004	17
L AIC	−33	27	−1	5.79	< 0.001	31
R AIC	36	24	−1	6.05	< 0.001	112
MFC	9	21	47	5.81	< 0.001	51
SC	6	−30	−4	5.77	< 0.001	91
L occipito-temporal	−24	−51	−16	7.59	< 0.001	1,422
L IPS	−27	−60	53	5.54	0.001	
R occipito-temporal	30	−48	−10	6.92	< 0.001	1,159
Postdetection vs. search						
L AG	−48	−72	41	6.27	< 0.001	311
R AG	51	−66	44	6.24	< 0.001	302
MFC	−9	60	17	5.42	< 0.001	70
L SFG	−18	33	44	5.55	< 0.001	70
R SFG	24	27	53	5.70	< 0.001	45
Precuneus	9	−54	35	5.59	< 0.001	95
Response[a]						
L SMG	−54	−48	29	5.71	< 0.001	
R SMG	54	−39	26	5.98	< 0.001	

L, left; R, right; AIC, anterior insular cortex; AG, angular gyrus; MFC, medial frontal cortex; SC, superior colliculus; SFG, superior frontal gyrus; SMG, supra marginal gyrus.
[a]For the Response contrast we report only the statistics associated with a relevant peak in the inferior parietal cortex (i.e., the SMG, see also Fig. 2C), but note that this contrast revealed activation of a large set of regions (see *Whole brain analyses: search and detection*) that are not reported here.

Next we turned to brain activity associated with the change-detection and tested for transient activations time-locked to the subjects' response on Hit trials (see Fig. 1B and Methods section). This showed the expected activation of the left motor cortex and the right cerebellum, most likely associated with the right-hand button-press responses. In addition we found activation of the SMG, the thalamus, the medial frontal cortex, and the fusiform gyrus (Fig. 2C). The detection-related activation of the SMG was anterior to the deactivation of the AG observed during search. In Figure 2C, the two effects are rendered on the same anatomical selection to show this partial segregation (change detection in cyan; "Search < postDet" in yellow; cf. Fig. 2B). The signal plots in Figure 2C show the BOLD time course in the SMG time-locked to the subject's response.

Using repeated-measures ANOVAs we assessed whether the target-position (front vs. back) and/or the viewing-condition (bmC vs. mC) affected the patterns of brain activity associated with active search and with change-detection. At the whole brain level, we found only an effect of 3D vs. 2D viewing. During search, we found activation of the superior occipital gyrus, including V3A (left hemisphere: x, y, z = −18, −93, 23, Z-value = 5.32, cluster-size = 35; right hemisphere, x, y, z = 24, −84, 23, Z-value = 5.55, cluster-size = 78; see also ROI analyses, below). Neither the dorsal

fronto-parietal regions that were found to activate during search (Fig. 2A) nor the AG and SMG in the ventral attention system (Fig. 2B,C) showed any significant effect of viewing-condition, target-position, or interactions between these two factors. Accordingly, both dorsal and ventral attention-controlling networks appeared to engage to a similar degree irrespective of the search environment (mC or bmC viewing) and the detection of changes in foreground or background.

ROI Analyses in V3A and LOC

We further investigated the effects of target-position and viewing-condition, targeting specifically visual areas processing stereoscopic cues (V3A [Brouwer et al. 2005; Neri et al., 2004; Ogawa and Macaluso, 2013; Tsao et al., 2003]) and involved in object recognition (LOC [Grill-Spector, 2003; Grill-Spector et al., 1999; Malach et al., 1995]). These were identified in each individual subject using a functional localizer. Figure 3A shows the localization of V3A and LOC in the both hemispheres, see also Table III. Within these ROIs, first we tested for the effect of search ("search" vs. "postDet" phase) and for any change-detection activity using one sample t tests (cf. whole brain analyses). Then we used repeated-measures ANOVAs to
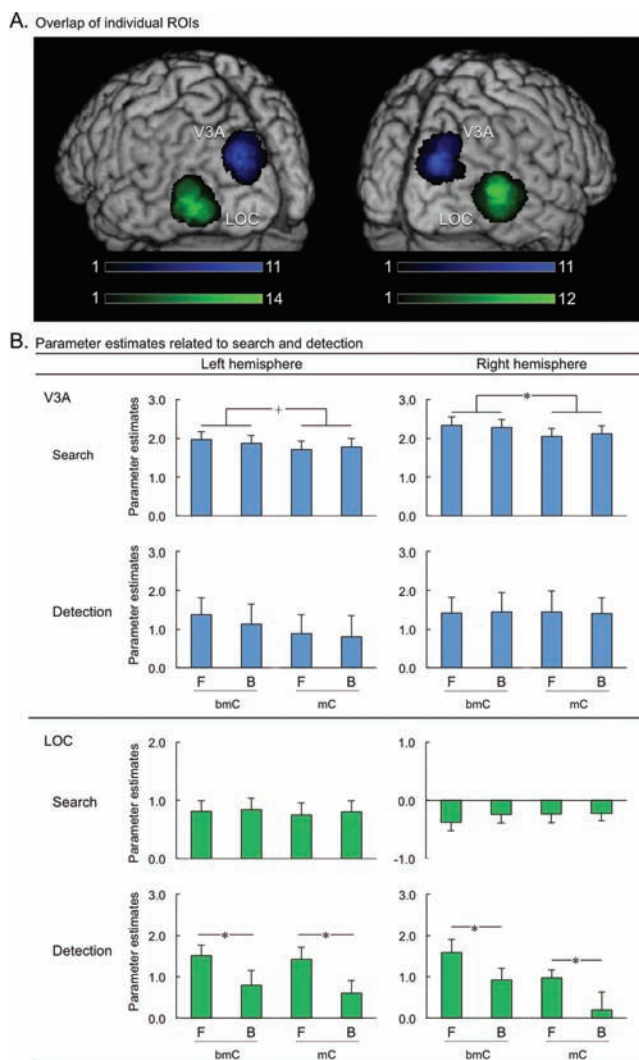
Figure 3.

The result of ROI analyses that targeted the V3A and LOC in the visual occipital cortex. **A**: The localization of the ROIs corresponding to the V3A and LOC are shown on the 3D brain template of MRIcron. Color-bars indicate the number of individual ROIs that overlapped in each voxel (cf. independent localizer scan). **B**: Parameter estimates for the search and the detection phases of the trial. During search, the V3A responded more in the bmC than mC viewing-condition. By contrast, the LOC showed a significant effect of target-position during change detection, with greater activity for the detection of foreground than background target objects. Error bars are standard error of the mean. "*": significant effects, Bonferroni corrected for the four ROIs ($P < 0.0125$); "+" the effect was significant ($P < 0.05$), but did not survive correction for multiple comparisons.

investigate any modulation according to F/B-position and mC/bmC-viewing, separately in the search and detection phase of the trial.

We found that the V3A and the LOC were activated during both search and change detection, but for the right LOC that showed a deactivation during search (Fig. 3B; and Table IV, "*t tests*"). In the right V3A, the ANOVA revealed a significant main effect of bmC vs. mC viewing during search. In the left V3A the same effect was also present but did not survive correction for multiple comparisons (Fig. 3B, plots in the top raw; and Table IV, "ANOVA Search"). During change-detection, the viewing-condition and target-position did not modulate activity in area V3A (cf. Fig. 3B, plots in the second row). By contrast, the LOC did not show any significant effect during search, but change-detection activity was larger when participants detected a change in the foreground compared with the detection of changes in the background (see Fig. 3B lower panel; and Table IV, "ANOVA Detection").

Control Analyses Modeling a Single Trial Phase

The main analyses presented above considered multiple trial-phases with variable durations (i.e., "search," "detection," "post-detection"; see Fig. 1B). Because of this, we had to construct two different models to investigate the effects of target-position and viewing-condition on search- and detection-related activity, and we compared conditions with different durations (e.g., shorter search for front- than back-targets, cf. Fig. 1C). Here, we sought to verify that our main results were not dependent on this specific, multiple-phases modeling approach. We recomputed all the first-level models now using a single-phase for each trial (fixed-duration = 6 s) and compared the different trial-types in a group-level ANOVA, see Methods section.

We tested for search-related effects by comparing "Misses and CR" vs. "Hits." This confirmed the activation the left IPS and the occipital cortex (fully significant after correction for multiple comparisons, see also Supporting Information Fig. S1A), while in left premotor cortex we found activation at an uncorrected level of significance (P-unc. < 0.001, Z-value = 3.7; x, y, $z = -57$, -3, 41; also note that the cluster was located more ventrally than that observed in the main analysis). The reverse comparison ("Hits" vs. "Misses and CR") was used to identify any detection-related effect and revealed an extensive pattern of activation including both the AG and the SMG in the inferior parietal cortex (Supporting Information Fig. S1B). Thus, unlike our main analyses with multiple phases, this single-phase approach was unable to distinguish between areas that activated when the target was detected ("Hits," in SMG; cf. Fig. 2C) versus areas that deactivated during search (i.e., negative parameter estimates for "Misses and CR," in AG; cf. Fig. 2B,C).

Finally, we examined the patterns of activation in V3A and LOC using this single-phase control approach. Considering the search-related effects (i.e., "Misses + CR vs. Hits" trials), the ROI analyses confirmed an effect of

TABLE III. Mean coordinates (± standard deviations) of the functional ROIs V3A and LOC

Left V3A			Right V3A		
x	y	z	x	y	z
−19 ± 4.7	−86 ± 3.9	29 ± 5.5	27 ± 4.2	−83 ± 5.0	29 ± 5.5
Left LOC			Right LOC		
x	y	z	x	y	z
−52 ± 4.3	−75 ± 3.9	3 ± 5.7	58 ± 3.3	−62 ± 4.6	7 ± 5.7

Coordinates (mm) refer to the standard Montreal Neurological Institute template space.

viewing-condition in bilateral V3A (bmC > mC; see Table IV, and Supporting Information Fig. S2 upper panel). The comparison designed to capture detection-related activity (i.e., "Hits vs. Misses + CR") revealed again an effect of target-position in the LOC (F > B), albeit now significant only in the left hemisphere (see Table IV, and Supporting Information Fig. S2 lower panel). Thus, the results of the control analyses that used a single-phase approach and directly compared different trial-types were overall in good agreement with our main findings modeling multiple trial phases (cf. Fig. 2 and Supporting Information Fig. S1; and Fig. 3A and Supporting Information Fig. S2).

DISCUSSION

We investigated the effects of foreground/background target-position and of binocular depth cues on change-detection within naturalistic scenes. Analyses of fixation patterns revealed that the participants gave higher search priority to front/near locations and, consistent with this, they were faster to detect color changes in foreground compared with background objects. The presence of binocular cues did not affect fixation patterns and, while overall improving search performance, the viewing-condition did not interact with the effect of target-position. During search, the functional imaging results showed activation of dorsal fronto-parietal regions (IPS and FEF) and deactivation of the AG; while the SMG activated transiently upon the detection of the changes. These patterns of activation/deactivation in dorsal and ventral attention-controlling networks were unaffected by mC/bmC viewing-condition and front/back target-position. By contrast, ROI analyses based on subject-specific localization of visual areas V3A and LOC highlighted that the V3A activated during search under stereoscopic viewing, while the LOC was associated with the detection of targets in the foreground. These findings demonstrated that target-position and binocular cues jointly contribute to search in naturalistic scenes, but do so in an additive, rather than interactive, manner via bottom-up processing in separate visual areas.

The search for changes within complex scenes requires the strategic allocation of attention, particularly so in the

flickering paradigm when the presentation of an interleaved blank image disrupts the bottom-up sensory representation of the change [Becker et al. 2000]. Here, we found that regions of the dorsal fronto-parietal attention control network activated during the search phase of the trial. These results are in agreement with previous imaging studies that showed the activation of this network during search using simple and stereotyped visual stimuli [e.g., Pessoa and Ungerleider, 2004; Shulman et al., 2003]. Our result also highlighted search-related activation in the superior colliculus and the anterior insula, consistent with previous studies that associated highlighted the role of these areas in with attention control [Bell et al., 2004; Fecteau et al., 2004; Sapir et al., 1999; Talsma et al., 2010].

However, the level of activation within these regions was unaffected by the bmC/mC viewing-condition. This seems to disagree with previous behavioral data indicating that visuo-spatial attention control takes "depth-information" into account [Marrara and Moore, 2000; Reppa et al., 2010; see also Kimura et al., 2009]. Moreover, using functional imaging, Chen et al. recently highlighted activation of the dorsal premotor cortex, when participants reoriented visuo-spatial attention between different depths rendered using stereoscopic cues [Chen et al., 2012]. This should be somewhat analogous to the shifts of spatial attention that participants here had to perform when searching for changes in the bmC condition. However, it should be noticed that in Chen's study the shifts of attention were triggered by targets shown at "invalid" locations after the presentation of predictive pre-cues. An extensive literature indicates that such "invalid" trials involve a set of complex operations, including both spatial and non-spatial processes (e.g., breaches of expectation, resetting of current attention priorities [Corbetta et al., 2008; Doricchi et al., 2010; Geng et al. 2006; Vossel et al., 2006]). By contrast, here visual search entailed primarily endogenous/strategic control, at least before the detection of the change in Hit trials (see also below).

Also the analyses of the detection phase did not reveal any significant effect of viewing-condition or target-position in the attention-controlling networks. When participants detected the change, we found that: (a) the SMG activated transiently; (b) activity decreased in the dorsal attention system; and (c) activity in the AG, which was found to deactivate during active search, returned to baseline. The interplay between dorsal and ventral parietal regions during search/target-detection has been previously described by Shulman et al. [2003]. The authors suggested that during active search the deactivation of AG reflects filtering operations that help reducing the likelihood that distracter-related signals are feed-forward from lower ("sensory") areas to higher ("associative") brain regions [see also Huettel et al., 2001; Shulman et al., 2007]. Here we replicated these patterns of activation/deactivation and found a dissociation between the SMG and the AG, when the modeling of the trial included multiple trial phases (cf. Fig. 1B, and see below). Unlike the AG, the SMG did not get deactivated significantly during search,

TABLE IV. Summary of the results of the ROI analysis

		Left V3A		Right V3A		Left LOC		Right LOC	
		z	P	z	P	z	P	z	P
Main analyses, with multiple trial phases and variable durations									
t tests	Search−PostDet	**5.09**	**<0.001**	**5.29**	**<0.001**	**2.64**	**0.008**	**2.79**	**0.005**
	Detect	*2.23*	*0.026*	**2.98**	**0.003**	**3.78**	**<0.001**	**3.42**	**<0.001**
ANOVA search	bmC vs. mC	**2.41**	**0.016** [(+)]	**3.16**	**0.002**	*1.29*	*0.199*	*1.20*	*0.229*
	F vs. B	*0.23*	*0.818*	*0.16*	*0.871*	*1.16*	*0.248*	*1.86*	*0.063*
ANOVA detection	bmC vs. mC	*1.42*	*0.156*	*0.02*	*0.983*	*0.01*	*0.996*	*2.24*	*0.025*
	F vs. B	*0.57*	*0.571*	*0.01*	*0.994*	**4.35**	**<0.001**	**2.50**	**0.012**
Control analysis, with single phase and fixed trial-durations									
Search (Miss>CR > Hit)	bmC vs. mC	**3.25**	**0.001**	**2.93**	**0.003**	*1.21*	*0.209*	*1.00*	*0.319*
	F vs. B	*2.03*	*0.043*	**2.34**	**0.018** [(+)]	*0.04*	*0.968*	*0.03*	*0.972*
Detection (Hit > Miss+CR)	bmC vs. mC	*1.05*	*0.292*	*0.22*	*0.822*	*0.53*	*0.596*	*2.26*	*0.024*
	F vs. B	*0.80*	*0.425*	*1.73*	*0.083*	**2.92**	**0.003**	*2.01*	*0.044*

(Bold) Significant effects, Bonferroni corrected ($P < 0.00125$). (+) Significant effects ($P < 0.05$), but did not survive correction of multiple comparisons. (italic) No significant effects.

but instead showed a robust transient response when the target was detected. This fits with the proposal that the anterior part of the inferior parietal lobule engages when the participants detect a behaviorally relevant stimulus [Bledowski et al., 2004; Corbetta and Shulman 2002; Downar et al., 2001; Kiehl et al., 2001; Linden et al., 1999; see also Kubit and Jack, 2013; for a recent meta-analysis linking AG with re-orienting of attention and SMG with target detection]. Here we extend these previous data demonstrating the engagement of these activation/deactivation mechanisms during search and detection in complex naturalistic scenes.

The null-finding concerning the viewing-condition and target-position in the dorsal and the ventral attention systems should be interpreted with caution. For instance other methods, such as multivariate analyses [e.g. Davis, et al., 2014; Jimura and Poldrack, 2012], may enable revealing effects of viewing-condition and/or target-position in these fronto-parietal regions. Moreover, we should note that—together with binocular disparity—the current stimuli (naturalistic pictures) included many other types of depth cues. For example, occlusions, perspective, relative size, and illumination can all provide the attention systems with information about the position of different objects in depth and, thus, contribute to the allocation of attention in 3D space. Indeed, previous fMRI studies that investigated depth perception [e.g., Shikata, et al., 2001; Taira, et al., 2001] and attention to depth information [Inui, et al., 2000] using only monocular cues, often reported activation of the posterior parietal cortex and the intra-parietal sulcus. Here, monocular cues were present in both viewing-conditions (i.e., bmC and mC) and any effect associated with these types of depth cues would be cancelled when comparing the two viewing conditions.

While not impacting on the level of activation of the fronto-parietal networks, binocular cues were found to affect the speed of the target detection (see Fig. 1C), suggesting that these cues did nonetheless contribute to attention orienting. The result of the analyses of the fixation sequences revealed that the participants searched front locations first, progressively moving towards more distant objects, and this irrespective of viewing-condition (Fig. 1E). Accordingly, we suggest that the binocular cues did not directly influence the overall search strategy but rather may have contributed to the processing / analysis of the incoming sensory signals (e.g., figure-ground segmentation).

Consistent with this proposal, our ROI analyses revealed an effect of both viewing-condition and target-position in the occipital visual cortex. Using a localizer scan, we specifically targeted the visual areas V3A and LOC, because previous studies highlighted the role of feature-selective visual areas during change detection tasks [e.g., Beck et al., 2001; Reddy et al., 2006; Schwarzkopf et al., 2010]. In the V3A, the ROI analyses confirmed an overall effect of search in bmC vs. mC, consistent with the whole brain results and previous data on the processing of binocular cues in dorsal occipital cortex [Brouwer et al. 2005; Neri et al., 2004; Ogawa and Macaluso, 2013; Tsao et al., 2003]. The activation of V3A during search was not significantly different from the activation during the postdetection phase of the trial (i.e., viewing-condition × trial-phase interaction: [bmC_Search—mC_Search] vs. [bmC_postDet—mC_post-Det]; data not shown). This indicates that the responses of V3A reflected primarily the processing of the binocular cues, rather than any interaction between the bottom-up sensory input and any top-down signal associated with the active search of the target. Nonetheless, previous studies showed that stereoscopic-viewing supports the segmentation of objects at different depths [Finlayson et al., 2013], which can facilitate the representation and selection of objects in complex 3D scenes [Lee and Saunders, 2011; Valsecchi et al. 2013]. In turns, this bottom-up effect at the sensory level could be the basis for the enhanced change-

detection performance that we found here when comparing bmC vs. mC conditions (see Fig. 1C).

Segmentation-related processes may also underlie the effect of target-position ("front > back") that we found in the LOC during the detection phase of the trial. The LOC has been previously involved in the representation of contextual information about objects' position within complex scenes [cf. Preston et al., 2013]. Furthermore, the LOC can make use of disparity signals to represent the 3D shape of objects [Preston et al., 2008], as well as processing disparity gradients that contribute to the recognition of 3D objects [Chandrasekaran et al., 2007; Cumming, 2002; Gilaie-Dotan et al., 2002]. Here, together with the fully significant effect of target-position, the right LOC also showed a marginal effect of "bmC > mC" (not significant after correction for multiple comparisons but observed both in the main and the control analyses; see Table IV) indicating a possible contribution of binocular signals also in the detection phase of the trial. The latter would be consistent with previous studies showing that targets including binocular disparity can facilitate visual search, over monocular cues only [Finlayson, et al., 2012].

In the LOC, the most robust effect (cf., also "single-phase" control analyses, in Table IV) concerned the target-position during the detection phase of the trial, which is when the subjects became aware of the target identity. This is consistent with previous studies about the role of LOC in object recognition [Grill-Spector et al., 1999; Malach et al., 1995]. The result of current study does not allow deciding whether here the activation of LOC was a "cause" of the object-recognition or, rather, it was a "consequence" of the participants becoming aware of the presence of the target-object (e.g., Tsubomi et al., 2011; Vanni et al., 1996; see also Schwarzkopf et al., 2010, who showed that TMS over the lateral occipital cortex (area LO) decreased the number of false alarms in a change detection task, possibly implying a role of this region in object-representation rather than detection). Nonetheless, our current finding of a differential effect between foreground/background conditions suggests that it is unlikely that LOC activation here reflected a general (i.e., not "depth-specific") effect of target awareness.

Together with LOC, which was targeted here using a specific functional localizer scan, also a region in the transverse occipital sulcus (TOS) has been previously associated with scene perception and the object processing [Dilks, et al., 2013; Grill-Spector and Malach, 2004; Schwarzlose, et al., 2008] and might have responded differentially to foreground/background targets in the current study. Therefore, we also run additional tests targeting TOS with the same ROI approach used for V3A and LOC (left TOS: x, y, $z = -34$, -87, 9, see Turk-Browne et al., 2012; right TOS: x, y, $z = 34$, -87, 9; i.e. symmetric in the right hemisphere). These additional tests confirmed that TOS was indeed activated in the processing of the naturalistic scenes, but there was no significant difference between the experimental condi-

tions (bmC/mC and F/B), neither in the search-phase nor in the detection-phase of the trial.

Thus, overall the ROI analyses indicated that binocular cues and object-segmentation jointly contributed to the enhanced detection of the foreground targets in 3D space, but did so in an additive, rather than interactive, manner (see also behavioral results) and via processing in different visual areas, i.e., V3A and LOC.

All our findings discussed above are based on data analyses that separated search and detection processes by modeling each experimental trial using multiple predictors (see Fig. 1B). As noted in the method section, this entailed constructing different models to investigate the effects of target-position and viewing-condition during search and detection (Models A and B, cf. Table II) and comparing trials with different durations across conditions (e.g., shorter search for front- vs. back-targets, see Fig. 1C). This approach allowed us to take into account the individual performance on a trial-by-trial basis, using trail-specific search times for the fMRI modeling. While the use of variable-duration modeling has been shown to be appropriate for the analyses of imaging data [e.g., Grinband, et al., 2008; Helfinstein, et al., 2014], it might be argued that the differences of durations could have affected our results. Accordingly, we carried out additional control analyses modeling each trial with a single predictor and directly compared different trial-types, now all with the same fixed 6-s duration (cf. Model C, in Table I). The results of these control analyses confirmed all our main results, both at the whole-brain level (see fronto-parietal networks, in Supporting Information Fig. S1) and in the prespecified visual ROIs (V3A and LOC, see Supporting Information Fig. S2 and Table IV). This demonstrated that the results of the variable-durations modeling did not arise merely because of some unwanted confound (e.g., comparisons between conditions with different durations, or unexplained variance in one or the other multiple-phases model). Moreover, it should be noticed that the single phase approach did not allow distinguishing between detection-related effects in the SMG vs. search-related deactivation in the AG (see Fig. 2B,C, and compare with Supporting Information Fig. S1B; see also discussion above).

In summary, we found that both target position in depth and binocular depth cues contributed to change-detection within naturalistic scenes. Behaviorally, participants were faster to detect changes in the foreground than in the background and when the visual input included binocular cues. However, the two factors did not interact, suggesting additive mechanisms. The imaging results associated the effects of binocular viewing and of foreground target position with activity in the visual cortex. The two factors influenced activity in separate brain regions, with bmC viewing activating area V3A during search and the detection of targets in the foreground activating area LOC. We suggest that these areas contribute to change-detection performance primarily via by bottom-up processes, including binocular representations in the V3A and object

segmentation in the LOC. By contrast, target position and binocular cues did not affect activity in higher order frontoparietal attention-controlling networks. Active search activated the dorsal attention network (FEF and IPS) and deactivated the AG, while target detection activated the SMG in the ventral attention system. We conclude that bottom-up processing within sensory regions, rather than strategic control in fronto-parietal attention networks, mediates the influence of target-position and binocular-cues during search and detection in 3D real-world scenes.

ACKNOWLEDGEMENT

REFERENCES

Arnold DH, Birt A, Wallis TS (2008): Perceived size and spatial coding. J Neurosci 28:5954–5958.

Beck DM, Rees G, Frith CD, Lavie N (2001): Neural correlates of change detection and change blindness. Nat Neurosci 4:645–650.

Becker MW, Pashler H, Anstis SM (2000): The role of iconic memory in change-detection tasks. Perception 29:273–286.

Bell AH, Fecteau JH, Munoz DP (2004): Using auditory and visual stimuli to investigate the behavioral and neuronal consequences of reflexive covert orienting. J Neurophysiol 91:2172–2184.

Bledowski C, Prvulovic D, Goebel R, Zanella FE, Linden DE (2004): Attentional systems in target and distractor processing: A combined ERP and fMRI study. Neuroimage 22:530–540.

Brainard DH (1997): The psychophysics toolbox. Spat Vis 10:433–436.

Brett M, Anton J-L, Valabregue R, Poline J-B (2002): Region of interest analysis using an SPM toolbox. Organization for Human Brain Mapping annual meeting, Sendai, Japan.

Brouwer GJ, van Ee R, Schwarzbach J (2005): Activation in visual cortex correlates with the awareness of stereoscopic depth. J Neurosci 25:10403–10413.

Cavanaugh J, Wurtz RH (2004): Subcortical modulation of attention counters change blindness. J Neurosci 24:11236–11243.

Chandrasekaran C, Canon V, Dahmen JC, Kourtzi Z, Welchman AE (2007): Neural correlates of disparity-defined shape discrimination in the human brain. J Neurophysiol 97:1553–1565.

Chau AW, Yeh YY (1995): Segregation by color and stereoscopic depth in three-dimensional visual space. Percept Psychophys 57:1032–1044.

Chen Q, Weidner R, Vossel S, Weiss PH, Fink GR (2012): Neural mechanisms of attentional reorienting in three-dimensional space. J Neurosci 32:13352–13362.

Corbetta M, Shulman G (2002): Control of goal-directed and stimulus-driven attention in the brain. Nat Rev Neurosci 3: 201–215.

Corbetta M, Shulman GL (1998): Human cortical mechanisms of visual attention during orienting and search. Philos Trans R Soc Lond B Biol Sci 353:1353–1362.

Corbetta M, Patel G, Shulman GL (2008): The reorienting system of the human brain: From environment to theory of mind. Neuron 58:306–324.

Cottereau BR, McKee SP, Ales JM, Norcia AM (2011): Disparity-tuned population responses from human visual cortex. J Neurosci 31:954–965.

Cumming B (2002): Stereopsis: Where depth is seen. Curr Biol 12: R93–R95.

Davis T, LaRocque KF, Mumford JA, Norman KA, Wagner AD, Poldrack RA (2014): What do differences between multi-voxel and univariate analysis mean? How subject-, voxel-, and trial-level variance impact fMRI analysis. Neuroimage 97:271–283.

Dilks DD, Julian JB, Paunov AM, Kanwisher N (2013): The occipital place area is causally and selectively involved in scene perception. J Neurosci 33:1331–1336.

Doricchi F, Macci E, Silvetti M, Macaluso E (2010): Neural correlates of the spatial and expectancy components of endogenous and stimulus-driven orienting of attention in the Posner task. Cereb Cortex 20:1574–1585.

Downar J, Crawley AP, Mikulis DJ, Davis KD (2001): The effect of task relevance on the cortical response to changes in visual and auditory stimuli: An event-related fMRI study. Neuroimage 14:1256–1267.

Fairhall SL, Macaluso E (2009): Spatial attention can modulate audiovisual integration at multiple cortical and subcortical sites. Eur J Neurosci 29:1247–1257.

Fecteau JH, Bell AH, Munoz DP (2004): Neural correlates of the automatic and goal-driven biases in orienting spatial attention. J Neurophysiol 92:1728–1737.

Finlayson NJ, Remington RW, Grove PM (2012): The role of presentation method and depth singletons in visual search for objects moving in depth. J Vis 12:13.

Finlayson NJ, Remington RW, Retell JD, Grove PM (2013): Segmentation by depth does not always facilitate visual search. J Vis 13:11.

Friston KJ, Glaser DE, Henson RN, Kiebel S, Phillips C, Ashburner J. (2002): Classical and Bayesian inference in neuroimaging: applications. Neuroimage 16:484–512.

Geng JJ, Eger E, Ruff CC, Kristjánsson A, Rotshtein P, Driver J (2006): On-line attentional selection from competing stimuli in opposite visual fields: Effects on human visual cortex and control processes. J Neurophysiol 96:2601–2612.

Gilaie-Dotan S, Ullman S, Kushnir T, Malach R (2002): Shape-selective stereo processing in human object-related visual areas. Hum Brain Mapp 15:67–79.

Grill-Spector K (2003): The neural basis of object perception. Curr Opin Neurobiol 13:159–166.

Grill-Spector K, Kushnir T, Edelman S, Avidan G, Itzchak Y, Malach R (1999): Differential processing of objects under various viewing conditions in the human lateral occipital complex. Neuron 24:187–203.

Grill-Spector K, Malach R (2004): The human visual cortex. Annu Rev Neurosci 27:649–77.

Grinband J, Wager TD, Lindquist M, Ferrera VP, Hirsch J (2008): Detection of time-varying signals in event-related fMRI designs. Neuroimage 43:509–520.

Helfinstein SM, Schonberg T, Congdon E, Karlsgodt KH, Mumford JA, Sabb FW, Cannon TD, London ED, Bilder RM, Poldrack RA (2014): Predicting risky choices from brain activity patterns. Proc Natl Acad Sci USA 111:2470–2475.

Huettel SA, Güzeldere G, McCarthy G (2001): Dissociating the neural mechanisms of visual attention in change detection using functional MRI. J Cogn Neurosci 13:1006–1018.

Inui T, Tanaka S, Okada T, Nishizawa S, Katayama M, Konishi J (2000): Neural substrates for depth perception of the Necker

cube: A functional magnetic resonance imaging study in human subjects. Neurosci Lett 282:145–148.

Jimura K, Poldrack RA (2012): Analyses of regional-average activation and multivoxel pattern information tell complementary stories. Neuropsychologia 50:544–552.

Kiehl KA, Laurens KR, Duty TL, Forster BB, Liddle PF (2001): Neural sources involved in auditory target detection and novelty processing: An event-related fMRI study. Psychophysiology 38:133–142.

Kimura T, Miura T, Shinohara K (2009): Effect of fixation point distances on allocation of attention in real three-dimensional space. Percept Mot Skills 109:327–337.

Kubit B, Jack AI (2013): Rethinking the role of the rTPJ in attention and social cognition in light of the opposing domains hypothesis: Findings from an ALE-based meta-analysis and resting-state functional connectivity. Front Hum Neurosci 7:323.

Lathrop WB, Bridgeman B, Tseng P (2011): Perception in the absence of attention: perceptual processing in the Roelofs effect during inattentional blindness. Perception 40:1104–1119.

Lee YL, Saunders JA (2011): Stereo improves 3D shape discrimination even when rich monocular shape cues are available. J Vis 11:6.

Linden DE, Prvulovic D, Formicaio E, Völlinger M, Zanella FE, Goebel R, Dierks T (1999): The functional neuroanatomy of target detection: An fMRI study of visual and auditory oddball tasks. Cereb Cortex 9:815–823.

Liu C, Yuen J (2011): SIFT Flow: Dense correspondence across scenes and its applications. IEEE T Pattern Anal 33:978–994.

Lowe DG (1999): Object recognition from local scale-invariant features. Proceedings of the International Conference on Computer Vision, Corfu, Greece: IEEE Computer Society, Washington, DC, USA, pp. 1150–1157.

Malach R, Reppas JB, Benson RR, Kwong KK, Jiang H, Kennedy WA, Ledden PJ, Brady TJ, Rosen BR, Tootell RB (1995): Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. Proc Natl Acad Sci USA 92:8135–8139.

Marrara MT, Moore CM (2000): Role of perceptual organization while attending in depth. Percept Psychophys 62:786–799.

Mazza V, Turatto M, Umiltà C (2005): Foreground-background segmentation and attention: A change blindness study. Psychol Res 69:201–210.

Nakayama K, Silverman GH (1986): Serial and parallel processing of visual feature conjunctions. Nature 320:264–265.

Neri P, Bridge H, Heeger DJ (2004): Stereoscopic processing of absolute and relative disparity in human visual cortex. J Neurophysiol 92:1880–1891.

Ogawa A, Macaluso E (2013): Audio-visual interactions for motion perception in depth modulate activity in visual area V3A. Neuroimage 71C:158–167.

Ogawa A, Bordier C, Macaluso E (2013): Audio-visual perception of 3D cinematography: An fMRI study using condition-based and computation-based analyses. PLOS ONE 8:e76003.

Ozkan K, Braunstein ML (2010): Change detection for objects on surfaces slanted in depth. J Vis 10:12.

Parkhurst D, Law K, Niebur E (2002): Modeling the role of salience in the allocation of overt visual attention. Vision Res 42:107–123.

Pessoa L, Ungerleider LG (2004): Neural correlates of change detection and change blindness in a working memory task. Cereb Cortex 14:511–520.

Pomplun M, Garaas TW, Carrasco M (2013): The effects of task difficulty on visual search strategy in virtual 3D displays. J Vis 13:24.

Preston TJ, Li S, Kourtzi Z, Welchman AE (2008): Multivoxel pattern selectivity for perceptually relevant binocular disparities in the human brain. J Neurosci 28:11315–11327.

Preston TJ, Guo F, Das K, Giesbrecht B, Eckstein MP (2013): Neural representations of contextual guidance in visual search of real-world scenes. J Neurosci 33:7846–7855.

Reddy L, Quiroga RQ, Wilken P, Koch C, Fried I (2006): A single-neuron correlate of change detection and change blindness in the human medial temporal lobe. Curr Biol 16:2066–2072.

Rensink RA (2002): Change detection. Annu Rev Psychol 53:245–277.

Rensink RA, O'Regan JK, Clark JJ (1997): To see or not to see: The need for attention to perceive changes in scenes. Psychol Sci 8:368–373.

Reppa I, Fougnie D, Schmidt WC (2010): How does attention spread across objects oriented in depth? Atten Percept Psychophys 72:912–925.

Sapir A, Soroker N, Berger A, Henik A (1999): Inhibition of return in spatial attention: Direct evidence for collicular generation. Nat Neurosci 2:1053–1054.

Schwarzkopf DS, Silvanto J, Gilaie-Dotan S, Rees G (2010): Investigating object representations during change detection in human extrastriate cortex. Eur J Neurosci 32:1780–1787.

Schwarzlose RF, Swisher JD, Dang S, Kanwisher N (2008): The distribution of category and location information across object-selective regions in human visual cortex. Proc Natl Acad Sci USA 105:4447–4452.

Seegmiller JK, Watson JM, Strayer DL (2011): Individual differences in susceptibility to inattentional blindness. J Exp Psychol Learn Mem Cogn, 37:785–791.

Shikata E, Hamzei F, Glauche V, Knab R, Dettmers C, Weiller C, Büchel C (2001): Surface orientation discrimination activates caudal and anterior intraparietal sulcus in humans: An event-related fMRI study. J Neurophysiol 85:1309–1314.

Shulman GL, McAvoy MP, Cowan MC, Astafiev SV, Tansy AP, d'Avossa G, Corbetta M (2003): Quantitative analysis of attention and detection signals during visual search. J Neurophysiol 90:3384–3397.

Shulman GL, Astafiev SV, McAvoy MP, d'Avossa G, Corbetta M (2007): Right TPJ deactivation during visual search: Functional significance and support for a filter hypothesis. Cereb Cortex 17:2625–2633.

Shulman GL, Astafiev SV, Franke D, Pope DL, Snyder AZ, McAvoy MP, Corbetta M (2009): Interaction of stimulus-driven reorienting and expectation in ventral and dorsal frontoparietal and basal ganglia-cortical networks. J Neurosci 29:4392–4407.

Taira M, Nose I, Inoue K, Tsutsui K (2001): Cortical areas related to attention to 3D surface structures based on shading: an fMRI study. Neuroimage 14:959–966.

Talsma D, Coe B, Munoz DP, Theeuwes J (2010): Brain structures involved in visual search in the presence and absence of color singletons. J Cogn Neurosci 22:761–774.

Tsao DY, Vanduffel W, Sasaki Y, Fize D, Knutsen TA, Mandeville JB, Wald LL, Dale AM, Rosen BR, Van Essen DC, Livingstone MS, Orban GA, Tootell RB (2003): Stereopsis activates V3A and caudal intraparietal areas in macaques and humans. Neuron 39:555–568.

Tsubomi H, Ikeda T, Hanakawa T, Hirose N, Fukuyama H, Osaka N (2012): Dissociable neural activations of conscious visibility and attention. J Cogn Neurosci 24:496–506.

Turatto M, Angrilli A, Mazza V, Umiltà C, Driver J (2002): Looking without seeing the background change: Electrophysiological correlates of change detection versus change blindness. Cognition 84:B1–10.

Turk-Browne NB, Simon MG, Sederberg PB (2012): Scene representations in parahippocampal cortex depend on temporal context. J Neurosci 32:7202–7207.

Valsecchi M, Caziot B, Backus BT, Gegenfurtner KR (2013): The role of binocular disparity in rapid scene and pattern recognition. i-Perception 4:122–136.

Vanni S, Revonsuo A, Saarinen J, Hari R (1996): Visual awareness of objects correlates with activity of right occipital cortex. Neuroreport 8:183–186.

Vossel S, Thiel CM, Fink GR (2006): Cue validity modulates the neural correlates of covert endogenous orienting of attention in parietal and frontal cortex. Neuroimage 32:1257–1264.